ELSEVIER

Hearing Research 169 (2002) 97–111

HEARING RESEARCH

www.elsevier.com/locate/heares

Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex

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

[a] *Electrophysiology Laboratory, House Ear Institute, 2100 West Third Street, Los Angeles, CA 90057, USA*
[b] *Department of Communication Sciences, Northwestern University, Evanston, IL, USA*
[c] *Department of Neurobiology and Physiology, Northwestern University, Evanston, IL, USA*
[d] *Department of Otolaryngology, Northwestern University, Evanston, IL, USA*

Received 22 August 2001; accepted 11 February 2002

Abstract

Speech perception depends on the auditory system's ability to extract relevant acoustic features from competing background noise. Despite widespread acknowledgement that noise exacerbates this process, little is known about the neurophysiologic mechanisms underlying the encoding of speech in noise. Moreover, the relative contribution of different brain nuclei to these processes has not been fully established. To address these issues, aggregate neural responses were recorded from within the inferior colliculus, medial geniculate body and over primary auditory cortex of anesthetized guinea pigs to a synthetic vowel–consonant–vowel syllable /ada/ in quiet and in noise. In noise the onset response to the stop consonant /d/ was reduced or eliminated at each level, to the greatest degree in primary auditory cortex. Acoustic cue enhancements characteristic of 'clear' speech (lengthening the stop gap duration and increasing the intensity of the release burst) improved the neurophysiologic representation of the consonant at each level, especially at the cortex. Finally, the neural encoding of the vowel segment was evident at subcortical levels only, and was more resistant to noise than encoding of the dynamic portion of the consonant (release burst and formant transition). This experiment sheds light on which speech-sound elements are poorly represented in noise and demonstrates how acoustic modifications to the speech signal can improve neural responses in a normal auditory system. Implications for understanding neurophysiologic auditory signal processing in children with perceptual impairments and the design of efficient perceptual training strategies are also discussed.

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

1. Introduction

In everyday listening situations, accurate speech perception depends on the auditory system's ability to process complex acoustic elements in the presence of background noise. Studies evaluating the perception of specific speech features in background noise have shown that consonants are more vulnerable to the masking effects of noise because they are typically lower in intensity than vowels. In particular, place of articulation cues are most susceptible to perceptual confusion (Miller and Nicely, 1955). However, because consonants provide most of the acoustic information needed for word meaning, their role is essential for optimal speech intelligibility.

In poor acoustic environments (i.e. background noise, reverberation, or when the listener is hearing impaired or a non-native speaker), speakers naturally alter the acoustic characteristics of their speech in order to

* Corresponding author. Tel.: +1 (213) 483 4431; Fax: +1 (213) 413 6739.
E-mail address: jcunningham@hei.org (J. Cunningham).

Abbreviations: LP, children with learning problems; ABR, auditory brainstem response; CV, consonant–vowel syllable; FFT, fast Fourier transform; SNR, signal-to-noise ratio; HL, hearing level; RMS, response amplitude calculated by root-mean-squared; SOA, stimulus onset asynchrony; ISI, interstimulus interval; SR, spontaneous rate (auditory nerve fibers); STR, stochastic resonance; IC, inferior colliculus; MG, medial geniculate body; AC, auditory cortex

0378-5955/02/$ – see front matter
PII: S0378-5955(02)00344-1

preserve speech clarity. Increasing the intensity of the release burst and lengthening the stop gap duration (silence) prior to a consonant are two acoustic modifications of stop consonants that are predominantly used to produce 'clear' speech (Picheny et al., 1986). The perceptual benefits of cue enhanced speech are well documented (Picheny et al., 1985; Gordon-Salant, 1986; Hazan and Simpson, 1998); yet because studies evaluating cue enhanced speech have been limited to a psychophysical design, the neural representation of these signals remains to be determined.

The neurophysiologic mechanisms governing these processes in background noise, both in normal listeners and especially in individuals with perceptual impairments, are largely unknown. In a recent study, the neurophysiologic representation of 'conversational' and 'clear' /ada/ stimuli in noise was measured in a group of normal children and a group of children with learning problems (LP children) (Cunningham et al., 2001). The decision to include the stop consonant /d/ and a brief formant transition of 40 ms in the design of the stimuli was based on reports that LP children have difficulty perceiving stop consonants in general (reviewed in Bradlow et al., 1999) and /da/–/ga/ specifically (Kraus et al., 1996). Neurophysiologic responses obtained in the LP children revealed abnormalities in the fundamental sensory representation of sound in noise at brainstem and cortical levels. Specifically, LP children lacked precision in the neural representation of key stimulus features (e.g. onset and harmonic elements). Importantly, the study also demonstrated that specific cue enhancement improved the deficient neural response to speech in noise in this population. These findings provide physiologic evidence to support the use of exaggerated stimuli as an effective tool in the remediation of children with language impairments (Merzenich et al., 1996; Tallal et al., 1996).

In accordance with understanding the physiologic properties of the normal auditory system and developing hypotheses about possible impaired neural mechanisms in the LP population, the major goal of the current study was to characterize speech encoding in noise within specific brain nuclei in an animal model using the identical speech-like stimuli that proved to be efficacious in distinguishing between the two groups of children. Intracranial recordings provide direct information about level-specific speech-sound processing which can only be inferred from far-field scalp recordings in humans. Moreover, the favorable signal/noise characteristics of intracranial responses make it possible to distinguish neural activity in response to transient, harmonic and steady-state aspects of the speech stimuli to a greater degree than is possible in humans.

The first aim was to examine aspects of the neural representation of a speech-like stimulus composed of consonant and vowel elements in background noise. Specifically, the neurophysiologic response that is well synchronized to the stimulus and across neurons was considered. The following specific questions were addressed. (1) Does the neural encoding of the consonant (release burst and formant transition) suffer a greater degree of degradation than vowel segments in noise? (2) Does the effect of noise on the neural encoding of speech differ across levels of the primary auditory pathway (inferior colliculus (IC), medial geniculate body (MG) and primary auditory cortex (AC))?

The second aim was to determine whether the neural representation of acoustic elements that are degraded in noise can be improved by specific cue enhancements. The following questions were addressed. (1) Does the neural encoding of speech in noise improve with acoustic cue enhancements? (2) Does the effect of cue enhancement differ across levels of the primary auditory pathway?

Taken together, we hypothesize that cortical onset responses will show a greater effect of background noise and cue enhancement compared to onset responses from the inferior colliculus and medial geniculate. This hypothesis is based on recent findings that there is a maintenance and possible improvement in temporal coding at the cortex compared to the auditory nerve (Heil and Irvine, 1997). This process may result from a greater degree of neural inhibition and the convergence of multiple inputs at the cortex. In combination, these mechanisms may limit neural jitter and make onsets and offsets more salient (Phillips, 1995). Moreover, similar to findings in normal children (frequency following response), we hypothesize that background noise and cue enhancement will have minimal effect on later components of the response (steady-state portion).

The answers to these questions should determine which speech-sound elements are more vulnerable to noise, how acoustic modifications to the speech signal can improve neural responses, and the contribution of different auditory pathway nuclei in these processes.

2. Materials and methods

2.1. Animals and electrode placement

Animals were 11 albino guinea pigs, each weighing between 300 and 450 g. The eight males ranged in age from 27 to 42 days and three females from 39 to 45 days. All animals had normal hearing thresholds measured by auditory brainstem response (ABR). The ABR was elicited by a click stimulus at 70, 40 and 10 dB hearing level (HL) (re: normal guinea pig click thresholds) from a recording site located at the posterior vertex/midline of the scalp using an EMG needle electrode.

Animals were anesthetized with ketamine hydrochloride (100 mg/kg i.m.) and xylazine (7 mg/kg i.m.). The anesthetic state was maintained for the entire duration of the experiment by the administration of supplementary doses (15 mg/kg of ketamine; 3 mg/kg of xylazine), typically given every hour. Body temperature was maintained at 37 ± 1°C. Experimentation was approved by the Northwestern University Animal Care and Use Committee.

