ELSEVIER

Clinical Neurophysiology 112 (2001) 758–767

www.elsevier.com/locate/clinph

Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement

Jenna Cunningham[a,*], Trent Nicol[a], Steven G. Zecker[a], Ann Bradlow[b], Nina Kraus[a,c,d]

[a]*Department of Communication Sciences, Northwestern University, 2299 North Campus Drive, Evanston, IL 60208-3550, USA*
[b]*Department of Linguistics, Northwestern University, 2299 North Campus Drive, Evanston, IL 60208-3550, USA*
[c]*Department of Neurobiology and Physiology, Northwestern University, 2299 North Campus Drive, Evanston, IL 60208-3550, USA*
[d]*Department of Otolaryngology, Northwestern University, 2299 North Campus Drive, Evanston, IL 60208-3550, USA*

See Editorial, pages 721–723

Abstract

Objectives: Some children with learning problems (LP) experience speech-sound perception deficits that worsen in background noise. The first goal was to determine whether these impairments are associated with abnormal neurophysiologic representation of speech features in noise reflected at brain-stem and cortical levels. The second goal was to examine the perceptual and neurophysiological benefits provided to an impaired system by acoustic cue enhancements.

Methods: Behavioral speech perception measures (just noticeable difference scores), auditory brain-stem responses, frequency-following responses and cortical-evoked potentials (P1, N1, P1′, N1′) were studied in a group of LP children and compared to responses in normal children.

Results: We report abnormalities in the fundamental sensory representation of sound at brain-stem and cortical levels in the LP children when speech sounds were presented in noise, but not in quiet. Specifically, the neurophysiologic responses from these LP children displayed a different spectral pattern and lacked precision in the neural representation of key stimulus features. Cue enhancement benefited both behavioral and neurophysiological responses.

Conclusions: Overall, these findings contribute to our understanding of the preconscious biological processes underlying perception deficits and may assist in the design of effective intervention strategies.

Keywords: Children; Learning and attention disorders; Central auditory physiology; Noise; Speech-sound perception; Acoustic cue enhancement

1. Introduction

Recent studies indicate that speech-sound perception deficits may contribute to the learning problems (LP) of some children. In particular, these children have difficulty discriminating between acoustically similar sounds (Tallal and Piercy, 1974; Tallal, 1980; Elliott et al., 1989; Stark and Heinz, 1996; Kraus et al., 1996; Bradlow et al., 1999). Moreover, these deficits become worse in the presence of background noise (Nabelek and Pickett, 1974; Elliott, 1979; Bellis, 1996; Chermak and Musiek, 1997). While the underlying cause of speech-sound perception deficits remains controversial (Tallal and Piercy, 1974; Tallal, 1980; Nittrouer, 1992; Studdert-Kennedy and Mody, 1995; Denneberg, 1999), new evidence suggests that basic neurophysiologic processes related to stimulus encoding and discrimination may be involved. Three recent studies have begun to elucidate the biological bases of impaired speech perception in some individuals with LP. First, poor readers differed from good readers in neural recovery time of auditory cortical responses to rapidly presented stimuli (Nagarajan et al., 1999). Second, dyslexic individuals displayed significantly smaller far-field EEG amplitude modulated following responses than normal subjects (McAnally and Stein, 1997). Finally, a subset of children with LP showed a significant reduction in cortical responses to speech-sound contrasts differing in rapid spectro-temporal elements, consistent with their impaired behavioral discrimination of those stimuli (Kraus et al., 1996).

Despite general acknowledgement that background noise excessively taxes perception in most children with LP, little is known about the underlying neurobiologic processes. The first goal of this investigation was to determine whether speech perception deficits in some LP children are associated with abnormal neurophysiologic representation of rapidly changing speech features in noise reflected by potentials generated at brain-stem and cortical levels. To accomplish this aim, auditory brain-stem responses (ABR), frequency-

* Corresponding author. Tel.: +1-847-491-2465; fax: +1-847-491-2523.
E-mail address: jennacu@msn.com (J. Cunningham).

1388-2457/01/$ - see front matter
PII: S1388-2457(01)00465-5

CLINPH 2000139

following responses (FFR) and cortical-evoked potentials were studied in a group of LP children and compared to responses in normal children.

Evaluation of these electrophysiologic measures separately and in combination provides a unique opportunity to assess the integrity of central auditory stimulus-timing mechanisms at various levels of the auditory pathway. For instance, the ABR reflects neural activity synchronized to the stimulus onset. It is generated by action potentials traveling along axons in a pattern of short-duration, biphasic responses. The magnitude of the ABR depends on a high degree of synchronized firing among the neurons, such that deviations of tenths of milliseconds are considered diagnostic of brain-stem pathology (Starr and Don, 1988). If there is excessive neural 'jitter', which might occur in an impaired auditory system, the separation of individual neural responses by even a fraction of a millisecond could cause responses to cancel each other out. The FFR also depends on a high degree of neural synchrony. It reflects brain-stem-generated, phase-locked responses to the low frequency components of a stimulus (less than 800 Hz) (Worden and Marsh, 1968). Differences between brain-stem and cortical responses are particularly apparent in the overall spectra of the evoked responses (ABR ~1 kHz, Boston and Moller, 1985; cortical potentials ~10 Hz, Moller, 1994). Cortical responses reflect the summation of excitatory post-synaptic potentials originating from multiple generator sites in response to stimulus onset and other acoustic features of the stimulus. These slow dendritic events can be separated by several milliseconds and will still sum constructively. Nevertheless, cortical potentials do depend on stimulus-locked synchronous firing across neural ensembles.

The second goal of this study was to examine the perceptual and neurophysiological benefits provided to an impaired system by acoustic cue enhancements typical of 'clear' speech (Picheny et al., 1986) in order to gain a deeper understanding of how processing deficits can be overcome by the speech signal. Research has shown that speakers naturally alter the acoustic characteristics of their speech from a 'conversational' to a 'clear' speaking style when the listener is known to have speech perception difficulties. The acoustic characteristics of 'conversational' and 'clear' speech have been well described (Picheny et al., 1986). The perceptual benefits of 'clear' speech have also been established (Picheny et al., 1985; Gordon-Salant, 1986; Hazan and Simpson, 1998), and some of these features have been incorporated into commercially available auditory training programs designed for LP children (Merzenich et al., 1996; Tallal et al., 1996).

