Exp Brain Res (2004) 157: 431–441
DOI 10.1007/s00221-004-1857-6

RESEARCH ARTICLES

Catherine M. Warrier · Krista L. Johnson · Erin A. Hayes · Trent Nicol · Nina Kraus

Learning impaired children exhibit timing deficits and training-related improvements in auditory cortical responses to speech in noise

Received: 1 July 2003 / Accepted: 20 January 2004 / Published online: 6 April 2004

Abstract The physiological mechanisms that contribute to abnormal encoding of speech in children with learning problems are yet to be well understood. Furthermore, speech perception problems appear to be particularly exacerbated by background noise in this population. This study compared speech-evoked cortical responses recorded in a noisy background to those recorded in quiet in normal children (NL) and children with learning problems (LP). Timing differences between responses recorded in quiet and in background noise were assessed by cross-correlating the responses with each other. Overall response magnitude was measured with root-mean-square (RMS) amplitude. Cross-correlation scores indicated that 23% of LP children exhibited cortical neural timing abnormalities such that their neurophysiological representation of speech sounds became distorted in the presence of background noise. The latency of the N2 response in noise was isolated as being the root of this distortion. RMS amplitudes in these children did not differ from NL children, indicating that this result was not due to a difference in response magnitude. LP children who participated in a commercial auditory training program and exhibited improved cortical timing also showed improvements in phonological perception. Consequently, auditory pathway timing deficits can be objectively observed in LP children, and auditory training can diminish these deficits.

Keywords Auditory evoked potentials · Cross-correlation · Learning disorders · Neural timing · Speech perception · Auditory training

C. M. Warrier (✉) · K. L. Johnson · E. A. Hayes · T. Nicol · N. Kraus
Auditory Neuroscience Laboratory, Northwestern University, Frances Searle Building, 2240 Campus Drive, Evanston, IL 60208, USA
e-mail: cwarrier@northwestern.edu
Tel.: +1-847-491-2465
Fax: +1-847-491-2523

C. M. Warrier · E. A. Hayes · N. Kraus
Institute for Neuroscience, Northwestern University, Evanston, IL 60208, USA

C. M. Warrier · K. L. Johnson · E. A. Hayes · T. Nicol · N. Kraus
Department of Communication Sciences, Northwestern University, Evanston, IL 60208, USA

N. Kraus
Department of Neurobiology and Physiology, Northwestern University, Evanston, IL 60208, USA

N. Kraus
Department of Otolaryngology, Northwestern University, Evanston, IL 60208, USA

Introduction

Whether it is a classroom filled with children laughing and talking, a busy street corner, or a kitchen humming with appliances, normal everyday listening environments are filled with noise. Accurately perceiving speech in such noise-filled environments challenges the auditory system's ability to process the complex acoustic elements of speech. The portions of the speech signal that are most susceptible to perceptual disruption by masking noise are those that are low intensity and/or aperiodic. As the acoustic characteristics distinguishing consonants are typically low in intensity and aperiodic, consonant identification is typically more challenged by the presence of background noise than vowel identification. Consonants provide most of the information needed for word meaning; shorthand writing often includes consonants but drops vowel symbols. Consequently, proper neural encoding of the acoustic signal giving rise to consonant percepts is paramount to understanding speech in the presence of background noise.

Many people are able to perceive speech relatively well in poor listening environments, but some listeners demonstrate a great deal of difficulty. Individuals with learning impairments, including language-based learning problems and attention deficit hyperactivity disorder, form

part of this latter group. Previous research has shown that auditory perceptual deficits contribute to learning problems in some children (Tallal and Piercy 1974; Tallal 1980; Elliott et al. 1989; Kraus et al. 1996; Stark and Heinz 1996; Alain et al. 1998; Bradlow et al. 1999). These deficits are exacerbated when background noise is introduced (Nabelek and Pickett 1974; Elliott 1979; Bellis 1996; Chermak and Musiek 1997; Bradlow et al. 2003). Furthermore, it is believed that some children with learning problems exhibit fundamental differences in the way they physiologically encode complex acoustic information, such as speech, compared to normally developing children (Kraus et al. 1996; Cunningham et al. 2001; King et al. 2002; Wible et al. 2002).

Auditory evoked responses can be used to identify neurophysiological differences between the way normal children (NL) and children with learning problems (LP) encode sounds. For example, a study by Cunningham and colleagues (2001) found similar fundamental sensory representation of speech between NL and LP populations when evoked responses were recorded in quiet, but LP children demonstrated neurophysiological abnormalities at both cortical and subcortical levels when the speech was presented in background noise. In that study, noise degraded subcortical encoding of both transient and periodic stimulus features in the auditory brainstem and frequency following responses in LP children. Additionally, cortical responses in noise showed a dramatic amplitude decrease in LP children with respect to NL children in the P2N2 complex occurring between 150–250 ms post-stimulus onset.

Motivated by these findings, the present study aims to further describe the physiological mechanisms that contribute to the encoding deficits of LP children when speech is presented in noise. We recorded cortical evoked responses to a speech stimulus in NL and LP children in quiet and in the presence of background noise. Although cortical responses are inherently variable between subjects, even within the normal population, comparing a subject's response in noise to his or her own response in quiet reduces variability and makes it possible to assess the degree to which that individual's response is affected by noise. A technique of response cross-correlation (described below) was used in individual subjects to evaluate changes in the morphology and timing of cortical auditory evoked waveforms when noise was added to the stimulus. Response correlation has been used previously to evaluate the timing mechanisms of auditory evoked responses (Ponton et al. 2001; Wible et al. 2002; Hayes et al. 2003). This type of comparison is advantageous because it is a sensitive measure of cortical timing across a specified range of the response, in this case 150 ms. Timing differences were contrasted with overall response magnitude as measured by root mean square (RMS) amplitude. Both cross-correlation and RMS amplitude are computationally objective measures that avoid the relatively subjective process of choosing peaks. Amplitudes and latencies of prominent peaks within the specified time range were assessed to inform the correlation and RMS amplitude results.

A previous study assessing behavioral and neurophysiological changes due to commercial auditory training showed that such training can affect cortical responses to speech sounds collected in the presence of background noise (Hayes et al. 2003). This study also reported improved phonological processing resulting from the same training. This improvement was seen on the Auditory Processing score, a composite score of the Woodcock-Johnson–Revised test battery that includes scores from two tests of phonemic awareness: the Sound Blending and Incomplete Words subtests (Woodcock and Johnson 1989). Improvements were observed on both of these tests. Improvements in Sound Blending were significantly greater than improvements seen in an untrained control group, and a similar group difference of gains in Incomplete Words approached significance. Sound Blending requires subjects to synthesize phonemes to form a word, and Incomplete Words requires subjects to complete words that are missing phonemes. Overall training effects were assessed in the Hayes et al. (2003) study, but the relationship between neurophysiological and perceptual changes was not examined. The present study includes a second experiment to determine whether training-related cortical and perceptual changes co-occur, or if these changes are independent of one another. We recorded children's auditory evoked potentials to a speech sound presented in quiet and in the presence of background noise both before and after participation in an auditory training program. We were then able to compare these responses in order to assess training-related changes. We also tested these subjects on speech perception measures before and after training and related changes on these measures to neurophysiological changes.