A rostro-caudal incision was made along the scalp surface and the tissue was retracted to expose the skull. Holes were drilled in the skull under an operating microscope and tungsten needle (1 MΩ, 25 μm tip) recording microelectrodes were placed stereotaxically within the left IC (central nucleus) and the left MG (ventral division). A 500 kΩ electrode has a recording range of 1 mm^3 in volume, therefore it can be assumed that a 1 MΩ electrode records from a volume less than 1 mm^3 (Kraus et al., 1994; King et al., 1999). For recording IC responses, locations were approximately 0.3 mm caudal to the interaural line, 1.5 mm left of the sagittal suture and 4.0 mm ventral to the surface of the brain. For recording MG responses, locations were approximately 4.8 mm rostral to the interaural line, 4.0 mm left of the sagittal suture and 7.2 mm ventral to the surface of the brain. The placement of the IC and MG electrodes was performed using audiovisual assessment. Specifically, neural activity was elicited by a 70 dB HL click stimulus and delivered to a loud speaker outside the test booth. Depth electrode penetration was halted when the click-evoked neural response became audible and an averaged electrophysiologic response, characteristic of that primary auditory region, was recorded (Kraus et al., 1994; McGee et al., 1992; King et al., 1996, 1999, 2000). Visual inspection of the response size and waveform morphology was considered. If the response was small in amplitude and broad in shape, electrode penetration was continued. This process was repeated until the morphology of the waveform conformed to the large amplitude, sharp onset response commonly observed in recordings obtained from the central nucleus of the IC and the ventral division of the MG. Previous studies (King et al., 1996, 1999, 2000) have shown a 100% hit rate for the central nucleus of the IC and the ventral division of the medial geniculate using the stereotaxic and physiologic criteria described above.

Epidural silver bead electrodes (10 kΩ, 0.5 mm diameter) were used to record surface responses over the left primary AC using procedures previously described (Kraus et al., 1988). The position of the surface electrode was approximately 3 mm caudal to bregma and 10 mm left of the sagittal suture. The reference electrode was placed over the olfactory bulb approximately 15 mm rostral to bregma and 1 mm to the right of the sagittal suture. Previous experience has shown that minimal auditory activity is recorded at this location. The ground electrode was placed on the posterior scalp surface.

2.2. Stimuli

Two five-formant synthetic /ada/ speech syllables (one 'conversational', one 'clear') were constructed using a Klatt cascade-parallel formant synthesizer (Klatt, 1980). The two /ada/ stimuli shared the following characteristics: the first /a/ vowel was 50 ms 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 for the initial and final /a/ were 720, 1240, 2500, 3600 and 4500 Hz, for F1 through F5. Starting frequencies for the consonant were 220, 1700, 2580, 3600 and 4500 Hz for F1 through F5. Stimuli had a fundamental frequency of 125 Hz. Voicing amplitude was stable across the stimulus. Differences in the construction of the two /ada/ stimuli centered on the length of the stop gap duration and the intensity of the release burst. Specifically, the first /ada/ stimulus was modeled according to the acoustic characteristics of 'conversational' speech while the second /ada/ was designed with two cue enhancements characteristic of 'clear' stop consonants (Picheny et al., 1986). Natural speech production from four 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. The 'conversational' stimulus had a stop gap duration of 50 ms, the consonant release burst to vowel intensity ratio (CV ratio) was −18 dB and total stimulus duration of 230 ms shown in Figs. 1a and 2. For the 'clear' stimulus, the stop gap duration was increased to 130 ms, the release burst intensity was increased by 10 dB so that the CV ratio was −8 dB and total stimulus duration was 310 ms shown in Fig. 2.

2.3. Procedure

Aggregate neural responses were collected from the three electrode locations (IC, MG and AC) simultaneously. Responses were obtained to both 'conversational' and 'clear' stimuli in quiet and background noise. Stimuli were presented monaurally to the right ear through hollow earbars in a stereotaxic device using Etymotic ER-3 insert earphones at an overall intensity of 85 dB SPL. Continuous broadband (up to 5 kHz) Gaussian white noise was presented at 80 dB SPL.

The recording window was 75 ms pre-stimulus and

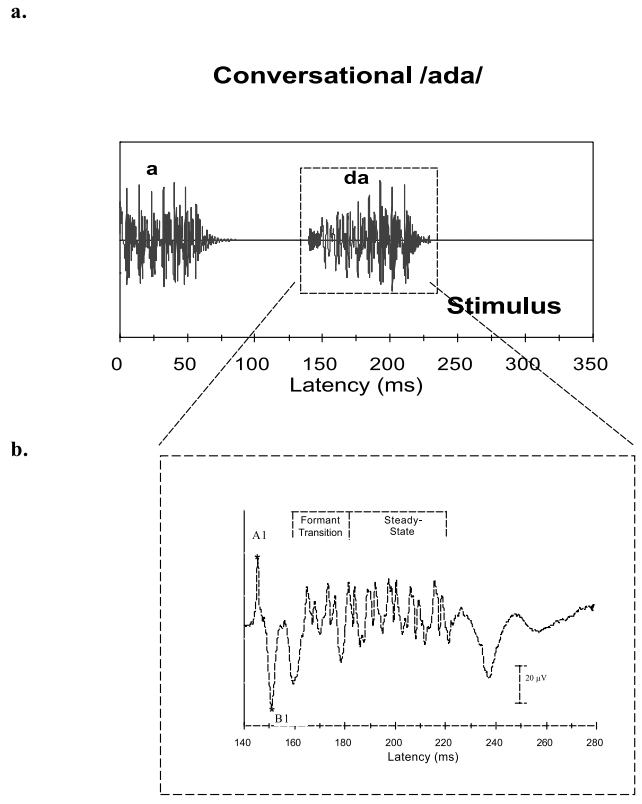

Fig. 1. (a) Stimulus waveform illustrates the acoustic characteristics of the 'conversational' /ada/ stimulus. (b) A representative trace of an evoked response recorded at the level of the IC to the 'conversational' /da/ portion of the stimulus is provided to illustrate the three components of the response (onset, formant transition, steady-state). A labels the positive peak of the onset response. B is placed at the negative trough of the onset response. The peak-to-trough amplitude is measured between points A and B. The formant transition segment is the 20 ms response following the onset response. The steady-state portion is the 50 ms segment following the formant transition.

500 ms post-stimulus, with a sampling rate of 20 kHz. Responses were on-line bandpass filtered from 0.05 to 1000 Hz (12 dB/octave). Stimuli were presented with a constant stimulus onset asynchrony (SOA) interval of 1000 ms. The selection of a constant SOA sacrificed the consistency of the interstimulus interval (ISI) for the two /ada/ stimuli. Differences in total stimulus duration of the two stimuli meant that the ISI for the 'conversational' /ada/ was 770 ms and the ISI for 'clear' /ada/ was 690 ms. Because longer ISIs are associated with larger amplitude responses (Spenner and Urbas, 1986; Arehole et al., 1987; Näätänen and Picton, 1987), we reasoned that if the amplitude of the onset response to the 'clear' /d/ was larger than the 'conversational' /d/, the amplitude differences would be attributable to the acoustic cue enhancements of 'clear' speech rather than to the length of the ISI. Each /ada/ stimulus was presented in a block of 100 presentations. Blocks were alternated until 500 responses could be averaged for each stimulus condition. Animal research has documented that ketamine and the combination of ketamine and xylazine affects auditory evoked response amplitude for several hours post-injection (Pineda et al., 1989; Dodd and Capranica, 1992; Smith and Kraus, 1987; Crowther et al., 1990). Consequently, alternation of stimulus blocks was used to control for any possible amplitude effects on the evoked responses resulting from changes in the anesthetic state of the animal.

2.4. Data analysis

The purpose of the initial vowel in the /ada/ stimuli was to permit the insertion and manipulation of the silent stop gap period (one of the 'clear' speech cue enhancements) before the stop consonant. It also served as a control measure to confirm that the initial /a/ in the 'conversational' and 'clear' stimuli yielded identical neural responses and furnished an additional control for the ISI effects discussed above. As expected, there were no significant differences in latency or amplitude between the evoked responses elicited by the initial /a/ to the 'conversational' and 'clear' stimuli either in quiet or in noise. Consequently, any differences found between responses to the 'conversational' and 'clear' /ada/ can be attributed to the intended manipulations on the /da/ portion of the stimulus.

Further evaluation of onset response attributes in quiet and in noise revealed that latencies were significantly prolonged in noise on the order of 6 ms (mean difference) at the level of the IC, 4 ms at the level of the MG and over primary AC. Because noise had a uniform effect on the response to the initial /a/ in the midbrain, thalamus and cortex, and because responses to 'conversational' and 'clear' stimuli also did not differ in response to this portion of the stimulus, any differences in responses to /ada/ can be attributed specifically to the effects of noise and/or of 'clear' speech on the neural representation of /da/.

All subsequent analyses were limited to responses to the consonant–vowel portion of the stimulus (/da/). This segment of the neural response was the region of interest because it depicted neural encoding of the consonant, the formant transition, the following vowel and cue enhanced portions of the signal. In particular, analyses focused on the examination and comparison across levels of the effects of noise and cue enhancement on neural encoding of the onset of the consonant /d/, the spectrally dynamic formant transition and the steady-state vowel.