2. Methods

2.1. Subjects

Subjects consisted of normal children ($n = 9$) and children with LP ($n = 9$). The normal group included children between the ages of 10 and 13 years (3 female, 6 male) with no reported history of learning or attention problems and scores within normal limits on standardized tests of learning and academic achievement (Woodcock and Johnson, 1977, 1989; Wilkinson, 1993). The group with LP included age-matched children (two female, 7 male) diagnosed clinically with a reading-based learning disability (LD). Six of the 9 children were also clinically diagnosed with attention deficit disorder (LD/ADD). Because perceptual deficits are known to cut across diagnostic categories, children with different clinical diagnoses (LD and LD/ADD) were combined in the LP group (Kraus et al., 1996). Five of the LP children had a history of expressive language deficits and 3 children had a history of middle ear infections. All subjects demonstrated normal intelligence measured by the Brief Cognitive Scale (IQ measure) of the Woodcock–Johnson Psycho-Educational Battery (scores >85) and normal peripheral hearing sensitivity (<25 dB HL re: normal hearing level) from 500 to 4000 Hz. Children with LP performed significantly poorer than normal children on measures of auditory processing (Mann–Whitney: $U = 0$, $P = 0.0002$, two-tailed; Woodcock and Johnson, 1977, 1989), reading ($U = 4$, $P = 0.0009$, two-tailed; Wilkinson, 1993), spelling ($U = 3$, $P = 0.0007$, two-tailed; Wilkinson, 1993) and behavioral discrimination of fine-grained acoustic differences along a /da/ to /ga/ continuum in a quiet background ($U = 1$, $P = 0.0003$, two-tailed; Kraus et al., 1996).

2.2. Behavioral perception

2.2.1. Stimuli

Five-formant synthetic speech syllables along an /ada/ to /aga/ continuum were produced with a Klatt cascade/parallel formant synthesizer (Klatt, 1980). The stimuli were chosen because it has been previously shown that some children with LP have difficulty perceiving stop consonants in general (for review, see Bradlow et al., 1999) and /da/–/ga/ specifically (Kraus et al., 1996). The first /a/ vowel was 50 ms in duration with a 40 ms formant transition into the stop gap. The stop gap was followed by a 40 ms formant transition containing a 10 ms release burst at the initial portion of the transition and a 50 ms final vowel /a/. Vowel formant frequencies were 720, 1240, 2500, 3600 and 4500 Hz for F1, F2, F3, F4 and F5, respectively. Starting frequencies for the consonant were 220, 1700, 3600 and 4500 Hz for F1, F2, F4 and F5, respectively. F3 varied as described below. Stimuli had a fundamental frequency of 125 Hz. The voicing amplitude was stable across the stimulus.

These acoustic parameters were used to design 4 separate /ada/ to /aga/ continua. The stop gap duration and release burst intensity of the stimuli in the first continuum were modeled according to the acoustic characteristics of 'conversational' speech while the second matched 'clear' speech features relevant to stop consonants (Picheny et

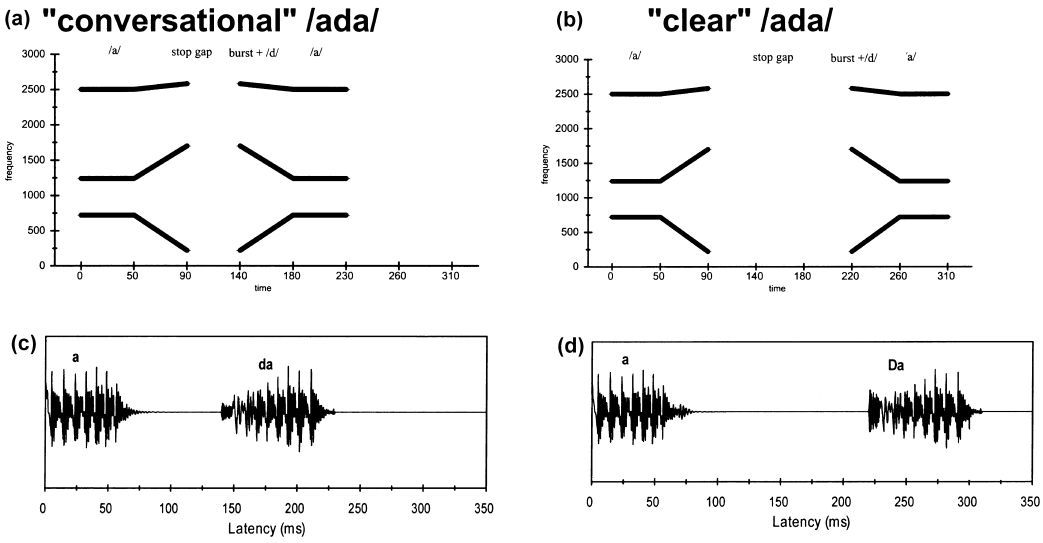

Fig. 1. (a) Spectrogram represents the first 3 formant frequencies (F1, F2 and F3) for the 'conversational' /ada/ stimulus. (b) Spectrogram represents the first 3 formant frequencies (F1, F2 and F3) for the 'clear' /ada/ stimulus. (c) Stimulus waveform illustrates the acoustic characteristics of the 'conversational' /ada/ stimulus. (d) Stimulus waveform illustrates the acoustic characteristics of the 'clear' /ada/ stimulus. Note that the stop gap duration is lengthened to 130 ms and the release burst intensity is increased by 10 dB compared to the 'conversational' /ada/ stimulus.

al., 1986). Natural speech production from 4 young adult speakers saying 'I am going to say /ada/ now' conversationally and again in a clear manner confirmed findings by Picheny et al. (1986) that speakers increase the stop gap duration and release burst intensity of plosives in 'clear' speech. Thus, in accordance with these measurements, the synthesized 'conversational' stimuli had a stop gap duration of 50 ms, the consonant release burst to vowel intensity ratio (CV ratio) was −18 dB sound pressure level (SPL) and the total stimulus duration was 230 ms. The acoustic characteristics and spectrogram of the 'conversational' /ada/ are displayed in Fig. 1a,c. For the 'clear' stimuli, the stop gap duration was increased to 130 ms, the release burst intensity was increased by 10 dB SPL so that the CV ratio was −8 dB SPL and the total stimulus duration was 310 ms. The acoustic characteristics and spectrogram of the 'clear' /ada/ are illustrated in Fig. 1b,d.

Both cue enhancements were applied uniformly to all stimuli in the second continuum. In order to evaluate the effects of each cue enhancement separately, only the stop gap duration was lengthened to 130 ms in the third continuum. In the fourth continuum, only the release burst intensity was increased by 10 dB SPL. Stimuli were presented binaurally through Sennheiser HD 540 earphones at an overall intensity of 65 dB SPL. Continuous broadband Gaussian white noise was presented at 60 dB SPL.