Experiment 1

Materials and methods

Subjects

One hundred twelve children between the ages of 8 and 13 years participated in this study. All children had normal pure tone thresholds as assessed by screening at 20 dB HL for octaves from 500 to 4,000 Hz. Eighty children (25 females, 55 males) were classified as having learning problems (LP) based on diagnosis by an outside professional and an in-house psychoeducational test battery. Thirty-two children (14 females, 18 males) were classified as normal (NL) based on subject history and their ability to score average or above on all subtests of an in-house psychoeducational test battery. The psychoeducational test battery, given to all children, consisted of standardized measures of learning and academic achievement taken from subtests of the Woodcock-Johnson Revised (Woodcock and Johnson 1989): Listening Comprehension, Memory for Words, Cross-out, Word Attack, and Auditory Processing (a composite score based on Incomplete Words and Sound Blending subtests). Reading and spelling were evaluated using subtests from the WRAT-3 Test Battery (Wilkinson 1993). The Brief Cognitive Scale (an IQ estimate) was required to be at least 85 for all subjects to eliminate the confound of a low IQ (Woodcock and Johnson 1977). Institutional Review Board approval for this study was

obtained from Northwestern University, ensuring that all methods were performed in accordance with the ethical guidelines set forth in the Declaration of Helsinki. Written and oral informed assent/consent was given by each child and his/her parent or guardian. The children were compensated for their participation.

Stimuli

A Klatt formant synthesizer was used to synthesize a 40 ms speech syllable /da/ at a sampling rate of 10 kHz (Klatt 1980). It consisted of five formants with an onset burst frication during the first 10 ms at F_3, F_4, and F_5 (Fig. 1).

Neurophysiological procedures

Stimuli were presented by a PC-based stimulus delivery system (NeuroScan Stim) that controlled time of delivery and stimulus intensity. The stimuli were presented to the right ear through Etymotic ER-3 earphones at 80 dB SPL. For noise conditions, continuous broadband white noise was generated by a Bio-logic Navigator system and mixed with the /da/ stimulus (0 dB signal-to-noise ratio). Each condition included 1,000 stimulus presentations with 590 ms interstimulus intervals. A PC-based evoked potentials averaging system (NeuroScan Acquire) was triggered at every stimulus onset.

Cortical responses were collected in the quiet and noise conditions with a recording window of 575 ms including a 75 ms prestimulus period. Responses were sampled at 2 kHz, and bandpass filtered online from 0.05 to 100 Hz. Recordings were made with silver-silver chloride electrodes (impedance <5 kΩ). Electrical responses were recorded from a central vertex electrode (Cz). A nose electrode served as reference, and a forehead electrode served as ground. A bipolar supraorbital-to-lateral canthus electrode monitored eye blinks. Sweeps with levels exceeding ±100 μV were rejected online.

To ensure subject cooperation, all subjects watched videotaped programs such as movies or cartoons with the sound presented at a low level (<40 dB SPL) in sound field, and heard through the unoccluded left ear. They were instructed to stay awake and attend to the video rather than to the stimulus.

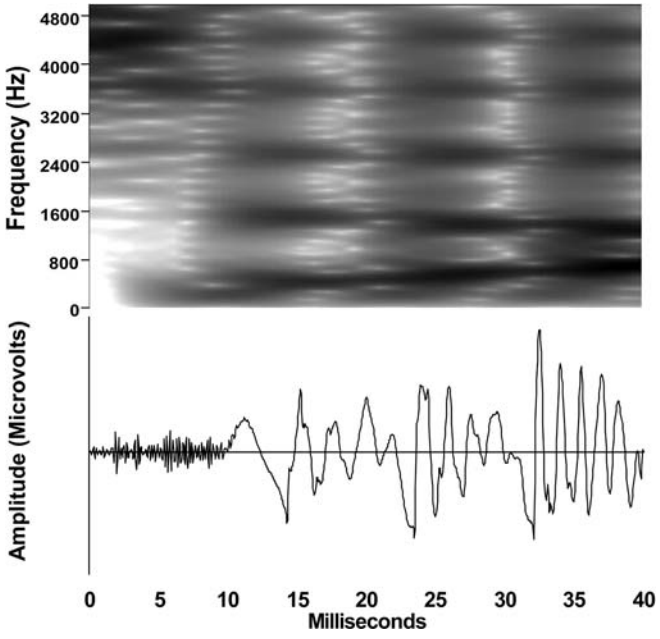

Fig. 1 A synthesized /da/ syllable containing five formants with burst enhancement in F4 and F5. *Top* spectrogram; *bottom* time waveform

Analysis

Waveforms recorded in quiet and in background noise were analyzed using three measurement techniques: cross-correlation, root-mean-square (RMS) amplitude, and measurement of peak amplitudes and latencies. Correlation and RMS amplitude are complementary analysis methods, each operating over the duration of a specified period. Cross-correlating two waveforms measures how well the events of those waveforms correspond in time irrespective of their magnitude, and RMS amplitude measures the magnitude of activity present irrespective of timing. The latency range from 100 to 350 ms post-stimulus onset was investigated. This range included the P2N2 complex, a robust response that was identifiable in most subjects. Prior to calculation of RMS amplitude, waveforms were shifted to a baseline of zero over the 100–350 ms range in order to remove DC drift.

The cross-correlation of responses involves shifting in time an individual's waveform recorded in noise with respect to that recorded in quiet until a "best fit" is established. The addition of noise is known to delay the latency of events within a waveform (Whiting et al. 1998; Martin et al. 1999). Thus, time-shifting the waveform recorded in noise allows an assessment of the stability of the relative timing between morphological features of the two waveforms. Therefore, this measure can be interpreted as a measure of the preservation of response timing between the quiet and noise conditions. In order to obtain a high correlation, after an appropriate shift, the morphological features of the two waves must occur at similar points. A high correlation indicates that the addition of background noise to the stimulus had a minimal effect on response timing, whereas a low value indicates a larger effect. The maximum correlation between waveforms (Pearson r value), as well as the amount of shift required to achieve it (in ms), was recorded for each child (Fig. 2). Automatic cross-correlations were performed in a PC-based statistical system (NeuroScan Stats). Verification that the correct correlation value was chosen was accomplished by visual estimation of the delay between quiet and noise responses. To transform correlation values to a normal distribution for the purposes of statistical analysis, r values were converted to z' scores using Fisher's transformation. Although all statistics were performed on these z' scores, raw r values are used in figures and text for ease of interpretation.

Amplitudes and latencies of peaks contained within the 100–350 ms window (N1, P2, and N2) were assessed for all waveforms recorded in quiet and in the presence of background noise to inform the correlation and RMS amplitude data. Three subjects' waveforms (one from each group: Control, LP-In, and LP-Out; defined below) did not exhibit identifiable peaks in this time range and were therefore not included in this analysis. In addition, some waveforms did not exhibit all three peaks. Therefore, when performing statistics involving specific peaks, if those peaks were not seen in a particular waveform, that waveform was not included in the analysis.