2.4.1. Response amplitude related to stimulus onset, formant transition and steady-state

Response amplitude to the /da/ portion of the stimulus was measured using two procedures: peak-to-

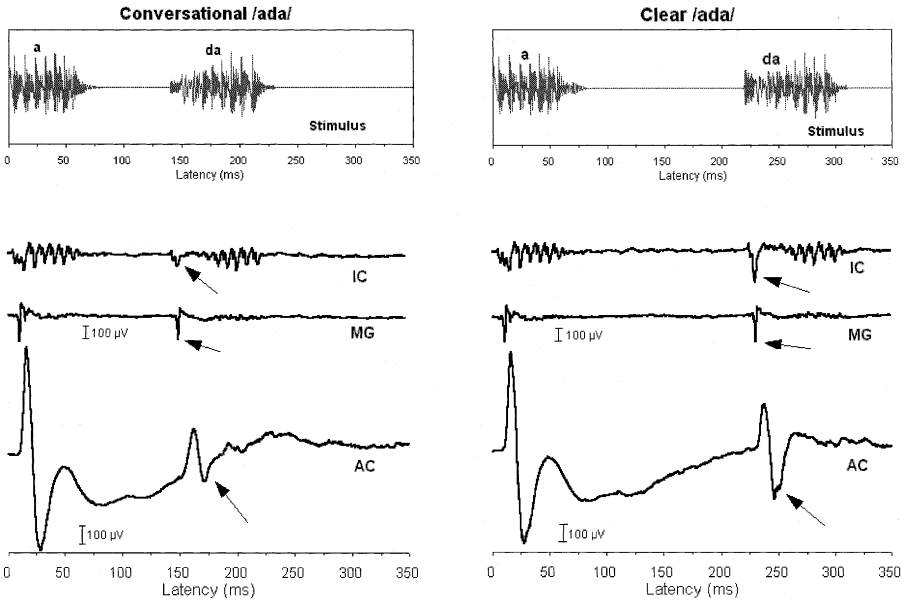

Fig. 2. Effect of 'clear' speech cue enhancements on the amplitude of the onset response to /d/ in quiet. Representative aggregate neural responses to the 'conversational' and 'clear' speech exemplars, /ada/ recorded from the IC, MG and primary AC. Arrows indicate the onset response to the consonant /d/. Onset response amplitude was significantly larger to the 'clear' /d/ compared to the 'conversational' /d/ at all levels (IC, MG, AC; $P < 0.01$). There were no significant differences in the amplitude of the onset response to initial vowel /a/ between stimulus conditions.

trough amplitude for the single-peaked onset response and root-mean-squared (RMS) measurements for multi-peaked responses. Using the first procedure, the overall amplitude between the positivity (wave A) and the subsequent negativity (wave B) of the guinea pig auditory response was obtained for the consonant /d/ onset response as shown in Fig. 1b. Onset amplitude was assessed at the IC, MG and AC. Because most onset responses elicited by 'conversational' /d/ were significantly reduced or eliminated in noise, the peak amplitude values were determined by measuring the response amplitude at a latency consistent with that observed in quiet plus the amount of latency shift observed for the initial vowel in noise (reported in Section 2.3). In the second procedure, RMS was calculated for the neural response to the formant transition (the 20 ms portion following the onset response) and steady-state vowel (50 ms segment following the formant transition) at the IC and MG. Formant transition and steady-state segments are also illustrated in Fig. 1b. These portions of the response were not seen in recordings from the level of the cortex, consistent with the expected absence of phase-locking within this frequency range in anesthetized animals (Phillips, 1995).

2.4.2. Comparison across speech features in noise

In order to determine whether neural encoding of the 'conversational' consonant (reflected in the onset response to /d/), the formant transition or the following vowel (displayed in the steady-state phase-locking) was more affected by the addition of background noise, the percent amplitude reduction was obtained for these three portions of the response.

To calculate the percent reduction value, the amplitude of the response to the 'conversational' /d/ in noise was divided by the amplitude of the response to the same stimulus in quiet, minus one. Subsequently, comparisons were made between percent reduction values for the response segments within a single level (i.e. percent reduction value for the stimulus onset versus percent reduction value for the steady-state).

2.4.3. Comparison across levels (noise and cue enhancement)

In order to determine whether there were differences along the auditory pathway pertaining to the neural encoding of the consonant onset response, the percent change in onset amplitude to the stop consonant /d/ was assessed.

First, because noise often entirely eliminated the cortical onset response in the 'conversational' stimulus, analyses were performed on responses to 'clear' stimuli only. The percent response amplitude reduction from quiet to noise was calculated for each individual subject. This value was determined by dividing the amplitude of the onset response to the 'clear' /d/ in noise by the response to 'clear' /d/ in quiet and then subtracting one.

Second, the percent increase in the onset amplitude attributable to cue enhancements was obtained by di-

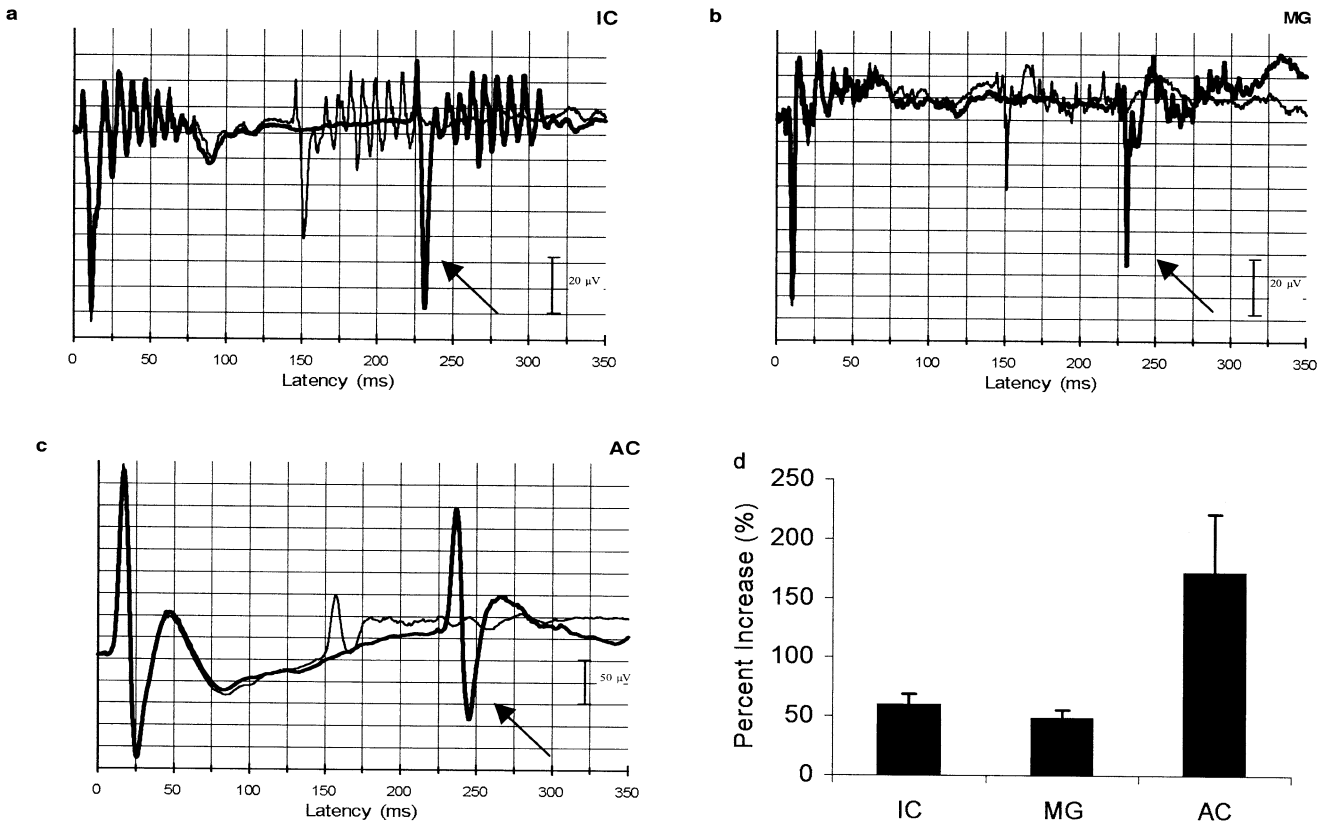

Fig. 3. Effect of 'clear' speech cue enhancements on the amplitude of the onset response to /d/. Grand average response to the 'conversational' and 'clear' speech exemplar /ada/, recorded from the (a) IC, (b) MG, and (c) primary AC in quiet. Arrows indicate the onset response to the 'clear' consonant /d/. Onset response amplitude was significantly larger to the 'clear' /d/ compared to the 'conversational' /d/ at all levels. (d) Mean and standard errors of the percent increase in the amplitude of the onset response from 'conversational' /d/ to 'clear' /d/ in quiet at the level of the IC, MG and AC. The percent increase in onset response amplitude was significantly greater over primary AC compared to subcortical levels.

viding the amplitude of the onset response to 'clear' /d/ by the amplitude of the onset response to the 'conversational' /d/ in quiet and subtracting one.