2.2.2. Procedure

Just noticeable difference scores (JNDs) were obtained using an adaptive tracking algorithm (parameter estimation by sequential tracking, PEST) (Taylor and Creelman, 1967). Ideal examples of the syllables /ada/ and /aga/ served as endpoints of the continua and values of the synthesis parameters were interpolated linearly to generate the intermediate stimuli (Kraus et al., 1996; Walley and Carrell, 1983; Carrell et al., 1999). The continua reflected a change in the consonant's third formant (F3) onset frequency. F3 changed from 2580 Hz for /ada/ to 2180 Hz for /aga/ in 40 steps of 10 Hz each. A detailed description of behavioral test procedures has been previously reported (Kraus et al., 1999).

Both normal and LP subjects first discriminated along the 'conversational' continuum in quiet. This procedure, coupled with experience in discriminating /da/–/ga/ in another study (Kraus et al., 1996), familiarized subjects with the task so that subsequent differences could be attributed to true perceptual differences and not task performance. Following the quiet presentation, all children discriminated the 'conversational' and 'clear' continua in noise presented in random order. In addition, 8 of the 9 LP children also discriminated the third and fourth continua in noise to determine which cue enhancement (stop gap duration or release burst intensity) was more beneficial in ameliorating speech discrimination deficits in the LP population.

2.3. Electrophysiology

2.3.1. Stimuli

A 5-formant synthetic speech syllable, /da/, was produced with a Klatt cascade/parallel formant synthesizer (Klatt, 1980). Requisite to recording ABR is the use of a short stimulus with a brief rise time and a rapid rate of presentation. Thus, a 40 ms /da/ was constructed using the identical acoustic parameters as the consonant and formant transitions into the final vowel of the /ada/ stimuli used in the

behavioral experiment. Specifically, /da/ was composed of a 10 ms release burst and formant frequencies transitioning in 40 ms from /d/ to the vowel /a/. Stimuli had a fundamental frequency of 125 Hz. Starting frequencies were 220, 1700, 2580, 3600 and 4500 Hz transitioning to 720, 1240, 2500, 3600 and 4500 Hz, for F1, F2, F3, F4 and F5, respectively. Voicing amplitude was stable across the stimulus.

A cue enhanced /da/ condition, consisting of 10 dB SPL amplification of the release burst intensity (determined to be the more important characteristic of 'clear' speech from the behavioral experiment, see below), was also studied.

2.3.2. Procedure

Aggregate neural responses were collected to both stimuli in quiet and in noise. Speech stimuli were presented to the right ear through an Etymotic ER-3 insert earphone at an overall intensity of 80 dB SPL. Continuous broadband Gaussian white noise was presented at 75 dB SPL.

Electrophysiologic measures included the ABR, FFR, and cortical-evoked potentials. For the electrophysiologic procedures, subjects sat in a reclining chair in an electrically-shielded, sound-treated booth and viewed a videotaped movie of their choice. The audio portion of the videotape was held constant below 40 dB SPL and was monitored by the subject's unoccluded left ear. All subjects were instructed to ignore the sounds presented to the right ear and to sit as quietly as possible. None of the subjects experienced difficulty in complying with these instructions.

2.4. Brainstem responses

The ABR and FFR were obtained with Cz active, forehead ground and right earlobe reference electrodes. Both evoked potentials were collected simultaneously using a 10 ms pre-stimulus and 60 ms post-stimulus recording window with a 20 kHz sampling rate. On-line bandpass filtering was from 100 to 2000 Hz. Stimuli were presented with an interstimulus interval of 51 ms. Approximately 4000 responses were averaged for each condition.

2.4.1. Data analysis

2.4.1.1. FFR. Examination of the FFR was employed in order to provide an overall assessment of the magnitude of phase-locking to the stimulus fundamental frequency and its harmonics. Visual inspection of the FFR revealed that the phase-locked response was largest between 20 and 40 ms across subjects. A Fast Fourier Transform (FFT) was performed on this portion of the averaged FFR in quiet and in noise. Following the generation of the FFT, the spectral components were divided into 3 frequency bins (0–200, 250–400 and 450–750 Hz) and amplitude values within each bin were summed. Bin widths were designed to segregate the major frequency component at 125 Hz corresponding to the stimulus fundamental frequency into bin 1, and the first formant frequencies transitioning from 220 to 720 Hz into bins 2 and 3.

2.4.1.2. Stimulus-to-response correlations. Correlations between the stimulus and response waveforms were obtained as a measure of the 'precision' of phase-locking. Each individual FFR waveform was correlated with a stimulus waveform template. Correlation coefficients were successively calculated as the stimulus waveform was shifted in time relative to the FFR waveform. Subsequently, at a peak delay between 5.6 and 8.1 ms, the correlation coefficient (r) value was assigned to the maximal r value within this latency range. This range was determined by aligning the grand-averaged FFR elicited by /da/ and the enhanced /da/ for all normal subjects in both quiet and noise with the formant transition portion of the stimulus waveform. Consistent with previous reports, we observed that the peak latencies of the FFR were equivalent in quiet and noise (Moushegian et al., 1978; Bledsoe and Moushegian, 1980). Accordingly, the same lag time range was used for both conditions.

2.4.1.3. ABR. The ABR was evaluated in order to assess the integrity of the neurophysiologic response to the onset of the stimulus. The amplitudes and latencies of wave V were measured for each individual subject from their average waveforms obtained in quiet and in noise. Wave V latencies ranged from 5.6 to 6.6 ms in quiet and from 6.1 to 8.1 ms in noise for all subjects combined. Evidence to support the labeling of this response as wave V is provided by its response latency in quiet and in continuous background noise. A pilot study evaluating changes in the stimulus onset response to /da/ (80 dB SPL) in various levels of background noise (70, 75 and 80 dB SPL) showed that wave V latency increased and wave V amplitude decreased systematically as the intensity of the noise increased. This finding is consistent with past research characterizing ABR behavior in noise (Burkhard and Hecox, 1983).

2.5. Cortical potentials

Cortical-evoked responses reflecting behavioral auditory perception (Cunningham et al., 2000) were obtained using Cz active, forehead ground and nasal reference electrodes. Eye movements were monitored and on-line rejected using a bipolar supraorbital-to-lateral canthus electrode montage. The recording window was 75 ms pre-stimulus and 500 ms post-stimulus, with a sampling rate of 20 kHz. Responses were off-line bandpass filtered from 0.1 to 100 Hz. Speech stimuli were presented with an interstimulus interval of 550 ms. Approximately 1000 responses were averaged for each condition.