Results

Response correlations

Inspection of NL and LP cross-correlation scores showed that there were a number of LP subjects who fell below the worst NL score (Fig. 3). In order to evaluate these poor-correlating subjects as a separate group, we defined poor correlation scores as being below the mean −1 SD of the NL group's correlation scores (mean: r=0.79; mean −1 SD: r=0.61). Twenty-three percent (n=18) of the LP

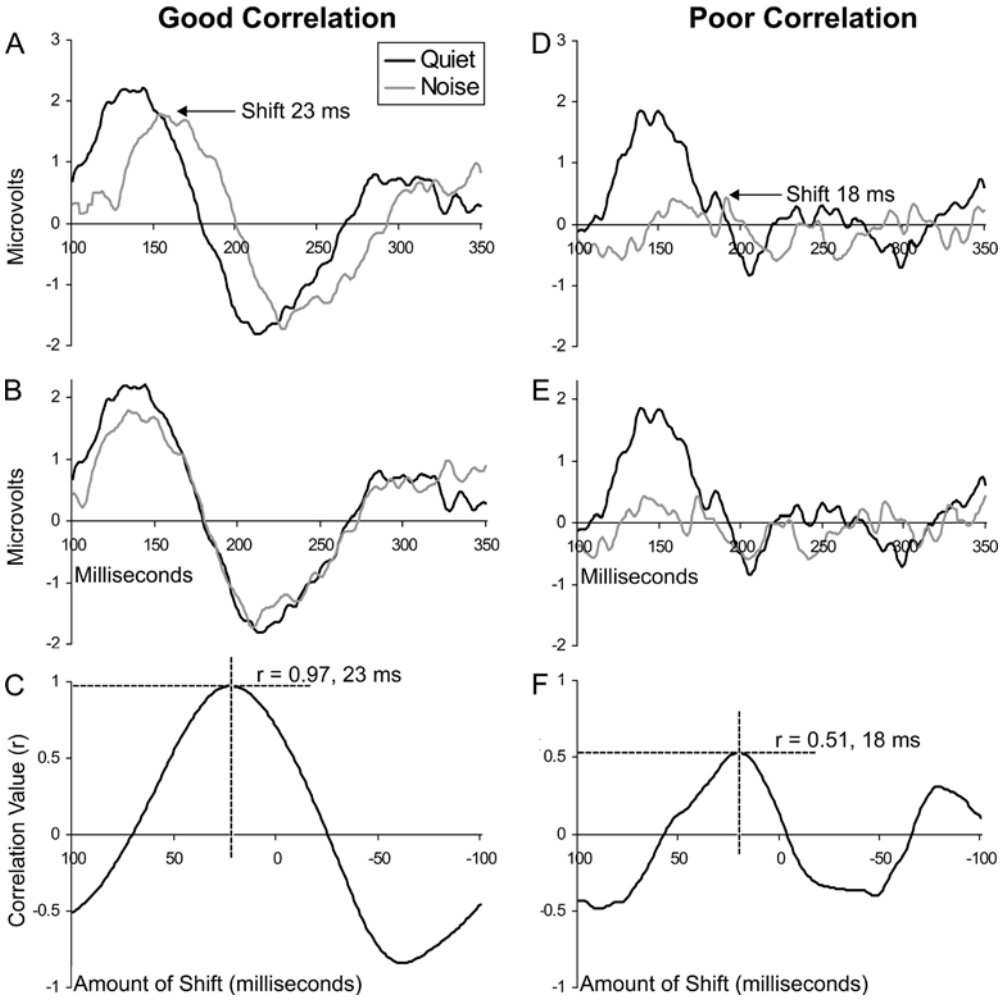

Fig. 2A–F Schematic of the quiet-to-noise cross-correlation technique used in this study. A subject with good correlation **(A–C)** has similar wave morphology in both quiet (*black line*) and in background noise (*gray line*). This subject's original waves are shown in **A**. The noise response is shifted relative to the quiet response to a degree that allows the best fit, in this case, 23 ms. The shifted noise waveform is shown with the original quiet waveform in **B**. The cross-correlation function is shown in **C**, where each point along the wave is a correlation value at a given latency shift. The maximum correlation value here is 0.97 with a 23 ms shift. **D–F** show similar figures for a subject with poor quiet-to-noise correlation

subjects exhibited poor correlations. The proportion of LP children falling below the normal range exceeded that found in the normal population (binomial test, p=0.05). We therefore divided our LP group in two based on correlation scores. The 18 children with poor correlations became our LP-Out group, and the remaining LP children became our LP-In group. The time shift required for maximum correlation did not differ between groups [Mean (SD): NL 7.5 (19.7) ms, LP-In 13.3 (13.6) ms, LP-Out 10.2 (14.2) ms; ANOVA F=1.54, p=0.22]. The male/female ratio did not differ across the three subject groups as assessed by Pearson's chi-square test (% male: NL=56%, LP-In=66%, LP-Out=84%; χ^2=2.58, p>0.2). Each group's grand averaged waveforms of responses recorded in quiet and in noise are shown in Fig. 4.

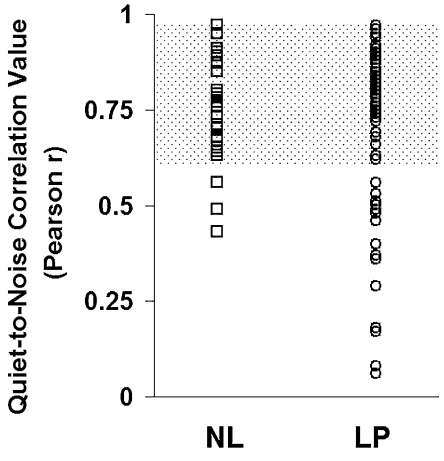

Fig. 3 Individual correlation values for normal and learning impaired children. The highlighted normal range was determined by the mean ± 1 SD of the NL group's correlation values. Notice the larger proportion of LP subjects compared to NL subjects falling below this range

Response magnitude

The effect of noise on RMS amplitudes was assessed using a repeated measures ANOVA with three levels of group as a between subjects factor and two levels of noise condition as a within subjects factor. The absence of an interaction effect indicated that the addition of noise decreased the response magnitudes of all subject groups to the same degree (F=0.20, p>0.8) (Fig. 5). RMS amplitude of the LP-Out group did not differ from that of the NL group, indicating that differences in response timing between

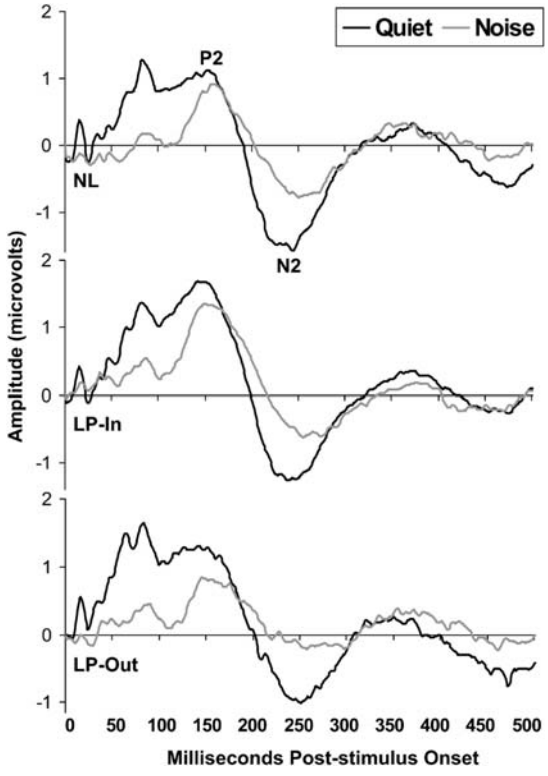

Fig. 4 Averaged cortical responses elicited by /da/ in quiet (*black line*) and in background noise (*gray line*) for three subject groups. The noise response in the normal and LP-In groups is relatively well preserved, whereas the noise response for the LP-Out group appears degraded

these two groups could not be attributed to differences in overall response magnitude.