2.4.4. Magnitude of spectral content

Visual inspection of the responses at each level revealed that the IC response displayed the greatest degree of phase-locking. In order to assess the magnitude of phase-locking to the stimulus frequency components, a fast Fourier transform (FFT) of responses to 'conversational' and 'clear' /da/ was performed in quiet and in noise using a 250 ms time window (beginning at the response to the /d/ and extending to the end of the recording window) and a spectral resolution of 4 Hz. Following the generation of the FFT, the spectral components were divided into two frequency bins (4–160 Hz, 164–300 Hz) and the amplitude values within each bin were summed. Bin widths were chosen to segregate the major frequency components at 125 Hz corresponding to the stimulus fundamental frequency (bin 1) and the first formant frequencies transitioning from 220 to 720 Hz (bin 2), and were identical to bins used in a related study with LP children using the same stimuli (Cunningham et al., 2001). Subsequently, the integrated spectral magnitude values relating to stimulus ('conversational' and 'clear') and environmental conditions (quiet and noise) were compared.

2.4.5. Neural precision of phase-locking

Correlations between the final /a/ (stimulus waveform) and the final 50 ms of the IC response were obtained for each individual subject in quiet and in noise as a measure of the 'precision' of phase-locking. Specifically, the extracted portion of the response waveform was correlated with a stimulus waveform template. Correlation coefficients were successively calculated as the stimulus waveform was shifted in time relative to the response waveform. The value chosen was the maximum r value found between a 5 and 9 ms offset. This range was determined by aligning the grand-averaged final 50 ms portion of the IC response elicited by the 'conversational' and 'clear' stimuli for all subjects in both quiet and noise with the final /a/ of the stimulus waveform. This latency range encompassed the

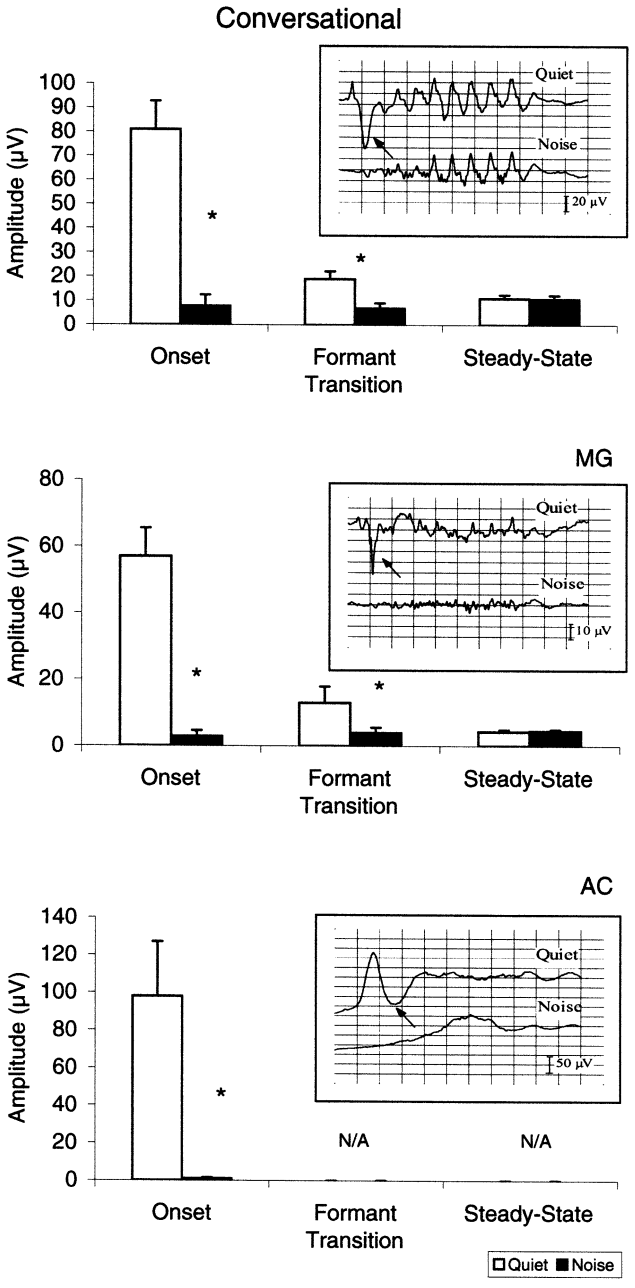

Fig. 4. Effects of noise on the amplitude of the /da/ response recorded from the IC, MG and primary AC. (Insets) Average responses to the 'conversational' speech exemplar /da/ in quiet and in noise. Arrows indicate the onset response to the consonant /d/. (Bar graphs) Mean response amplitudes and standard errors to stimulus onset, formant transition and steady-state in response to /da/ in quiet and in noise. There was a significant reduction in the onset response amplitude in noise across all levels, and in the response to the formant transition at the IC and MG. The response to the steady-state portion of the stimulus was stable in noise at the IC and MG. There was no response to the formant transition and steady-state portions of the stimulus in either quiet or noise at the cortex.

peak response in both quiet and noise conditions and is consistent with latencies reported for brainstem-generated activity (Dobie and Wilson, 1984).

2.4.6. Individual subjects

In quiet, aggregate neural responses to both stimuli were similar in all 11 subjects. However in noise, individual differences were evident to transient aspects of the stimuli at the level of the IC. Specifically, eight of the subjects demonstrated an elimination of the onset response to /d/ and response amplitude reduction to the formant transition. Visualization of the response to the onset and the formant transition in noise were obscured in three of the subjects due to the high-amplitude phase-locking to the fundamental frequency of the signal.

For statistical purposes, the subset of eight subjects was used for analyses involving transient response measures (onset and formant transition) in noise whereas analyses involving predominantly steady-state response measures included all 11 subjects. This was done to ensure that the experimental results were based on uniform data. All data were evaluated using a non-parametric statistic because the assumptions of normal distribution and equal variance in the sample could not be met. Accordingly, data were analyzed using the Wilcoxon signed ranks test, two-tailed with an a priori significance level of 0.05. Specifically, the Wilcoxon signed ranks test is used to assess the direction and magnitude of the difference between two conditions (Siegel and Castellan, 1988). Bonferroni correction factors were applied to significance values for multiple comparisons.

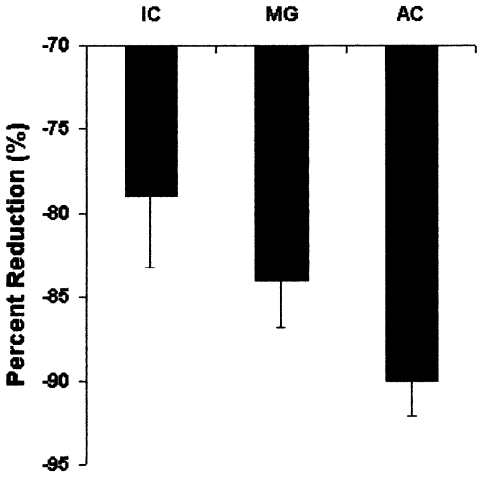

Fig. 5. Percent reduction (mean and standard errors) in the amplitude of the onset response from 'clear' /d/ in quiet to 'clear' /d/ in noise across levels. Reduction of onset response amplitude was significantly greater over primary AC compared to the subcortical levels.

3. Results

3.1. Responses in quiet

3.1.1. Response amplitude of stimulus features

Illustrated in Fig. 2 are the 'conversational' and 'clear' /ada/ stimulus waveforms and a representative trace of the electrophysiologic response to these stimuli recorded from subcortical (IC, MG) and cortical (AC) levels. In quiet, the 'conversational' consonant information encoded in the onset response to /d/ is well-represented at all levels. Acoustic features of the formant transition and vowel are encoded by phase-locking at the IC and to a lesser extent at the MG. The amplitude of the onset response to the 'clear' consonant appears larger than that of the 'conversational' consonant at all levels. No other amplitude differences were noted to responses between 'conversational' and 'clear' speech stimuli.

Fig. 3a–c depicts average responses of all subjects to 'conversational' and 'clear' /ada/ at the levels of the IC, MG and AC. Onset response amplitude was significantly larger to the 'clear' /d/ than to the 'conversational' /d/ at each level (IC: $T^+ = 66$, $P = 0.003$, MG: $T^+ = 66$, $P = 0.003$, AC: $T^+ = 66$, $P = 0.003$). There were no significant amplitude differences between neural responses elicited by the two stimuli for formant transition (IC: $T^+ = 34$, $P = 0.929$, MG: $T^+ = 53$, $P = 0.075$) or the steady-state vowel portion (IC: $T^+ = 42$, $P = 0.423$, MG: $T^+ = 37$, $P = 0.722$).