2.5.1. Data analysis

Cortical-evoked responses elicited by the /da/ stimulus displayed a double-peaked response labeled P1 and P1′, the latter of which was more prominent (see Fig. 6). The latencies (P1, N1, P1′ and N1′) and amplitudes (P1-to-N1

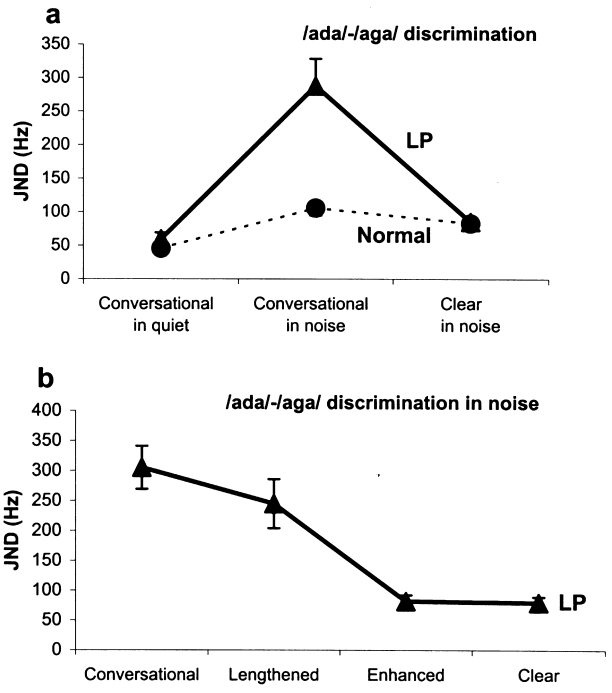

Fig. 2. (a) Mean behavioral JNDs and standard errors for normal children ($n = 9$) and LP children ($n = 9$) along /ada/ to /aga/ continua modeled according to the acoustic characteristics of 'conversational' and 'clear' speech. LP children demonstrated poorer speech discrimination than normal children in noise and regained normal performance with cue enhanced stimuli. (b) Mean behavioral JNDs and standard errors for children with LP ($n = 8$) along 4 /ada/ to /aga/ continua in noise (conversational, lengthening stop gap duration (lengthened), increasing release burst intensity (enhanced) and clear (combination of cue enhancement features)). Speech discrimination significantly improved in the enhanced condition. Increasing the release burst intensity was determined to be the more important cue enhancement for improving perception in this population.

and P1′-to-N1′) were identified for each subject by two experienced testers who were 'blind' to the subject's diagnostic category. P1 was identified as the relative positivity occurring within the range of approximately 50–90 ms. N1 was considered the early negativity between 100 and 120 ms. P1′ was labeled as the positivity occurring between 150 and 170 ms. Finally, the large negativity within the range of 210–250 ms was labeled N1′ for this study.

All non-parametric statistical analyses were two-tailed comparisons with an a priori significance level of 0.05. Bonferroni corrections were applied to significance values for multiple comparisons.

3. Results

3.1. Behavioral perception

In the behavioral experiment, JNDs were obtained to the 'conversational' speech continuum in quiet and in noise and to the 'clear' speech continuum in noise for normal and LP children. As expected, a comparison between 'conversational' speech in quiet and in noise scores indicated that both groups showed more difficulty discriminating speech in noise (Wilcoxon signed ranks test: normal, $T^+ = 45$, $P = 0.008$; LP, $T^+ = 45$, $P = 0.008$). That is, a comparison of 'conversational' speech in noise scores between normal and LP children indicated that the LP children were significantly worse than their normal counterparts (Mann–Whitney: $U = 4$, $P = 0.0009$). Importantly, the groups performed similarly in their perception of 'conversational' speech in quiet[1] and 'clear' speech in noise as shown in Fig. 2a.

Follow-up testing with 8 of the LP children, using the third and fourth /ada/ to /aga/ continua designed to focus on each of the 'clear' speech enhancements in isolation, (1) lengthening the stop gap duration and (2) increasing the intensity of the release burst, demonstrated that increasing the release burst intensity was the more important acoustic cue manipulation in noise. That is, in LP children the perception of stimuli with the increased burst intensity in noise was significantly better than the perception of 'conversational' stimuli in noise (Wilcoxon signed ranks test: $T^+ = 36$, $P = 0.012$). However, the perception of stimuli with lengthened stop gap duration in noise was not significantly different from the perception of 'conversational' stimuli in noise. These results are illustrated in Fig. 2b.

3.2. Electrophysiology

3.2.1. Brainstem responses

3.2.1.1. FFR. The average brain-stem response for normal children to /da/ in quiet is shown in Fig. 3. In quiet, the magnitude of the spectral components (FFT) for the response to /da/ were not significantly different between normal and LP children. In noise, normal children displayed a significant reduction in the magnitude of the spectral content of bin 1 (Wilcoxon signed ranks test: bin 1, $T^+ = 45$, $P = 0.007$), whereas bins 2 and 3 remained stable. In contrast, all 3 bins, reflecting the fundamental frequency and F1 transition (0–750 Hz), were significantly diminished in LP children in noise (Wilcoxon signed ranks test: bin 1, $T^+ = 44$, $P = 0.011$; bin 2, $T^+ = 43$, $P = 0.015$; bin 3, $T^+ = 45$, $P = 0.007$). Consequently, Fig. 4a illustrates that the normal and LP children differed

[1] This result contrasts previous behavioral findings using the stimulus pair /da/–/ga/ in quiet and may be explained by differences in the acoustic characteristics of each stimulus pair. In the behavioral discrimination of /ada/–/aga/, the listener may use acoustic information from 3 speech features to differentiate the speech-sound pair: the formant transition from the initial vowel /a/ to the consonant /d/ or /g/, the consonant release burst, and finally the formant transition from the consonant /d/ or /g/ to the final vowel /a/. In contrast, the original 'stripped down' /da/–/ga/ contrast consists of only one of these acoustic cues: the formant transition (/d/ to /a/ or /g/ to /a/). Thus, the minimal acoustic cues used in the /da/–/ga/ contrast are effective in delineating the perception of the normal and LP children in quiet, whereas the /ada/–/aga/ pair is acoustically different enough not to pose a perceptual challenge for the LP children.

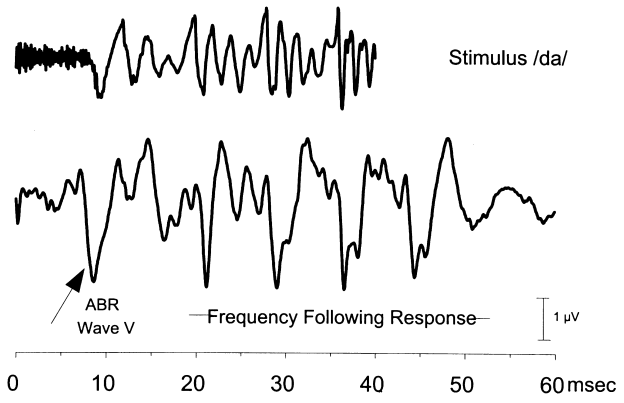

Fig. 3. Stimulus waveform (/da/) and averaged brain-stem response for normal children in quiet. Arrow indicates the onset response (wave V) of the ABR. The FFR is also shown.

significantly in the magnitude of the spectral content for bins 2 and 3 in noise (Mann–Whitney: bin 2, $U = 17$, $P = 0.031$; bin 3, $U = 7$, $P = 0.002$). In response to the enhanced /da/, the normal and LP children differed only for bin 3 ($U = 10$, $P = 0.005$), as depicted in Fig. 3b.