The addition of noise is widely known to diminish the magnitude of auditory evoked responses (Martin et al. 1997; Whiting et al. 1998). The expected main effect of noise condition was identified in which RMS amplitude was larger in quiet than in noise in all groups (F=73.459, p<0.001). A significant main effect of group was also found (F=5.38, p<0.01). Neither the LP-In nor the LP-Out group differed from NL, although they differed from each other in that the LP-Out group exhibited smaller RMS amplitude than the LP-In group (post hoc Tukey HSD, p<0.01).

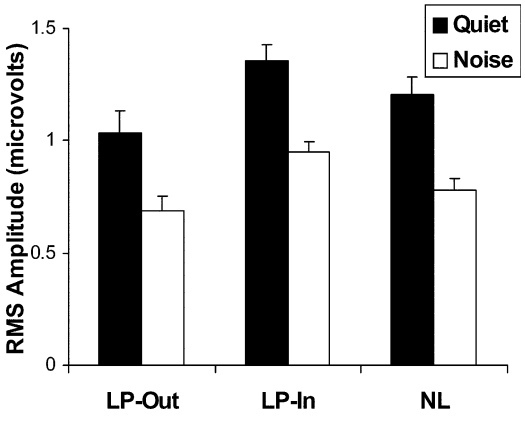

Fig. 5 RMS amplitude in quiet (*shaded bars*) and noise (*white bars*). The addition of noise decreased the response magnitude of all three groups to the same degree

Waveform peaks

Visual inspection of grand averaged waveforms shows that, in quiet, all groups had similar wave morphology; that is, the P2N2 complex was similar in shape and size (Fig. 4). However, the LP-Out group's response in the noise condition appears highly degraded relative to that of the NL and LP-In subjects, particularly in the N2 region. This degradation is not due to a disproportionate number of children not exhibiting the N2 peak in the LP-Out group: most subjects in all three groups retained robust N2 peaks in the noise condition (NL 87%, LP-In 97%, LP-Out 89%). Careful inspection of the amplitudes and latencies of individual waveform peaks suggested a link between low correlations and the effect of background noise on the N2 latency. A repeated measures ANOVA with two levels of N2 latency as a within subjects factor (quiet and noise conditions) and three levels of group as a between subjects factor (N=27 NL, 60 LP-In, and 16 LP-Out) revealed that the change in N2 latency between quiet and noise conditions differed between groups (F=3.00, p=0.054). The mean N2 latency of NL and LP-In groups was approximately 20 ms later in response to a stimulus presented in noise as opposed to in quiet, but the LP-Out group's N2 latencies did not change. Post hoc tests confirmed that although all groups had similar N2 latencies in the quiet condition (Tukey HSD, p>0.05 all comparisons), the LP-Out group's mean N2 latency differed from that of the NL and LP-In groups in the noise condition (Tukey HSD; NL vs. LP-Out: p<0.05, LP-In vs. LP-Out: p<0.05, NL vs. LP-In: p>0.05). No group differences in the latencies of earlier peaks were evident.

The interpeak amplitude of the P2N2 complex informed the RMS amplitude data. A repeated measures ANOVA with two levels of P2N2 amplitude as a within subjects factor (quiet and noise conditions) and three levels of group as a between subjects factor (N=24 NL, 52 LP-In, and 11 LP-Out) was applied to the data. The absence of an interaction effect (F=1.13, p>0.3) indicated that the addition of noise affected this response equally across groups. A main effect of noise (F=31.81, p<0.001) pointed to the decrease in amplitude with the addition of noise. A main effect of group was also found (F=4.45, p=0.015). As in the RMS amplitude data, post hoc testing revealed that neither LP group differed from NL on this measure (Tukey HSD, p>0.05 both comparisons). However, they differed from themselves such that the P2N2 amplitude in the LP-Out group was smaller than that of the LP-In group (Tukey HSD, p<0.05). Whereas RMS amplitude takes the magnitude of the entire waveform into account, the P2N2 amplitude measures a more focal region. The similarity of results between the two measures suggests that the group

differences seen with the RMS amplitude measure are due to differences in the P2N2 region.

Age did not differ between groups. Age did not correlate with any measure in which group differences were found.

Discussion: Experiment 1

Poor neural timing separates LP subjects

Our findings indicate that background noise can distort the timing of neural events corresponding to speech encoding in some learning-impaired children. As expected, in normal children the addition of background noise diminished the size of the cortical response, but did not affect the timing of morphological features of that waveform. In learning impaired children, the addition of noise decreased the amount of neural activity to a similar extent as in the normal children, but in approximately a quarter of this group, the timing of morphological features of the waveform was altered. Specifically, the latency of the N2 response elicited in noise was earlier in these children. Because only the latest portion of the waveform evoked in noise conflicted morphologically with that evoked in quiet, no amount of latency shift could align all the peaks of the two waveforms, thus producing low quiet-to-noise correlation values. Despite differences in response magnitude across groups, the equivalent decrease in neural activity with the introduction of noise, as revealed by RMS amplitude as well as the amplitude of the P2N2 complex, indicates that the poor cortical representation of the LP-Out children in noise cannot be attributed to an abnormal decrease in overall response activity. Rather, it appears that the activity associated with the neural encoding of speech sounds is being distributed differently over time across the responses recorded in noise in the LP-Out children.

Experiment 2

Materials and methods

Subjects

Twenty-four children who completed Experiment 1 went on to participate in Experiment 2. Thirteen LP children participated in a commercial auditory training program. Four LP children and seven NL children comprised the non-trained test/retest control group. Control subjects did not participate in any form of remediation during the study. Trained subjects were divided into two groups based on their initial quiet-to-noise correlation score. Seven of these children started out with correlation scores within the normal range as defined in Experiment 1 (Trained-In), and six children's correlation scores fell below this range (Trained-Out). Three control subjects (two NL, one LP) had correlation scores below the normal range. Cross-correlation data from an overlapping subject group has previously been reported in the context of a study assessing neural plasticity in children with learning impairments (Hayes et al. 2003) and with respect to auditory brainstem measures (King et al. 2002). This paper is the first to address the relationship between training-related changes in cortical responses and perception.