3.1.2. Comparison across levels (cue enhancement)

Comparison of the percent increase in the amplitude of the /d/ onset response at the three electrode sites attributable to 'clear' speech cue enhancements revealed that there were significant differences between subcortical and cortical levels as shown in Fig. 3d. Specifically, the percent increase at the AC was significantly larger than at the IC and MG (IC/AC: $T^+ = 59$, $P = 0.036$; MG/AC: $T^+ = 66$, $P = 0.019$) whereas the percent values at the IC were not significantly different from the MG (IC/MG: $T^+ = 35$, $P = 0.874$). Overall, these results indicated that the effect of cue enhancement was largest over the primary AC.

3.2. Effects of noise

3.2.1. Response amplitude of stimulus features

In order to determine the effect of noise on the neural encoding of specific speech features, comparisons were made between the amplitude of the response to 'conversational' /da/ in quiet and in noise for onset (release burst), formant transition, and steady-state (vowel) response segments. Results indicated that background noise significantly reduced the amplitude of the onset response to the consonant /d/ at the levels of the IC ($T^+ = 36$, $P = 0.011$), MG ($T^+ = 36$, $P = 0.011$) and AC ($T^+ = 36$, $P = 0.011$) as shown in Fig. 4. Similarly, the response to the formant transition was diminished at the IC ($T^+ = 36$, $P = 0.011$) and the MG ($T^+ = 36$, $P = 0.011$). There were no differences in the response

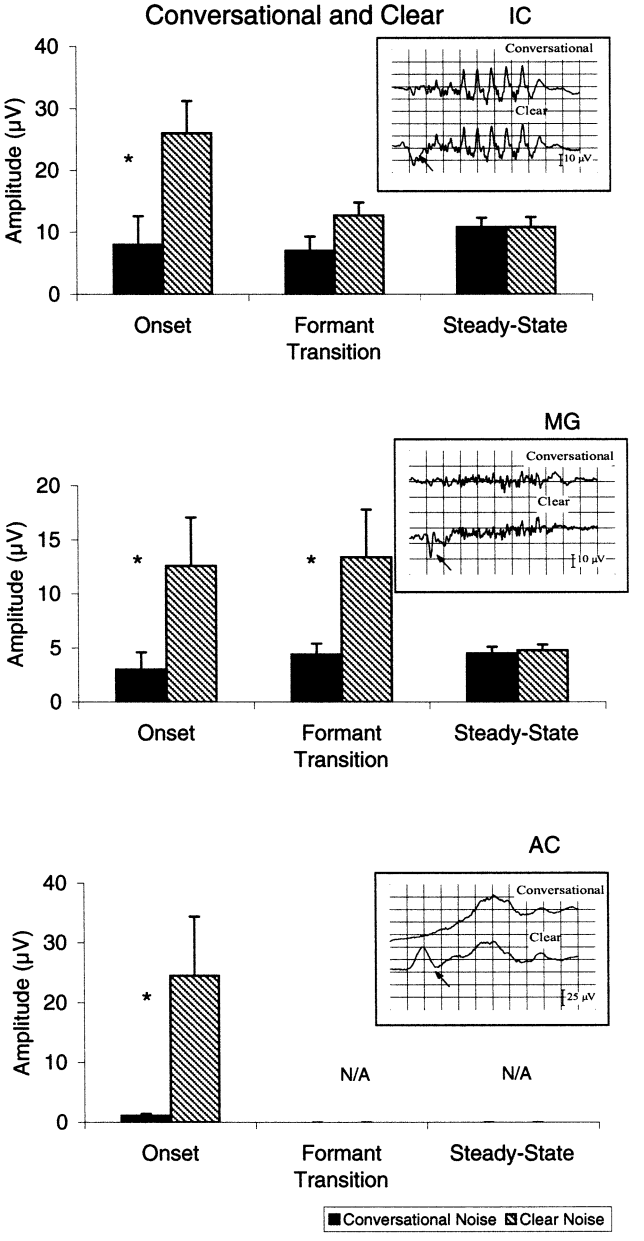

Fig. 6. Effects of acoustic cue enhancements on the amplitude of the /da/ response recorded from the IC, MG and primary AC. (Insets) Average responses to the 'conversational' and 'clear' speech exemplar /da/ in noise. Arrows indicate the onset response to the consonant /d/. (Bar graphs) Mean response amplitude and standard errors to stimulus onset, formant transition and steady-state portions of 'conversational' and 'clear' speech /da/ in noise. Cue enhancements preserved the onset response in noise at each level. The response to the formant transition was also significantly larger to cue enhanced stimuli.

amplitude from quiet to noise for the steady-state portion at the IC ($T^+ = 41$, $P = 0.476$) and the MG ($T^+ = 29$, $P = 0.722$).

3.2.2. Comparison across speech features in noise

Comparison of the percent reduction in response amplitude due to noise across speech features revealed that the amount of amplitude change was significantly larger for the onset compared to the formant transition (IC: $T^+ = 36$, $P = 0.027$, MG: $T^+ = 36$, $P = 0.027$) and the vowel (IC: $T^+ = 36$, $P = 0.027$, MG: $T^+ = 36$, $P = 0.027$). Moreover, the reduction in the response amplitude to the formant transition was also significantly greater than the vowel (IC: $T^+ = 35$, $P = 0.033$, MG: $T^+ = 33$, $P = 0.050$). Overall, the effect of noise was largest for the release burst information encoded by the stimulus onset and smallest for vowel information encoded by steady-state phase-locking.

3.2.3. Comparison across levels in noise

Comparison of the percent reduction in the amplitude of the /d/ onset response attributable to noise revealed that there were significant differences between subcortical and cortical levels as shown in Fig. 5. Specifically, the percent reduction in the onset response at the AC was significantly greater compared to the IC and MG (IC/AC: $T^+ = 36$, $P = 0.027$; MG/AC: $T^+ = 33$, $P = 0.050$), and IC and MG responses were not significantly different from each other (IC/MG: $T^+ = 29$, $P = 0.139$). Results indicated that the effect of noise was largest over the primary AC.

3.2.4. Response amplitude to 'clear' speech features in noise

Stimulus cue enhancements benefited the representation of the consonant onset response in noise at each level as shown in Fig. 6. Onset response amplitude to

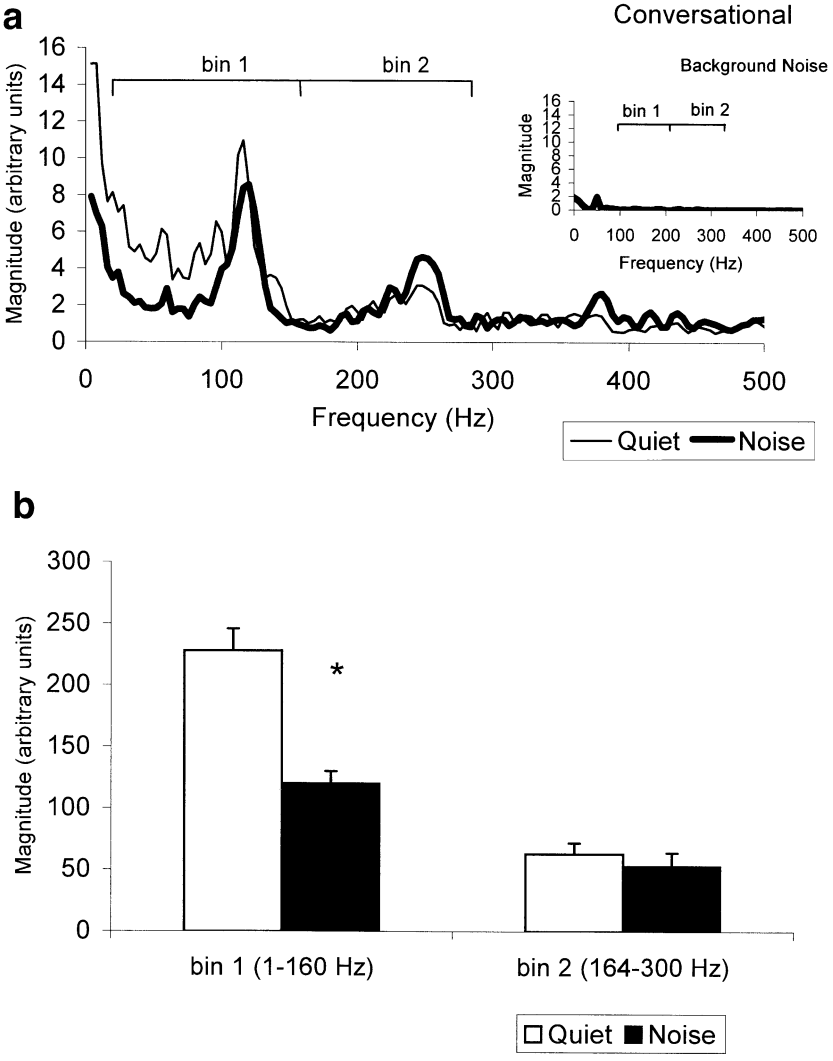

Fig. 7. (a) FFT of the averaged 'conversational' /da/ response recorded at the level of the IC in quiet and noise. (Inset) FFT of the background noise response recorded at the level of the IC. (b) Mean and standard errors represent spectral components of the FFT for two frequency bins (4–160 Hz, 164–300 Hz). There was a significant decrease in spectral energy from 1 to 160 Hz with the addition of background noise, whereas the spectral energy at the higher frequencies (164–300 Hz) remained constant.

the 'clear' /d/ was significantly larger than to 'conversational' /d/ in noise at each level (IC: $T^+ = 36$, $P = 0.011$, MG: $T^+ = 36$, $P = 0.011$, AC: $T^+ = 36$, $P = 0.011$). Moreover, the amplitude of the response to the formant transition was also significantly larger to the 'clear' stimulus at the MG ($T^+ = 32$, $P = 0.049$). In contrast, steady-state (IC: $T^+ = 42$, $P = 0.423$, MG: $T^+ = 29$, $P = 0.722$) portion showed no significant effect of cue enhancement.