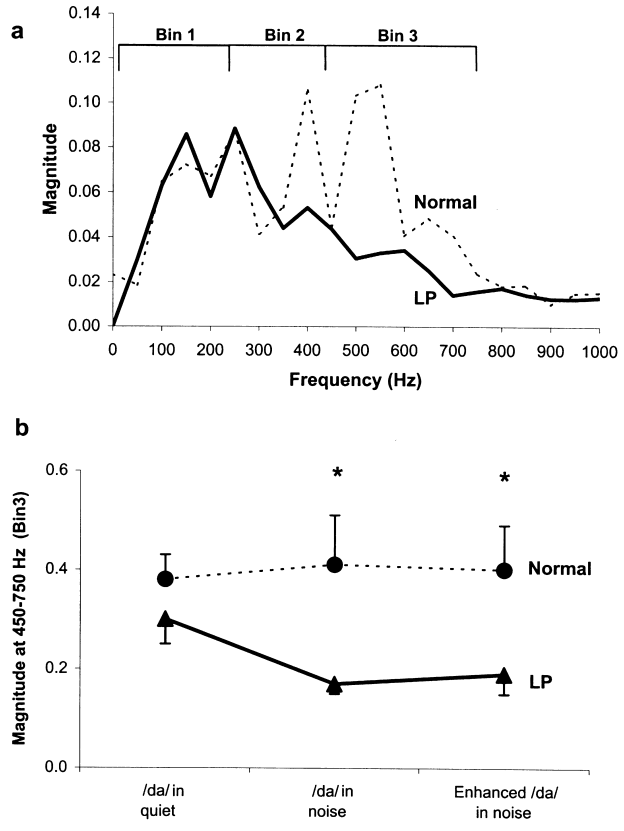

Fig. 4. (a) Magnitude of the spectral content of the FFR waveform as shown by FFT in normal and LP children to /da/ in noise. (b) Mean and standard errors of the magnitude of the spectral content of FFT components between 450 and 750 Hz for normal and LP children. The groups were significantly different in relation to both /da/ conditions in noise. No significant differences were seen in quiet.

3.2.1.2. Stimulus-to-response correlations. In quiet, the individual stimulus-to-response correlation coefficients were similar between normal and LP children. In contrast, correlation coefficients in noise were consistently lower in the LP children than in normal children (Mann–Whitney: /da/, $U = 10.5$, $P = 0.006$; enhanced /da/, $U = 4$, $P = 0.0009$). These correlation coefficients are plotted in Fig. 5. Interestingly in noise, the normal children demonstrated stability in the stimulus-to-response correlations, whereas correlations were decreased in LP children. The result seen in normal children mirrors that seen in an animal model using identical stimuli (Cunningham et al., 2001), and the beneficial rather than detrimental effect of external noise in facilitating neural activity in the auditory system (Lewis and Henry, 1995) and other sensory systems (Douglass et al., 1993; review in Wiesenfeld and Moss, 1995) is not unprecedented.

3.2.1.3. ABR. Again, the average brain-stem response for normal children to /da/ in quiet is shown in Fig. 3. In quiet, there were no significant differences in wave V latency or amplitude elicited by /da/ between normal and LP children. With the addition of background noise, both normal and LP children displayed a prolongation in wave V latency and a reduction in wave V amplitude (Wilcoxon signed ranks test: latency: normal, $T^+ = 44$, $P = 0.011$; LP, $T^+ = 45$, $P = 0.007$; amplitude: normal, $T^+ = 45$, $P = 0.007$; LP, $T^+ = 44$, $P = 0.011$). However, comparison of wave V latency to /da/ in noise between normal and LP children revealed that LP children exhibited significantly longer wave V latencies on the order of 0.41 ms (Mann–Whitney: $U = 19.5$, $P = 0.050$). There were no significant group differences in wave V amplitude in noise. In response to enhanced /da/ in noise, no significant latency and amplitude differences between normal and LP children were seen.

3.2.2. Cortical responses

With cortical-evoked responses a similar pattern of differences between normal and LP groups emerged as shown in

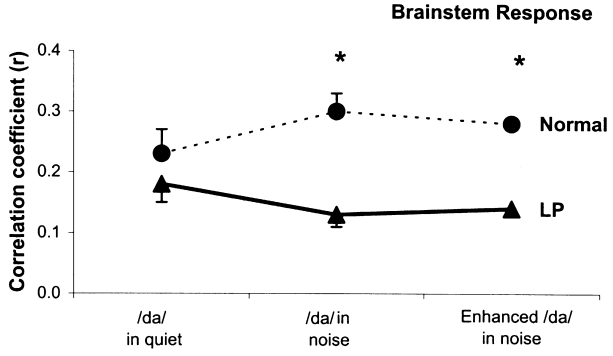

Fig. 5. Mean and standard errors of the stimulus-to-response correlation coefficient r values for normal and LP children. The groups differed significantly in both noise conditions. No significant differences were seen in quiet.

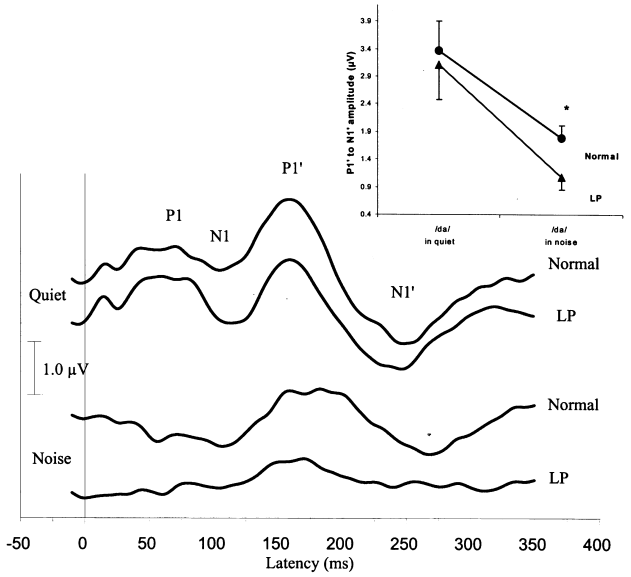

Fig. 6. Averaged cortical responses elicited by /da/ for normal and LP subjects. The top traces are the evoked responses recorded in quiet (normal children and LP children). Subsequent traces are the responses recorded in noise (normal children and LP children). In noise, the prominent P1′/N1′ component was attenuated to a greater extent in LP than in normal subjects.