Training

Auditory training began after completing initial neurophysiological and perceptual (described below) testing. Children in the LP trained group attended 35–40 1-h training sessions during an 8-week period. Training was conducted by an independent agency (Cognitive Concepts, Inc., Evanston, IL), and consisted of supervised work with Earobics Step I (developmental ages 4–7) and Step II (developmental ages 7–10) software (Morrison 1998; Diehl 1999). The Earobics program provides training on phonological awareness, auditory processing and language processing skills through interactive games. Specifically, the program consists of audiovisual exercises in quiet and noise that incorporate phoneme discrimination, auditory memory, auditory sequencing, auditory attention, rhyming, and sound blending skills. Because children presented a wide variety of phonological and auditory processing deficits, and were below age-level on at least one measure of phonological awareness or auditory processing, they all began training with Step I to ensure mastery of basic concepts and moved on to Step II after successfully completing Step I.

Neurophysiological procedures

Cortical neurophysiological responses to /da/ were measured as in Experiment 1.

Speech perception measures

In addition to the standardized measures comprising the psychoeducational battery described above, a speech-sound discrimination measure was used to evaluate training effects on speech perception. This measure, Parameter Estimation by Sequential Tracking (PEST) (Taylor and Creelman 1967; Carrell et al. 1999), used a 41 step speech-sound continuum ranging from /da/ to /ga/, differing only in F_3 onset frequency. Stimuli were 100 ms in duration including a 40 ms formant transition period. No onset burst was present. The PEST algorithm was used in a two interval, two alternative forced choice model to find a just noticeable difference (JND) score as defined by the distance between stimuli in the contrast pair which was correctly identified 69% of the time. The stimuli were presented bilaterally through headphones at 65 dB SPL, and subjects were instructed to press a button to indicate their responses.

Retest

Following training, subjects were retested on the neurophysiology and speech discrimination tests. In addition, the psychoeducational test battery was re-administered post-training. A clinically significant gain on the standardized tests of this battery was defined to be a change greater than one standard error of measurement. The two testing sessions occurred over a 3- to 6-month interval. Control subjects were also tested twice on all tests with a 3- to 6-month interval between testing sessions.

Analysis

Waveforms recorded in quiet and in background noise in both test and retest conditions were analyzed using the quiet-to-noise cross-correlation and RMS amplitude techniques. In addition, amplitudes and latencies of peaks N1, P2, and N2 were assessed for all waveforms. Although all subjects were included in this analysis, not

all subjects exhibited every peak in every waveform. As in Experiment 1, when performing statistics involving specific peaks, if those peaks were not seen in a particular waveform, that waveform was not included in the analysis. Due to the small sample sizes in this experiment, non-parametric statistics were used in all analyses.

Results

No differences in correlation scores between NL and LP control groups were found in either initial or retest sessions (Mann-Whitney U=14.0, p=1.0, Mann-Whitney U=9.5, p>0.3, respectively). Additionally, the change in correlation score from initial to retest session was assessed, again finding no differences between control groups (Mann-Whitney U=9.0, p>0.3).

Therefore, for experiment 2, the NL and LP controls were combined into a single control group (N=11).

Response correlations

The two groups of trained subjects, Trained-In and Trained-Out based on initial correlation scores, differed in how much their quiet-to-noise correlations changed with training (Mann-Whitney U=0.00; p=0.005). Specifically, the Trained-In group's correlation scores did not change (Wilcoxon Z=0.94, p>0.1); their response correlations remained in the normal range after training. The Trained-Out group, whose initial quiet-to-noise correlation scores were poor, improved this score with training (Wilcoxon Z=2.20, p<0.05) (also reported in Hayes et al. 2003). These scores improved to within the normal range (Fig. 6).

Waveform peaks

Training-related changes in the amplitudes and latencies of specific peaks of the response elicited in the quiet condition were assessed. The Trained-Out group's N2 latency evoked in the quiet condition became later with training (N=6; Wilcoxon Z=2.20, p<0.05) such that four of the six subjects showed latency increases of between 11 and 17 ms, and the other two subjects' latencies remained the same. Only one subject in each of the other two groups exhibited N2 latency increases greater than 11 ms. No other differences were seen between the responses recorded in quiet pre- and post-training.

Regretfully, the addition of noise eliminated enough peaks in the evoked waveforms to make fair statistical analysis of training-related amplitude and latency changes of these waveform peaks impossible. For example, based on Experiment 1 as well as the Experiment 2 results from the quiet condition, we wanted to focus our attention on the N2 region of the waveforms evoked in noise, but only three of the six children in the Trained-Out group exhibited an N2 response in the noise condition both pre- and post-training.

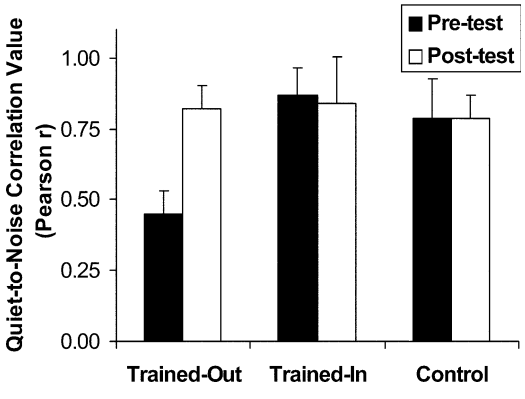

Fig. 6 Quiet-to-noise correlation values pre-test (*shaded bars*) and post-test (*white bars*) for trained subjects whose initial correlation fell outside the normal range (Trained-Out), trained subjects whose initial correlation fell within the normal range (Trained-In), and control subjects. *Error bars* indicate standard error of the means

Speech-sound perception

Before training, the Trained-Out group's scores on the speech-sound discrimination (PEST) test (JND) were significantly higher (worse) than those of the Control group (Mann-Whitney U=6.0, p<0.01), whereas the Trained-In group's scores did not differ from those of the Control group (Mann-Whitney U=20.5, p>0.1). After training, the Trained-Out group's improvements in quiet-to-noise correlations were accompanied by improvements on the speech-sound perception measure (Wilcoxon Z=1.99, p<0.05), suggesting a relationship between neurophysiology and speech perception (Fig. 7A). The Trained-In and Control groups showed no change on this measure (Wilcoxon Z=0.17, 1.07 respectively, p>0.1 both groups).