3.2.5. Magnitude of spectral content in noise

In Fig. 7a, the FFT of the IC response to 'conversational' /da/ in quiet shows a major frequency component at 125 Hz corresponding to the stimulus fundamental frequency. The magnitude of the spectral content of this peak is reflected in bin 1. In addition, the magnitude of a subsequent peak at 240 Hz is seen in bin 2. Comparison of the spectral content measured in bin 1 (4–160 Hz) in quiet and in noise, revealed that the integrated spectral magnitude in this range was significantly reduced in noise (bin 1: $T^+ = 66$, $P = 0.019$). In contrast, the integrated spectral magnitude of bin 2 was stable in noise as illustrated in Fig. 7b (bin 2: $T^+ = 52$, $P = 0.091$). Comparisons of the integrated spectral magnitude measured between 'conversational' and 'clear' /da/ in noise demonstrated that the two stimuli elicited the same response pattern in noise (bin 1: $T^+ = 39$, $P = 0.593$; bin 2: $T^+ = 45$, $P = 0.286$).

3.2.6. Neural precision of phase-locking in noise

Comparison of the stimulus-to-response correlation coefficients (the correlation of the final /a/ of the 'conversational' stimulus waveform to the final 50 ms portion of the IC response) obtained in quiet and in noise revealed that the correlation coefficients were not significantly different ($T^+ = 38$, $P = 0.285$) between the two conditions as illustrated in Fig. 8. In other words, the strength of the correlation coefficients was preserved in noise. As expected, there were no significant differences between the correlation coefficients obtained to the 'conversational' and 'clear' stimuli in noise ($T^+ = 14$, $P = 0.309$).

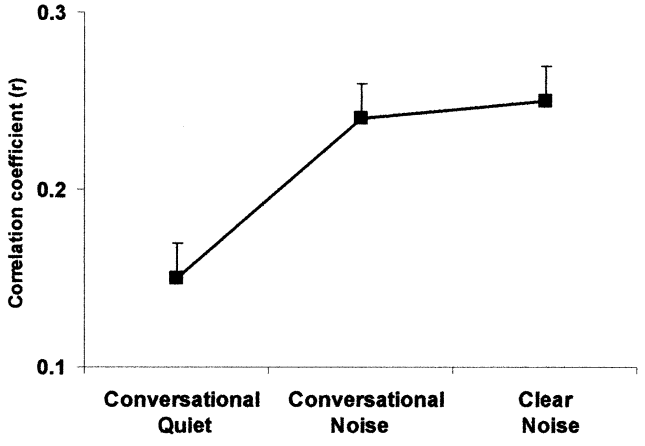

Fig. 8. Mean and standard errors represent correlations between the stimulus and response waveforms to 'conversational' /ada/ in quiet and in noise and to 'clear' /ada/ in noise at the level of the IC. Correlation coefficients were stable from quiet to noise for 'conversational' stimuli. There was no significant difference in the correlation coefficients between 'conversational' and 'clear' stimuli in noise.

4. Discussion

4.1. Transient responses

Results demonstrate that the neural representation of the stop consonant (release burst) is more vulnerable to disruption in noise than the later vowel portion. This effect was seen at each level (IC, MG and AC) and was most pronounced over AC. As the mechanisms underlying the neural encoding of consonant–vowel stimuli in noise at these levels are largely unknown, comparisons with past literature is difficult. Most studies of speech encoding have focused on responses in quiet (Young and Sachs, 1979; Sachs and Young, 1979; Sinex and Geisler, 1983; Delgutte and Kiang, 1984a,b,c; Carney and Geisler, 1986; Deng and Geisler, 1987; Steinschneider et al., 1982), vowel-in-noise encoding in the auditory nerve (Sachs et al., 1983; Delgutte and Kiang, 1984d) and consonant/vowel-in-noise encoding in the auditory nerve (Geisler and Gamble, 1989; Silkes and Geisler, 1991). Nevertheless, the examination of speech encoding at lower levels of the auditory system may aid in the interpretation of results from this experiment.

4.1.1. Subcortical

The stop consonant release is composed of transient, aperiodic acoustic energy. Most auditory nerve fibers respond with an increased discharge rate to this stimulus feature in quiet and in moderate levels of noise. Yet, at higher noise levels, the neural representation of the stop release becomes greatly reduced or eliminated (Geisler and Gamble, 1989; Silkes and Geisler, 1991). For example, in a study examining auditory nerve discharge rate profiles to /p/ and /t/, both low and high spontaneous rate (SR) fibers responded to the stop consonant onset at a 30 dB signal-to-noise ratio (SNR). As the intensity of the noise was raised (10 dB SNR), neural firing rates to the stimulus onset progressively decreased for both fiber types, but to a lesser degree for low SR fibers. Finally, at a 0 dB SNR, firing activity to the consonant onset was largely absent across fibers. Yet, a small number of low SR fibers displayed a weak onset response (Silkes and Geisler, 1991). Similar

auditory nerve fiber response patterns were also elicited by amplitude modulated signals in background noise (Frisina et al., 1996; reviewed in Frisina, 2001). The present data showed that the onset response, time-locked to the consonant release burst, was greatly reduced or eliminated in high levels of background noise (80 dB SPL; +5 dB SNR relative to the consonant) at the IC, MG and AC.

4.1.2. Cortex

Temporal processing of transient acoustic events is well-preserved at the cortex and possibly improved at the cortex (Heil and Irvine, 1997). Central timing mechanisms, mediated by neural inhibition and the convergence of multiple inputs likely contribute to the encoding of spectro-temporal aspects of consonants and serve to highlight phonetic distinctions based on voicing and place of articulation cues (Steinschneider et al., 1982). For instance, patients with bilateral primary AC lesions demonstrate impaired perception of stop consonants in conjunction with normal perception of steady-state vowels, supra-segmental aspects of speech and spectrally static consonants such as fricatives and glides (Phillips and Farmer, 1990). These data suggest that cortical areas are necessary for encoding stop consonants and further support the role of AC in the perception of transient speech elements.

In the present study, results indicated that the consonant onset response recorded at the cortex was more affected by noise than at subcortical levels. Research has shown that cortical neurons show elevations in neural thresholds and steeper rate–intensity profiles in continuous background noise. That is, once the signal has surpassed the neuronal threshold level, most cortical neurons display a rapid increase in neural discharge rate, frequently to a larger degree than to signals in quiet at the same intensity level (Phillips, 1990). In this experiment, we speculate that the greater effect of noise on onset amplitude at the cortex may be related to intensity encoding mechanisms in the cortex. However, differences in the electrodes and recording techniques used between subcortical and cortical sites cannot be ruled out as contributing to the effects observed.

4.1.3. Cue enhancement

Acoustic cue enhancements aided the neurophysiologic representation of speech by improving onset synchrony and preserving the spectral information in the release burst. Data showed that the acoustic feature that is most vulnerable in noise (release burst) may be preserved by using 'clear' speech cue enhancements (increasing the release burst intensity and lengthening the stop gap duration). The neurophysiologic benefit provided by these acoustic manipulations can be explained by basic auditory principles. First, the larger onset amplitude in response to the increased release burst intensity is a result of increased rates of neural firing and the recruitment of additional fibers from a neural population (Phillips, 1990; Brugge, 1992). Second, the larger onset amplitude in response to changes in the duration of the stop gap between the initial vowel and consonant may be attributed to a reduction in neural adaptation. By allowing additional time for neural recovery following vowel encoding, neurons would be more likely to fire in unison to the consonant onset. This effect has been shown previously using simple and complex stimuli. For instance, single unit recordings demonstrated a gradual increase in firing rate as the time interval between two stimuli was increased in eighth nerve fibers (Harris and Dallos, 1979), cochlear nucleus (Boettcher et al., 1990), and IC (Arehole et al., 1987; Walton et al., 1997; reviewed in Frisina, 2001). Similarly, data showed that the neural representation of MG neurons to stop consonants in naturally spoken German words [Stahl, Spott] could be modified by changing the stop gap duration placed between the /s/ and the /t/ or /p/ from 0 to 120 ms (Spenner and Urbas, 1986). Finally, numerous studies have shown that cortical responses are particularly susceptible to manipulation of the time interval between successive stimuli (Hocherman and Gilat, 1981; Phillips et al., 1989; Brosch and Schreiner, 1997; Brosch et al., 1999). There is a progressive increase in the effect of ISI from the cochlear nucleus to thalamus attributed to a larger amount of inhibitory properties at more central levels (Kitzes and Buchwald, 1969). While the neurophysiologic effects of increased stimulus intensity and lengthened gap duration are well-established at lower levels, the results from this study suggest that the cortex may capitalize on these effects. That is, stimulus modifications that improve the temporal precision of individual neural firing patterns can enhance neural synchrony across a population of cortical neurons, leading to larger amplitude aggregate neural responses.