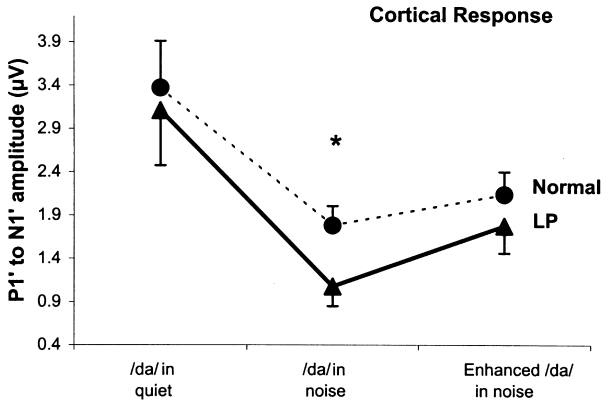

Fig. 7. Mean and standard errors of P1′-to-N1′ amplitude for normal and LP children. The groups differed significantly in P1′-to-N1′ amplitude elicited by /da/ in noise. No significant differences were seen in quiet or in relation to the enhanced /da/ in noise.

Fig. 6. In quiet, the characteristics of the cortical-evoked response complex, P1/N1/P1′/N1′ latency and amplitude did not differ between normal and LP children. In noise, both normal and LP children showed a reduction in P1-to-N1 and P1′-to-N1′ amplitude to /da/ (Wilcoxon signed ranks test: P1-to-N1 amplitude: normal, $T^+ = 40$, $P = 0.038$; LP, $T^+ = 41$, $P = 0.028$; P1′-to-N1′ amplitude: normal, $T^+ = 43$, $P = 0.015$; LP, $T^+ = 45$, $P = 0.007$) with no significant adjustment in latency. However, the P1′-to-N1′ response in noise demonstrated a significantly larger amplitude reduction in LP children compared to normal children (Mann–Whitney: $U = 16$, $P = 0.024$) (see Fig. 7). There were no significant differences between normal and LP children regarding P1′ or N1′ latency in noise. The earlier waveforms (P1 and N1) showed no group differences in latency and amplitude in noise. However, these waveforms were small to begin with and were nearly abolished in noise, thereby possibly obscuring any group differences. Finally, there were no differences between normal and LP children in latency and amplitude in response to the enhanced /da/ stimulus in noise. Thus, the group differences seen in P1′-to-N1′ amplitude in response to /da/ were eliminated by the use of a cue enhanced stimulus.

3.3. Relationship between behavioral perception and electrophysiology

Finally, it was of interest to determine whether the child's (normal and LP) ability to discriminate 'conversational' speech in noise had a relationship to their electrophysiologic responses in noise. The relationship was examined using a Spearman rank-order correlation coefficient. Results indicated that the JNDs to 'conversational' speech in noise were correlated with diminished magnitude of spectral information in the 450–750 Hz range ($r_s = -0.48$, $P = 0.050$), reduced stimulus-to-response correlations ($r_s = -0.53$, $P = 0.050$), prolonged wave V latency (Spearman $r_s = 0.52$, $P = 0.050$) and decreased amplitude of P1′-to-N1′ ($r_s = -0.59$, $P = 0.010$).

Overall, these results demonstrated that this group of LP children showed poorer speech discrimination coupled with diminished neurophysiologic responses in background noise compared to normal children. Moreover, both the perceptual and neurophysiological responses of these children were improved with cue enhanced stimuli. Table 1 provides a summary of the results of these findings.

Table 1
Summary of group differences pertaining to perceptual and neurophysiological results[a]

Response	Quiet	Noise	Cue enhanced noise
Behavioral JNDs	NL = LP	NL ≠ LP	NL = LP
FFR (FFT)	NL = LP	NL ≠ LP	NL ≠ LP
Stimulus-to-response correlation	NL = LP	NL ≠ LP	NL ≠ LP
Wave V latency	NL = LP	NL ≠ LP	NL = LP
P1′-to-N1′ amplitude	NL = LP	NL ≠ LP	NL = LP

[a] NL, normal children; LP, children with learning problems.

4. Discussion

Taken together, the electrophysiologic and behavioral results demonstrate a difference between these normal and LP children in the neurophysiologic representation and perception of speech in noise. Specifically, group differences arose in the JNDs of 'conversational' speech in noise, the magnitude of the spectral content in the FFR, the strength of the stimulus-to-response correlation coefficients reflected in the brain-stem response, the latency of wave V in the ABR and the amplitude of cortical potentials, P1′-to-N1′. Because there were no differences between the groups in quiet, the addition of competing noise provided the means for unveiling perceptual and neurophysiologic deficits in LP children and for understanding why such children may be particularly challenged in a real-world listening environment such as a classroom.

The neurophysiologic mechanisms that may underlie these perceptual deficits are little understood. Recent studies have shown that some individuals with LP have neurophysiologic deficits which may be seen in responses originating from the primary and non-primary divisions of the auditory cortex (Kraus et al., 1996; Nagarajan et al., 1999). The present study provides evidence that in certain LP children impaired neurophysiological mechanisms also may reside within sub-cortical auditory areas.

The FFR reflects the summation of neural activity synchronized to the frequency components and integral multiples of an auditory stimulus (Sohmer et al., 1977; Dobie and Wilson, 1984; Wilson and Dobie, 1987). The response can be elicited by a pure tone (Marsh et al., 1972; Moushegian et al., 1973, 1978; Bledsoe and Moushegian, 1980; Bledsoe et al., 1982), a temporally complex stimulus (Greenberg et al., 1987; Galbraith, 1994), vowels (Greenberg, 1980; Galbraith et al., 1998) and digitized words (Galbraith et al., 1995, 1997). Evidence pointing to a neural (brain-stem) generator source for the FFR is supported by the response latency (generally 6 ms), the upper frequency limit of the response (2 kHz), and its behavior under conditions of noise masking, eighth nerve sectioning, cooling of the cochlear nucleus and anoxia (Marsh et al., 1970, 1972; Moushegian et al., 1978; Bledsoe and Moushegian, 1980). Specifically, data from both an animal model and human studies using subjects with localized brain-stem lesions have implied neural involvement in the higher brain-stem centers located near the inferior colliculus (Sohmer et al., 1977; Smith et al., 1975). However, FFR recordings elicited by high intensity stimuli like those used in this study may also contain a far-field cochlear microphonic response (Sohmer et al., 1977; Stillman et al., 1978). This potential reflects the electrical activity generated by cochlear receptor cells in response to an auditory stimulus.