Specific attention was paid to the learning/perceptual tests composing the Auditory Processing score which was previously shown to improve with training: the Sound Blending and Incomplete Words tests (Hayes et al. 2003). Interestingly, although all groups showed improvement on the Auditory Processing (Trained-Out: Wilcoxon Z=2.21, p<0.05; Trained-In: Wilcoxon Z=2.20, p<0.05, Control: Wilcoxon Z=1.89, p=0.059) and Incomplete Words (Trained-Out: Wilcoxon Z=2.20; Trained-In: Wilcoxon Z=2.37; Control: Wilcoxon Z=2.30, p<0.05 all groups) (Fig. 7c) measures, improvement on the Sound Blending test appeared related to improved quiet-to-noise correlations as only the Trained-Out group showed significant gains on this measure (Trained-Out: Wilcoxon Z=2.21, p<0.05; Trained-In: Wilcoxon Z=1.36, p>0.1; Control: Wilcoxon Z=0.47, p>0.6) (Fig. 7B). Looking at the data another way, we found that all but one subject in each of the trained groups showed clinically significant improvement on the Auditory Processing score, and all but one trained child showed clinically significant improvement on the Incomplete Words score, but two out of seven Trained-In subjects (29%) compared to four out of six Trained-Out

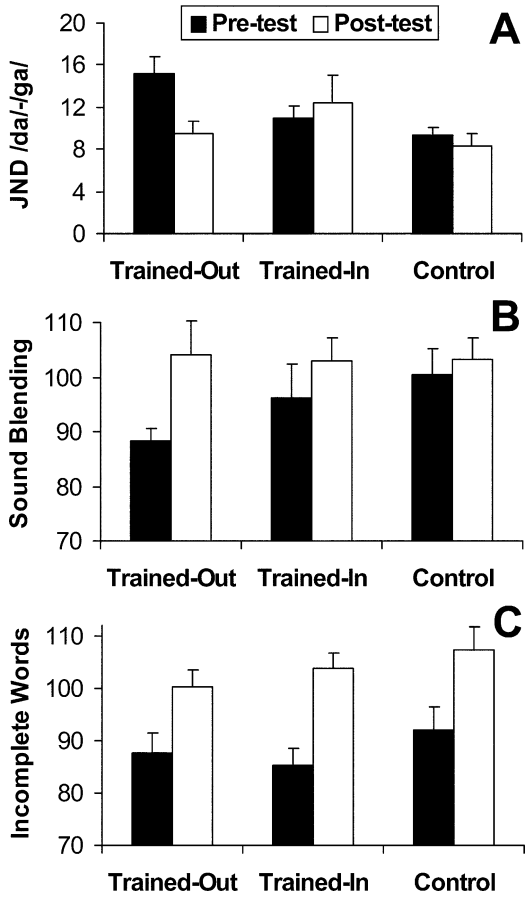

Fig. 7A–C Pre-test (*shaded*) and post-test (*white*) scores on phonemic perception tasks with subjects grouped by pre-test correlation. **A** Just noticeable difference (*JND*) scores for /da/-/ga/ continuum, **B** scores on Sound Blending subtest, and **C** scores on Incomplete Words subtest. No changes were seen in the Trained-In or control groups on either the /da/-/ga/ discrimination or Sound Blending task. The Trained-Out group showed a significant improvement on both tasks, bringing their post-test scores to within the normal range (/da/-/ga/ JND: lower scores are better; Sound Blending: higher scores are better). All groups improved on the Incomplete Words test. *Error bars* indicate standard error of the means

subjects (67%) showed clinically significant improvement on the Sound Blending score, again suggesting that improvement on this measure is related to improved neurophysiology (exhibited by the Trained-Out group only). A chi-square analysis on these data did not make significance (χ^2=2.2, p>0.05), possibly due to the small sample sizes of these groups. Although these results are encouraging, the small effect and sample sizes of each group recommend replication with larger subject groups.

Discussion: Experiment 2

Training improves neural responses and speech discrimination

Our findings suggest that LP subjects with poor quiet-to-noise response correlations are particularly able to benefit from commercial auditory training programs. LP subjects with poor cortical timing in noise improved to within the normal range after training. Their performance on the speech-perception measure also improved. This suggests that training programs aid in helping children with poor cortical encoding of speech in noise to more accurately preserve the representation of speech in challenging listening environments, and that this improvement may translate to an improvement in their speech perception.

The commercial training program associated with this study includes exercises on distinguishing phonemes in the presence of background noise. Because consonants are characterized by low amplitude fast formant transitions that are easily masked with noise, these exercises may help train the auditory system to extract this information from a noisy auditory stream. Thus this training may encourage the neural encoding of parts of the speech signal that were previously "overlooked" because the system could not discern critical rapid transitions. Our findings indicate that success in such training may be due to improved neural encoding of temporally important information vital to distinguishing the acoustic components that define speech sounds.

As in the previous training study (Hayes et al. 2003), the trained group as a whole improved their Auditory Processing scores, a composite of Sound Blending and Incomplete Words scores. However, whereas improvements on Auditory Processing and Incomplete Words were seen regardless of neurophysiological results, improvements on Sound Blending appeared related to cortical neurophysiological changes. Sound Blending tests phonemic awareness by requiring subjects to synthesize sounds to form a word. Incomplete Words assesses phonemic awareness by requiring subjects to complete words that are missing phoneme(s). Performance on discriminating tokens along a speech-sound continuum (/da/-/ga/) was also related to changes in quiet-to-noise correlations. This result was not reported in Hayes et al. (2003) because the trained group as a whole did not show improvement on this measure. Only when grouping the children based on quiet-to-noise correlations did this result become evident. Therefore, the correlation score was related to tests that required the decoding of sounds present in speech (Sound Blending, /da/-/ga/ discrimination) but not the generation of missing sounds (Incomplete Words), which requires more top-down processing. This implies that the fidelity of the timing of responses recorded in noise is related to the perceptual processing of speech.

General discussion

The goals of this study were to evaluate the degree to which normal learning and learning impaired individuals' neurophysiological cortical response morphology to speech syllables changed in background noise, and to examine whether training-related changes in cortical responses related to perception. Background noise affected the timing of the cortical representation of speech in a

higher proportion of children with learning problems than in normal children. This representation improved with auditory training, and those learning impaired children who showed the most improvement in response correlation also showed improved perceptual encoding of speech.

It is not surprising that only some learning impaired children showed a deficit in the timing of cortical evoked responses in the presence of background noise. The learning impaired population is heterogeneous. Although many children show similar learning deficits, such as dyslexia, the origin of these deficits probably differs from child to child. This study did not attempt to separate children with auditory-based learning disabilities from children whose difficulties stem from other sources. Consistent with previous studies, the percent of learning-impaired children with poor response correlations in this study matches the incidence of learning disabilities having an auditory basis (Kraus et al. 1996; King et al. 2002). Additionally, it was the trained children with poor initial auditory response correlations who also began with poor speech perception scores. This suggests that these children may comprise a group in which auditory difficulties underlie their learning disabilities.

The analyses used in this study focused on the 100–350 ms range, which encompasses the N1P2N2 complex. The N2 region of the LP-Outs' response was most affected by the addition of background noise. Multiple neural structures have been implicated in generating the N2 component (Näätänen et al. 1982; Näätänen and Picton 1986), but the chief neural generators for this response are found in auditory cortices located on the supratemporal plane bilaterally (Čeponienė et al. 2002). Therefore, this study suggests that in some learning impaired children the timing of firing of neuronal ensembles in cortical auditory regions can be altered by the addition of background noise to a listening environment. Additionally, it has been shown that the timing of responses originating in these regions can be adjusted through auditory training.

It is well documented that children with learning impairments show increased difficulty perceiving speech in noise. We have shown that the addition of noise alters the timing of neural mechanisms approximately 200–300 ms into the auditory processing stream (N2). Whereas earlier neural events (N1 and P2) are thought to reflect pre-attentive stimulus detection and encoding (Näätänen and Picton 1987; Martin et al. 1997; Shtyrov et al. 2000), this stage of processing has previously been associated with speech perception at the phonetic level (Maiste et al. 1995).