Visual inspection of the 'clear' consonant onset response in noise also demonstrated that cortical onset responses benefited from cue enhancements to a greater extent than lower levels. Again, this may be explained by steeper rate–intensity profiles in continuous noise compared to quiet for cortical neurons (Phillips, 1985) and/or differences in electrodes and recording techniques between subcortical and cortical sites.

4.2. Steady-state responses

4.2.1. Subcortical

The present data demonstrated that the steady-state response to the vowel showed greater resistance to noise masking compared to transient responses. This was

evident in (1) the preservation of response amplitude from quiet to noise, (2) the stability of spectral content (FFT) and (3) the increased strength of stimulus-to-response correlation coefficients in noise at the level of the IC.

Numerous studies have shown that in quiet, auditory-nerve fibers whose characteristic frequencies are tuned to the stimulus formant frequencies show increased firing rates to the peaks of the vowel spectrum based on 'rate-place' representation (average discharge rate versus characteristic frequency) (Sachs and Young, 1979; Sachs et al., 1983). However, the rate-place representation of neuronal discharge becomes saturated for most units in noise. Accordingly, the auditory system extracts a temporal code in order to preserve the neural representation of speech in noise. For instance, the precise encoding of steady-state vowel information is conserved by phase-locking of neural impulses for the duration of the stimulus (Sachs et al., 1983; Delgutte and Kiang, 1984d) and the synchronization of firing across a population of neurons tuned to the largest harmonics of each formant frequency (Shamma, 1985). In addition, the timing of neural spikes comprising this temporal code remains strong when stimulus levels exceed the intensity of the background noise (+9 dB SNR) (Sachs et al., 1983), (+10 dB SNR) (Delgutte and Kiang, 1984d) and when both signal and noise levels are equivalent (0 dB SNR) (Geisler and Gamble, 1989). This timing synchrony is demonstrated in the stability and/or the improvement of stimulus-to-response waveform correlations in noise for low SR fibers in the eighth nerve (Silkes and Geisler, 1991) and for IC responses seen in this study.

Other specialized mechanisms may also exist within the auditory system to ensure neurophysiologic representation and perceptual saliency of the steady-state aspects of a stimulus in noise. Recent data suggest that the presence of random noise may enhance rather than reduce neural response properties in the auditory (Frisina et al., 1994; Lewis and Henry, 1995; Frisina et al., 1996) and other sensory systems (Wiesenfeld and Moss, 1995; Levin and Miller, 1996). Stochastic resonance (STR) is a phenomenon whereby the detection and neurophysiologic representation of a weak periodic signal is improved by the addition of an optimal level of external random noise. For instance, gerbil auditory-nerve neurons tuned to high frequencies (CF $\geq$ 7 kHz) were capable of vigorous phase-locking to a low-frequency tone (70 dB SPL) upon the addition of background noise (Lewis and Henry, 1995). More recently, STR was seen in the auditory-nerve envelope response elicited by toneburst stimuli (0–5 dB SPL) in the presence of wideband noise (10–15 dB SPL) (Henry, 1999). The mechanism responsible for STR has not been fully established. However, research suggests that it may originate at the receptor level in the cochlea (Jaramillo and Wiesenfeld, 1998) and for near-threshold stimulus situations be derived from the Brownian motion of cochlear fluids (Enrenberger et al., 1999). For supra-threshold stimulus situations, neural adaptation to the background noise may underlie the maintenance and in some cases the enhancement of the steady-state response by raising rate–intensity functions to a greater decibel level and decreasing maximal firing rates (Geisler and Sinex, 1980; Costalupes et al., 1984).

4.3. Relationship between neural responses to speech in noise and speech perception (normal and impaired)

Studies indicate that some LP children show difficulties identifying and discriminating speech sounds with brief, dynamic frequency elements (Tallal and Piercy, 1974; Godfrey and Millay, 1981; Elliott et al., 1989; Reed, 1989; Kraus et al., 1996) and fine-grained acoustic differences (Stark and Heinz, 1996). Moreover, LP individuals exhibit poorer speech in noise perception (Cunningham et al., 2000, 2001) and a greater degree of perceptual backward masking (Wright et al., 1997). In a recent study using similar stimuli to those used here, we report abnormalities in the fundamental sensory representation of sound at brainstem and cortical levels in LP children when speech sounds are presented in noise (Cunningham et al., 2000, 2001). Specifically, group differences were seen in the magnitude of the spectral content in the frequency following response, the strength of the stimulus-to-response correlation coefficients reflected in the brainstem response, the latency of wave V in the ABR and the amplitude of cortical potentials, P1′–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 these children may be particularly challenged in a real-world listening environment such as a classroom. The neurophysiologic effects of noise and cue enhancement seen in the normal children mirrored findings from the IC (spectral patterns and stimulus-to-response correlation).

The current data on the neural representation of speech in noise complement our human studies using identical stimuli. In noise, normal subcortical responses (ICc and MGv) are chiefly composed of the harmonic representation of the formant transition and vowel portions of the signal whereas the onset response is reduced or eliminated. We hypothesize that possible safeguards related to improved neural synchrony in noise may be built into the auditory system to preserve the harmonic representation of these speech features. This was manifested in both normal children and animals through the

preservation of higher frequency spectral components measured in the FFT and enhancement of the stimulus-to-response correlation in noise. Although consonant discrimination in noise was still difficult for normal children without the onset burst information, it could be achieved with the spectral and temporal information provided in the formant transition. In noise, the neurophysiologic response of the LP children exhibited degraded harmonic representation of the formant transition. Poor discrimination of speech in noise in these children might be explained by reduced formant transition cues in addition to the (normally occurring) reduction or absence of release-burst spectral information. Finally, the present study revealed a particular role of the AC in the encoding of transient signals. That is, onset responses in AC were most reduced in noise and most boosted with enhanced signals. Perhaps this is related to the finding that in LP children, the effects of cue enhancement were also most dramatic in cortical compared to subcortical responses.

We speculate that in addition to the reduction or absence of release burst spectral information, LP children also suffer from degraded neurophysiologic representation of formant transition information. In an impaired system, neurons within a population may fire at different onset latencies or phase-lock imprecisely to more harmonic aspects of the signal, thus forming an incomplete or 'blurred' neural representation of such events in noise. Evidence provided by this study suggests that cue enhancements in noise aid the neurophysiologic representation of speech signals by improving onset synchrony and preserving the spectral information in the release burst. In addition, in response to cue enhanced stimuli, responses are also improved for the formant transition. Possibly, the presence of the release burst representation minimizes the backward masking effect of the following vowel on the formant transition. Taken together, data from the present study suggest that amplification of specific spectral regions in the speech signal improves the neurophysiologic representation of the consonant and may identify critical elements to target for improving impaired speech perception in noise. Isolating aspects of complex signal that are crucial for sound perception in challenging listening environments may also be used in speech-sound auditory training in individuals with perceptual impairments.

Acknowledgements

This work was supported by NIH-NIDCD-DC01510. We would like to acknowledge A. Bradlow and T. McGee for their valuable contributions.

References

Arehole, S., Salvi, R.J., Saunders, S.S., Hamernik, R.P., 1987. Evoked response 'forward masking' functions in chinchillas. Hear. Res. 30, 23–32.

Boettcher, F., Salvi, R.J., Saunders, S.S., 1990. Recovery from short-term adaptation in single neurons in the cochlear nucleus. Hear. Res. 48, 125–144.

Bradlow, A.R., Kraus, N., Nicol, T.G., McGee, T.J., Cunningham, J., Zecker, S.G., Carrell, T.D., 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.

Brosch, M., Schreiner, C., 1997. Time course of forward masking tuning curves in cat primary auditory cortex. J. Neurophysiol. 77, 923–943.

Brosch, M., Schulz, A., Scheich, H., 1999. Processing of sound sequences in macaque auditory cortex: response enhancement. J. Neurophysiol. 82, 1542–1559.