The observed reduction of the high frequency spectrum and diminished integrity of the FFR in LP children suggests that mechanisms within the brain-stem or cochlea may be responsible for the impaired representation of acoustic information. The importance of neural synchrony in representing sensory events has been established. Internally-generated synchronization of neural discharges has been linked to sensory/motor/cognitive processing that may be independent or loosely coupled to external stimulus events (Riehle et al., 1997; Stopfer et al., 1997; Barinaga, 1998). Nevertheless, it is the neural synchrony directly elicited by external stimulation that is considered here. In the auditory system, synchronous firing across a neural ensemble is important for the representation of temporally dynamic acoustic events prevalent in the speech signal, encoding differences between steady-state and dynamic stimuli (Eggermont, 1990), and is a means for preserving the representation of signal features at high stimulus levels and in noise (Sachs et al., 1983; Delgutte and Kiang, 1984). In an impaired system, neurons within a population may fire at different onset latencies or phase-lock imprecisely, thus forming an incomplete or 'blurred' neural representation of spectral events and fewer high frequency harmonics in noise.

Evoked potentials currently are the only way to evaluate neural synchrony in people. By definition, evoked responses require synchronous firing of neural ensembles. Consequently, these measures appear to be well suited for reflecting the response patterns of neurons responsible for encoding the acoustic complexities of speech in the normal auditory system. In addition, evoked potentials may provide diagnostic information for separating children with preconscious, auditory-based deficits from those with higher level processing problems.

This investigation also showed that the use of cue enhanced stimuli improves speech-sound perception and the neurophysiologic representation of stimulus features that were degraded in noise (summarized in Table 1). First, LP children performed significantly poorer than normal children on behavioral discrimination of 'conversational' speech in noise. Second, the LP children achieved normal performance on the same task using 'conversational' speech in quiet and 'clear' speech in noise. Cue enhanced stimuli are likely to improve speech perception either by increasing the audibility of critical acoustic features (release burst) that were masked in noise or by enhancing important phonemic contrasts. Moreover, improved speech perception may be attributed to a decrease in the exaggerated backward masking noted in the LP population (Wright et al., 1997). That is, an increased amplitude of the release burst (as seen in cue enhancement) may reduce backward masking produced by the following vowel.

The perceptual pattern prevails for the neurophysiologic data as well. LP children demonstrated significantly longer wave V latencies and smaller P1′-to-N1′ amplitudes in noise. Group differences in these evoked responses were eliminated in quiet and to enhanced stimuli in noise. The evoked responses that reflect stimulus onset encoding (ABR, cortical potentials) were improved with cue

enhanced stimuli. The only responses that did not improve with cue enhanced stimuli were the FFR (measured by the FFT and the stimulus–response correlations). Perhaps this is because the FFR predominantly reflects neural phase-locking to the stimulus fundamental frequency and spectral harmonics below 2 kHz and the release burst is composed of higher frequency energy (2580, 3600 and 4500 Hz) exceeding this upper frequency limit of this response.

This work has implications for speech-sound auditory training in LP children. Presently, commercial auditory training programs incorporate multiple cue enhancements (two of which were evaluated here) to facilitate speech-sound learning (Merzenich et al., 1996; Tallal et al., 1996). Data from the present study suggest that amplification of specific spectral regions in the speech signal improves auditory system encoding and perception, and may contribute to the success of these training programs in some children. Moreover, animal experiments have demonstrated that sensory cortex becomes restructured with training (Jenkins et al., 1990; Recanzone et al., 1993). In humans, pre-attentive neural responses to sound can be altered by short-term perceptual learning (Näätänen et al., 1993; Kraus et al., 1995; Tremblay et al., 1997, 1998, 2001). We speculate that auditory training to cue enhanced stimuli may alter the neurophysiologic responses in LP children and ameliorate speech perception deficits.

Acknowledgements

We would like to acknowledge T. McGee, D. Koch, C. King, the children, and their families for their valuable contributions. This work was supported by NIH-NIDCD-DC01510.

References

Barinaga M. Listening in on the brain (news). Science 1998;280:376–378.

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

Bledsoe Jr SC, Moushegian G. The 500 Hz frequency-following potential in kangaroo rat: an evaluation with noise masking. Electroenceph clin Neurophysiol 1980;48:654–663.

Bledsoe Jr SC, Rupert AL, Moushegian G. Response characteristics of cochlear nucleus neurons to 500 Hz tones and noise: findings relating to frequency-following potentials. J Neurophysiol 1982;47:113–127.

Boston J, Moller A. Brainstem auditory-evoked potentials. CRC Crit Rev Biomed Eng 1985;13:97–123.

Bradlow AR, Kraus N, Nicol TG, McGee TJ, Cunningham J, Zecker SG, Carrell TD. 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 1999;106:2086–2096.

Burkhard R, Hecox K. The effect of broadband noise on the human brainstem auditory evoked response. I. Rate and intensity effects. J Acoust Soc Am 1983;74:1204–1213.

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

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

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

Cunningham J, Nicol T, Kraus N. Neurophysiologic representation of clear speech in noise. Assoc Res Otolarygol 2001 in press.

Delgutte B, Kiang NY. Speech coding in the auditory nerve. V. Vowels in background noise. J Acoust Soc Am 1984;75:908–918.

Denneberg VH. Speech perception deficits in poor readers: auditory processing or phonological coding? J Learn Disabil 1999;32:379–383.

Dobie RA, Wilson MJ. Short-latency auditory responses obtained by cross correlation. J Acoust Soc Am 1984;76:1411–1421.

Douglass JK, Wilkens L, Pantazelou E, Moss F. Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance. Nature 1993;365:337–340.

Eggermont JJ. The correlative brain: theory and experiment in neural interaction. Studies of brain function, Berlin: Springer, 1990.

Elliott LL. Performance of children aged 9 to 17 on a test of speech intelligibility in noise using sentence material with controlled word predictability. J Acoust Soc Am 1979;66:651–653.

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

Galbraith GC. Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. Electroenceph clin Neurophysiol 1994;92:321–330.

Galbraith GC, Arbagey PW, Branski R, Comerci N, Rector PM. Intelligible speech encoded in the human brain stem frequency-following response. NeuroReport 1995;6:2363–2367.

Galbraith GC, Jhaveri SP, Kuo J. Speech-evoked brainstem frequency-following responses during verbal transformations due to word repetition. Electroenceph clin Neurophysiol 1997;102:46–53.

Galbraith GC, Bhuta SM, Choate AK, Kitahara JM, Mullen Jr TA. Brain stem frequency-following response to dichotic vowels during attention. NeuroReport 1998;9:1889–1893.

Gordon-Salant S. Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing. J Acoust Soc Am 1986;89:1599–1607.

Greenberg S. Temporal neural coding of pitch and vowel quality. UCLA Working Papers in Phonetics 1980;52:1–183.

Greenberg S, Marsh JT, Brown WS, Smith JC. Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. Hear Res 1987;25:91–114.