Performance on a spectrum of auditory processing measures including Auditory Processing, Incomplete Words, and Sound Blending has been found to predict speech-evoked N2 latency, with better scores predicting earlier latency (Cunningham et al. 2000). Additionally, abnormal N2 responses have been linked to other clinical diagnoses associated with auditory perceptual impairments; patients with severe language impairment (Tonnquist-Uhlen 1996a, 1996b), combined subtype of attention deficit/hyperactivity disorder (Johnstone et al. 2001), or sensorineural hearing loss (Oates et al. 2002) demonstrate delayed N2 latency, and children with suspected auditory processing disorder exhibit increased P2N2 amplitudes (Liasis et al. 2003). In the learning impaired population, the timing of speech evoked potentials in this latency region has previously been shown to be sensitive to the combined stresses of repetition and noise (Wible et al. 2002), and the amplitude of these responses has been linked to impaired speech perception in noise (Cunningham et al. 2001). Therefore, it is not unreasonable to assume that the N2 latency differences seen in this study are related to perceptual auditory processing, and phonemic perception specifically.

Changes in the N2 latency of speech-evoked responses in quiet were related to perceptual improvement on two phonemic perceptual tasks after auditory-based training. Training-related changes in the timing of events within the 100–350 ms range of the same responses evoked in noise were also evident by the increased response correlation scores of children showing these behavioral improvements. It was not possible to elucidate the exact nature of the timing changes seen in the noise condition, but learning-related changes within this latency range have been previously described. For example, speech-related training in the form of commercial auditory training or that specifically focusing on discriminating voice onset time has been shown to elicit changes in speech-evoked potentials from P1 to N2 (Tremblay et al. 2001; Tremblay and Kraus 2002; Hayes et al. 2003). In addition, changes in N1 and P2 have been related to other types of auditory training including fine frequency discrimination, complex tonal pattern discrimination, and long-term musical training (Menning 2000; Shahin et al. 2003; Atienza 2002; Bosnyak 2002). Auditory training can therefore affect a number of cortical neural mechanisms involved in the detection and discrimination of sound, including speech.

Previous data from an overlapping dataset has associated delayed brainstem onset responses to speech with training-related improvements in both cortical quiet-to-noise correlations evoked by speech and speech perception measures (King et al. 2002; Hayes et al. 2003). This relationship was also evident in our current data in that children with delayed speech-evoked brainstem onset responses exhibited improved cortical quiet-to-noise correlations, though children with brainstem responses within the normal range showed no such change.

Brainstem responses, occurring earlier in the auditory stream than cortical responses, reflect exact stimulus timing where differences of fractions of milliseconds can be clinically significant (Josey 1985; Musiek 1991). Cortical responses reflect the simultaneous firing of multiple cortical auditory regions along the primary auditory pathway (Scherg and Von Cramon 1985). The timing of events in cortical responses becomes more abstract, no longer relating directly to the stimulus, but must also demonstrate precise timing in order to evoke a clear perception of sound (Phillips 1993). One can imagine that an input with degraded timing would propagate through the system and affect later processing. It is

interesting to note that the timing of the cortical response to speech can be altered independently of the brainstem onset response which does not change with training (Hayes et al. 2003). This suggests that there is a mechanism for altering cortical stimulus encoding even when the input from the brainstem is deficient. However, new evidence now suggests that later brainstem measures, including a quiet-to-noise response correlation over the 12–30 ms time range may change with auditory training (Russo et al. 2003). Additionally, in a small subject sample, changes in this brainstem measure were found to parallel changes in the cortical response correlation described in this paper, suggesting a direct link between changes in the encoding of speech in noise at the brainstem and cortical levels.

Although the exact mechanisms of auditory learning-related plasticity remain unclear, a number of theories have been proposed. For example, some studies suggest that auditory cortical responses can be modified through the thalamo-cortical pathway (see Edeline 2003 for review). It has also been proposed that training induces changes in cortical auditory receptive fields that are mediated by acetylcholinergic pathways originating in the nucleus basalis of the basal forebrain (Kilgard and Merzenich 1998, 2002). Our study measured only changes in far-field evoked potentials and is therefore unequipped to address the possible mechanisms of the observed changes.

Changes in waveform morphology were addressed with two complementary analysis techniques in this study. Changes across the entire waveform were assessed with response correlation and RMS amplitude, and more focal changes were measured with amplitude and latency values of specific peaks within the waveform. The peak measurements were more subjective and time-intensive to acquire than the global measurements, but offered informative insights into specific morphological changes occurring in the waveforms. Response correlation proved sensitive to changes in waveform morphology within individuals, regardless of whether the responses conformed to an expected morphology or not. In one condition, high morphological variability combined with a small subject sample to render statistical analysis of peak measurements impractical, but assessment of changes in the timing of events across the waveform was made possible by response correlation. We therefore conclude that both global and focal measurements are useful tools, and that each can inform the other.

This study demonstrates that some children with learning problems exhibit abnormal timing of cortical responses to speech stimuli presented in background noise. Auditory training was shown to improve the timing of these responses in some children. Phonetic decoding skills improved in these same children. Further investigation may lead to clinical identification of children with poor cortical representation of speech sounds. Such identification may have diagnostic applications in determining whether sound perception deficits underlie learning problems in individual children, and in deciding whether a child would benefit from auditory training. Additionally, the analysis techniques outlined in this paper offer a method for objectively monitoring the neurophysiological effects of auditory training programs.

Acknowledgements We would like to thank the children and their families for participating in this study. This work was supported by National Institute of Health Grant R01DC01510.

References

Alain C, Woods D, Knight R (1998) A distributed cortical network for auditory sensory memory in humans. Brain Res 812:23–37

Atienza M, Cantero JL, Dominguez-Marin E (2002) The time course of neural changes underlying auditory perceptual learning. Learn Mem 9:138–150

Bellis TJ (1996) Assessment and management of central auditory processing disorders in the educational setting: from science to practice. Singular Publishing, San Diego

Bosnyak DJ, Eaton RA, Roberts LE (2002) Enhancement of multiple components of the auditory evoked potential in nonmusicians by training for pitch discrimination with 40-Hz amplitude modulated tones. Proceedings of the 13th International Conference on Biomagnetism, Jena, Germany, August

Bradlow AR, Kraus N, Nicol TG, McGee TJ, Cunningham J, Zecker SG, Carrell TD (1999) Effects of lengthened formant transition duration on discrimination and neural representation of synthetic CV syllables by normal and learning-disabled children. J Acoust Soc Am 106:2086–2096

Bradlow AR, Kraus N, Hayes E (2003) Speaking clearly for children with learning disabilities: sentence perception in noise. J Speech Lang Hear Res 46:80–97

Carrell TD, Bradlow AR, Nicol TG, Koch DB, Kraus N (1999) Interactive software for evaluating auditory discrimination. Ear Hearing 20:175–176

Čeponienė R, Rinne T, Näätänen R (2002) Maturation of cortical sound processing as indexed by event-related potentials. Clin Neurophysiol 113:870–882