Hazan V, Simpson A. The effect of cue-enhancement on the intelligibility of nonsense word and sentence materials presented in noise. Speech Commun 1998;24:211–226.

Jenkins WM, Merzenich MM, Ochs MT, Allard T, Guic-Robles E. Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. J Neurophysiol 1990;63:82–104.

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

Kraus N, McGee T, Carrell T, King C, Tremblay K. Central auditory system plasticity associated with speech discrimination training. J Cogn Neurosci 1995;7:27–34.

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

Kraus N, Koch DB, McGee TJ, Nicol TG, Cunningham J. Speech-sound discrimination in school-age children: psychophysical and neurophysiologic measures. J Speech Lang Hear Res 1999;42:1042–1060.

Lewis ER, Henry KR. Nonlinear effects of noise on phase-locked cochlear-nerve responses to sinusoidal stimuli. Hear Res 1995;92:1–16.

Marsh JT, Worden FG, Smith JC. Auditory frequency-following responses

in the central auditory pathway: neural or artifact? Science 1970;169:1222–1223.
Marsh JT, Smith JC, Worden FG. Receptor and neural response to auditory masking of low frequency tones. Electroenceph clin Neurophysiol 1972;32:63–74.
McAnally KI, Stein JF. Scalp potentials evoked by amplitude-modulated tones in dyslexia. J Speech Lang Hear Res 1997;40:939–945.
Merzenich MM, Jenkins WM, Johnston P, Schreiner C, Miller SL, Tallal P. Temporal processing deficits of language-learning impaired children ameliorated by training. Science 1996;271:77–81.
Moller A. Neural generators of auditory evoked potentials. In: Jacobson JT, editor. Principles and applications in auditory evoked potentials, Boston, MA: Allyn and Bacon, 1994. pp. 23–46.
Moushegian G, Rupert AL, Stillman RD. Laboratory note. Scalp-recorded early responses in man to frequencies in the speech range. Electroenceph clin Neurophysiol 1973;35:665–667.
Moushegian G, Rupert AL, Stillman RD. Evaluation of frequency-following potentials in man: masking and clinical studies. Electroenceph clin Neurophysiol 1978;45:711–718.
Näätänen R, Schroger E, Karakas S, Tervaniemi M, Paavilainen P. Development of a memory trace for a complex sound in the human brain. NeuroReport 1993;4:503–506.
Nabelek AK, Pickett JM. Reception of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids. J Acoust Soc Am 1974;56:628–639.
Nagarajan S, Mahncke H, Salz T, Tallal P, Roberts T, Merzenich MM. Cortical auditory signal processing in poor readers. Proc Natl Acad Sci USA 1999;96:6483–6488.
Nittrouer S. Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. J Phonetics 1992;20:1–32.
Picheny MA, Durlach NI, Braida LD. Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech. J Speech Hear Res 1985;28:96–103.
Picheny MA, Durlach NI, Braida LD. Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. J Speech Hear Res 1986;29:434–446.
Recanzone GH, Schreiner CE, Merzenich MM. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 1993;13:87–103.
Riehle A, Grun S, Diesmann M, Aertsen A. Spike synchronization and rate modulation differentially involved in motor cortical function. Science 1997;278:1950–1953.
Sachs MB, Voigt HF, Young ED. Auditory nerve representation of vowels in background noise. J Neurophysiol 1983;50:27–45.
Smith JC, Marsh JT, Brown WS. Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. Electroenceph clin Neurophysiol 1975;39:465–472.
Sohmer H, Pratt H, Kinarti R. Sources of frequency following responses (FFR) in man. Electroenceph clin Neurophysiol Suppl 1977;42:656–664.
Stark RE, Heinz JM. Vowel perception in children with and without language impairment. J Speech Hear Res 1996;39:860–869.
Starr A, Don M. Brain potentials evoked by acoustic stimuli. In: Picton TW, editor. Handbook of electroencephalography and clinical neurophysiology, Amsterdam: Elsevier, 1988. pp. 97–150.
Stillman RD, Crow G, Moushegian G. Components of the frequency-following potential in man. Electroenceph clin Neurophysiol 1978;44:438–446.
Stopfer M, Bhagavan S, Smith BH, Laurent G. Impaired odour discrimination on desynchronization of odour-encoding neural assemblies. Nature 1997;390:70–74.
Studdert-Kennedy M, Mody M. Auditory processing deficits in the reading impaired: a critical review of the evidence. Psychonom Bull Rev 1995;2:508–514.
Tallal P. Auditory temporal perception: phonics and reading disabilities in children. Brain Lang 1980;9:182–188.
Tallal P, Piercy M. Developmental aphasia: rate of auditory processing and selective impairment of consonant perception. Neuropsychologia 1974;12:83–93.
Tallal P, Miller SL, Bedi G, Byma G, Wang X, Nagarajan SS, Schreiner C, Jenkins WM, Merzenich MM. Language comprehension in language-learning impaired children improved with acoustically modified speech. Science 1996;271:81–84.
Taylor MM, Creelman CD. PEST: efficient estimates on probability functions. J Acoust Soc Am 1967;41:782–787.
Tremblay K, Kraus N, Carrell TD, McGee T. Central auditory system plasticity: generalization to novel stimuli following listening training. J Acoust Soc Am 1997;102:3762–3773.
Tremblay K, Kraus N, McGee T. The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. NeuroReport 1998;9:3557–3560.
Tremblay K, Kraus N, McGee TJ, Ponton C, Otis B. Central auditory plasticity: changes in N1/P2 complex following speech-sound training. Ear Hear 2001 in press.
Walley A, Carrell T. Onset spectra and format transitions in the adult's and child's perception of place of articulation in stop consonants. J Acoust Soc Am 1983;73:1011–1022.
Wiesenfeld K, Moss F. Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs. Nature 1995;373:33–36.
Wilkinson G. Wide range achievement test-3, Wilmington, DE: Jastak, 1993.
Wilson MJ, Dobie RA. Human short-latency auditory responses obtained by cross-correlation. Electroenceph clin Neurophysiol 1987;66:529–538.
Woodcock R, Johnson M. Woodcock–Johnson Psycho-Educational Battery – Revised. Tests of Cognitive Ability, Allen, TX: DLM Teaching Resources, 1977.
Woodcock R, Johnson M. Woodcock–Johnson Psycho-Educational Battery – Revised. Tests of Cognitive Ability, Allen, TX: DLM Teaching Resources, 1989.
Worden FG, Marsh JT. Frequency-following (microphonic-like) neural responses evoked by sound. Electroenceph clin Neurophysiol 1968;25:42–52.
Wright BA, Lombardino LJ, King WM, Puranik CS, Leonard CM, Merzenich MM. Deficits in auditory temporal and spectral resolution in language-impaired children. Nature 1997;387:176–178.