Chermak GD, Musiek FE (1997) Central auditory processing disorders: new perspectives. Singular Publishing, San Diego

Cunningham J, Nicol T, Zecker S, Kraus N (2000) Speech-evoked neurophysiologic responses in children with learning problems: development and behavioral correlates of perception. Ear Hearing 21:554–568

Cunningham J, Nicol T, Zecker SG, Kraus N (2001) Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. Clin Neurophysiol 112:758–767

Diehl S (1999) Listen & Learn? A software review of Earobics, language, speech and hearing services in schools. ASHA 30:108–116

Edeline JM (2003) The thalamo-cortical auditory receptive fields: regulation by the states of vigilance, learning and the neuromodulatory systems. Exp Brain Res 153:554–572

Elliott LL (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

Elliott LL, Hammer MA, Scholl ME (1989) Fine-grained auditory discrimination in normal children and children with language-learning problems. J Speech Hear Res 32:112–119

Hayes E, Warrier CM, Nicol T, Zecker SG, Kraus N (2003) Neural plasticity following auditory training in children with learning problems. Clin Neurophysiol 114:673–684

Johnstone SJ, Barry RJ, Anderson JW (2001) Topographic distribution and developmental timecourse of auditory event-related potentials in two subtypes of attention-deficit hyperactivity disorder. Int J Psychophysiol 42:73–94

Josey A (1985) Auditory brainstem response in site of lesion testing. In: Katz J (ed) Handbook of clinical audiology. Williams and Wilkins, Baltimore, MD, pp 534–548

Kilgard MP, Merzenich MM (1998) Cortical map reorganization enabled by nucleus basalis activity. Science 279:1714–1718

Kilgard MP, Merzenich MM (2002) Order-sensitive plasticity in adult primary auditory cortex. Proc Natl Acad Sci U S A 99:3205–3209

King C, Warrier CM, Hayes E, Kraus N (2002) Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. Neurosci Lett 319:111–115

Klatt DH (1980) Software for a cascade/parallel formant synthesizer. J Acoust Soc Am 67:971–995

Kraus N, McGee TJ, Carrell TD, Zecker SG, Nicol TG, Koch DB (1996) Auditory neurophysiologic responses and discrimination deficits in children with learning problems. Science 273:971–973

Liasis A, Bamiou DE, Campbell P, Sirimanna T, Boyd S, Towell A (2003) Auditory event-related potentials in the assessment of auditory processing disorders: a pilot study. Neuropediatrics 34:23–29

Maiste AC, Wiens AS, Hunt MJ, Scherg M, Picton TW (1995) Event-related potentials and the categorical perception of speech sounds. Ear Hearing 16:68–90

Martin BA, Sigal A, Kurtzberg D, Stapells DR (1997) The effects of decreased audibility produced by high-pass noise masking on cortical event-related potentials to speech sounds /ba/ and /da/. J Acoust Soc Am 101:1585–1599

Martin BA, Kurtzberg D, Stapells DR (1999) The effects of decreased audibility produced by high-pass noise masking on N1 and the mismatch negativity to speech sounds /ba/ and /da/. J Speech Lang Hear Res 42:271–286

Menning H, Roberts LE, Pantev C (2000) Plastic changes in the auditory cortex induced by intensive frequency discrimination training. Neuroreport 11:817–822

Morrison S (1998) Computer applications. Earobics Pro Child Lang Teaching Ther 14:279–284

Musiek FE (1991) Auditory evoked responses in site-of-lesion assessment. In: Rintelmann WF (ed) Hearing assessment. Pro-Ed, Austin, TX, pp 383–427

Näätänen R, Picton TW (1986) N2 and automatic versus controlled processes. Electroencephalogr Clin Neurophysiol Suppl 38:169–186

Näätänen R, Picton T (1987) The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24:375–425

Näätänen R, Simpson M, Loveless NE (1982) Stimulus deviance and evoked potentials. Biol Psychol 14:53–98

Nabelek AK, Pickett JM (1974) Reception of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids. J Acoust Soc Am 56:628–639

Oates PA, Kurtzberg D, Stapells DR (2002) Effects of sensorineural hearing loss on cortical event-related potential and behavioral measures of speech-sound processing. Ear Hear 23:399–415

Phillips DP (1993) Representation of acoustic events in the primary auditory cortex. J Exp Psychol Hum Percept Perform 19:203–216

Ponton CW, Vasama JP, Tremblay K, Khosla D, Kwong B, Don M (2001) Plasticity in the adult human central auditory system: evidence from late-onset profound unilateral deafness. Hear Res 154:32–44

Russo NM, Nicol T, Hayes E, Zecker S, Kraus N (2003) Auditory training can improve neural timing in the auditory brainstem. Program No. 400.33. 2003 Abstract Viewer/Itinerary Planner. Society for Neuroscience, Washington, DC, 2003. Online

Scherg M, Von Cramon D (1985) Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. Electroencephalogr Clin Neurophysiol 62:32–44

Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE (2003) Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. J Neurosci 23:5545–5552

Shtyrov Y, Kujala T, Lyytinen H, Ilmoniemi RJ, Näätänen R (2000) Auditory cortex evoked magnetic fields and lateralization of speech processing. Neuroreport 11:2893–2896

Stark RE, Heinz JM (1996) Vowel perception in children with and without language impairment. J Speech Hear Res 39:860–869

Tallal P (1980) Auditory temporal perception, phonics, and reading disabilities in children. Brain Lang 9:182–198

Tallal P, Piercy M (1974) Developmental aphasia: rate of auditory processing and selective impairment of consonant perception. Neuropsychologia 12:83–93

Taylor MM, Creelman CD (1967) PEST: efficient estimates on probability functions. J Acoust Soc Am 41:782–787

Tonnquist-Uhlen I (1996a) Topography of auditory evoked cortical potentials in children with severe language impairment. Scand Audiol Suppl 44:1–40

Tonnquist-Uhlen I (1996b) Topography of auditory evoked long-latency potentials in children with severe language impairment: the P2 and N2 components. Ear Hearing 17:314–326

Tremblay KL, Kraus N (2002) Auditory training induces asymmetrical changes in cortical neural activity. J Speech Lang Hear Res 45:564–572

Tremblay K, Kraus N, McGee T, Ponton C, Otis B (2001) Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. Ear Hear 22:79–90

Whiting KA, Martin BA, Stapells DR (1998) The effects of broadband noise masking on cortical event-related potentials to speech sounds /ba/ and /da/. Ear Hear 19:218–231

Wible B, Nicol T, Kraus N (2002) Abnormal neural encoding of repeated speech stimuli in noise in children with learning problems. Clin Neurophysiol 113:485–494

Wilkinson G (1993) Wide range achievement test-3. Jastak Assoc., Wilmington, DE

Woodcock R, Johnson M (1977) Woodcock-Johnson Psycho-Educational Battery. Tests of cognitive ability. DLM Teaching Resources, Allen, TX

Woodcock R, Johnson M (1989) Woodcock-Johnson Psycho-Educational Battery-Revised. Tests of cognitive ability. DLM Teaching Resources, Allen, TX