Brugge, J., 1992. An overview of central auditory processing. In: Popper, A.N., Fay, R.R. (Eds.), The Mammalian Auditory Pathway: Neurophysiology. Springer-Verlag, New York, pp. 1–33.

Carney, L.H., Geisler, C.D., 1986. A temporal analysis of auditory-nerve fiber responses to spoken stop consonant-vowel syllables. J. Acoust. Soc. Am. 79, 1896–1914.

Costalupes, J.A., Young, E.D., Gibson, D.J., 1984. Effects of continuous noise backgrounds on rate response of auditory-nerve fibers in cat. J. Neurophysiol. 51, 1326–1344.

Crowther, J.A., Miller, J.M., Kileny, P.R., 1990. Effect of anesthesia on acoustically evoked middle latency response in guinea pigs. Hear. Res. 43, 115–120.

Cunningham, J., Nicol, T., Kraus, N., 2000. Neurophysiologic representation of clear speech in noise. Assoc. Res. Otolaryngol. 23, 86.

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

Delgutte, B., Kiang, N.Y., 1984a. Speech coding in the auditory nerve: I. Vowel-like sounds. J. Acoust. Soc. Am. 75, 866–878.

Delgutte, B., Kiang, N.Y., 1984b. Speech coding in the auditory nerve: III. Voiceless fricative consonants. J. Acoust. Soc. Am. 75, 887–896.

Delgutte, B., Kiang, N.Y., 1984c. Speech coding in the auditory nerve: IV. Sounds with consonant-like dynamic characteristics. J. Acoust. Soc. Am. 75, 897–907.

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

Deng, L., Geisler, C.D., 1987. Responses of auditory-nerve fibers to nasal consonant-vowel syllables. J. Acoust. Soc. Am. 82, 1977–1988.

Dobie, R.A., Wilson, M.J., 1984. Short-latency auditory responses obtained by cross correlation. J. Acoust. Soc. Am. 76, 1411–1421.

Dodd, F., Capranica, R.R., 1992. A comparison of anesthetic agents and their effects on the response properties of the peripheral auditory system. Hear. Res. 62, 173–180.

Elliott, L.L., Hammer, M.A., Scholl, M.E., 1989. Fine-grained auditory discrimination in normal children and children with language-learning problems. J. Speech Hear. Res. 32, 112–119.

Enrenberger, K., Felix, D., Svozil, K., 1999. Stochastic resonance in cochlear signal transduction. Acta Otolaryngol. 119, 166–170.

Frisina, R.D., 2001. Subcortical neural coding mechanisms for auditory temporal processing. Hear. Res. 158, 1–27.

Frisina, R.D., Walton, J.P., Karcich, K.J., 1994. Dorsal cochlear nu-

cleus single neurons can enhance temporal processing capabilities in background noise. Exp. Brain Res. 102, 160–164.
Frisina, R.D., Karcich, K.J., Tracy, T.C., Sullivan, D.M., Walton, J.P., Colombo, J., 1996. Presevation of amplitude modulation colding in the presence of background noise by chinchilla auditory-nerve fibers. J. Acoust. Soc. Am. 99, 475–490.
Geisler, C.D., Sinex, D.G., 1980. Responses of primary auditory fibers to combined noise and tonal stimuli. Hear. Res. 3, 317–334.
Geisler, C.D., Gamble, T., 1989. Responses of 'high-spontaneous' auditory-nerve fibers to consonant-vowel syllables in noise. J. Acoust. Soc. Am. 85, 1639–1652.
Godfrey, J.J., Millay, K.K., 1981. Discrimination of the 'tempo of frequency change' cue. J. Acoust. Soc. Am. 69, 1446–1448.
Gordon-Salant, S., 1986. Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing. J. Acoust. Soc. Am. 89, 1599–1607.
Harris, D.M., Dallos, P., 1979. Forward-masking of auditory nerve fiber response. J. Neurophysiol. 42, 1083–1107.
Hazan, V., Simpson, A., 1998. The effect of cue enhancement on the intelligibility of nonsense word and sentence materials presented in noise. Speech Commun. 24, 211–226.
Heil, P., Irvine, D.R., 1997. First-spike timing of auditory-nerve fibers and comparison with auditory cortex. J. Neurophysiol. 78, 2438–2454.
Henry, K.R., 1999. Noise improves transfer of near-threshold, phase-locked activity of the cochlear nerve: evidence for stochastic resonance? J. Comp. Physiol. [A] 184, 577–584.
Hocherman, S., Gilat, E., 1981. Dependence of auditory cortex evoked unit activity on interstimulus interval in the cat. J. Neurophysiol. 45, 987–997.
Jaramillo, F., Wiesenfeld, K., 1998. Mechanoelectrical transduction assisted by Brownian motion: a role for noise in the auditory system. Nature Neuroscience 1, 384–388.
King, C., Nicol, T., McGee, T., Kraus, N., 1996. Neurophysiologic representation of comodulation masking release (CMR) in the auditory midbrain and thalamus. Assoc. Res. Otolarygol. 19.
King, C., Nicol, T., McGee, T., Kraus, N., 1999. Thalamic asymmetry is related to acoustic signal complexity. Neurosci. Lett. 267, 89–92.
King, C., Nicol, T., Kraus, N., 2000. Left vs. right response asymmetry in the auditory thalamus and midbrain of the guinea pig. Assoc. Res. Otolaryngol. 23, 255.
Kitzes, M., Buchwald, J., 1969. Progressive alterations in cochlear nucleus, inferior colliculus, and medial geniculate responses during acoustic habituation. Exp. Neurol. 25, 85–105.
Klatt, D.H., 1980. Software for a cascade/parallel formant synthesizer. J. Acoust. Soc. Am. 67, 971–995.
Kraus, N., McGee, T., Carrell, T., King, C., Littman, T., Nicol, T., 1994. Discrimination of speech-like contrasts in the auditory thalamus and cortex. J. Acoust. Soc. Am. 96, 2758–2768.
Kraus, N., McGee, T.J., Carrell, T.D., Zecker, S.G., Nicol, T.G., Koch, D.B., 1996. Auditory neurophysiologic responses and discrimination deficits in children with learning problems. Science 273, 971–973.
Kraus, N., Smith, D.I., McGee, T., 1988. Midline and temporal lobe MLRs in the guinea pig originate from different generator systems: a conceptual framework for new and existing data. Electroencephalogr. Clin. Neurophysiol. 70, 541–558.
Levin, J.E., Miller, J.P., 1996. Broadband neural encoding in the cricked sensory system enhanced by stochastic resonance. Nature 380, 165–168.
Lewis, E.R., Henry, K.R., 1995. Nonlinear effects of noise on phase-locked cochlear-nerve responses to sinusoidal stimuli. Hear. Res. 92, 1–16.
McGee, T., Kraus, N, Littman, T., Nicol, T., 1992. Contributions of medical geniculate body subdivisions to the middle latency response. Hear. Res. 61, 147–154.
Merzenich, M.M., Jenkins, W.M., Johnston, P., Schreiner, C., Miller, S.L., Tallal, P., 1996. Temporal processing deficits of language-learning impaired children ameliorated by training. Science 271, 77–81.
Miller, G.A., Nicely, P.E., 1955. An analysis of perceptual confusions among some english consonants. J. Acoust. Soc. Am. 27, 338–352.
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.
Phillips, D.P., 1985. Temporal response features of cat auditory cortex neurons contributing to sensitivity to tones delivered in the presence of continuous noise. Hear. Res. 19, 253–268.
Phillips, D.P., 1990. Neural representation of sound amplitude in the auditory cortex: effects of noise masking. Behav. Brain Res. 37, 197–214.
Phillips, D.P., 1995. Central auditory processing: a view from auditory neuroscience. Am. J. Otol. 16, 338–352.
Phillips, D.P., Farmer, M.E., 1990. Acquired word deafness, and the temporal grain of sound representation in the primary auditory cortex. Behav. Brain Res. 40, 85–94.
Phillips, D.P., Hall, S.E., Hollett, J.L., 1989. Repetition rate and signal level effects on neuronal responses to brief tone pulses in cat auditory cortex. J. Acoust. Soc. Am. 85, 2537–2549.
Picheny, M.A., Durlach, N.I., Braida, L.D., 1985. Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. J. Speech Hear. Res. 28, 96–103.
Picheny, M.A., Durlach, N.I., Braida, L.D., 1986. Speaking clearly for the hard of hearing. II: Acoustic characteristics of clear and conversational speech. J. Speech Hear. Res. 29, 434–446.
Pineda, J.A., Holmes, T.C., Swick, D., Foote, S.L., 1989. Brain-stem auditory evoked potentials in squirrel monkey (*Saimiri sciureus*). Electroencephalogr. Clin. Neurophysiol. 73, 532–543.
Reed, M.A., 1989. Speech perception and the discrimination of brief auditory cues in reading disabled children. J. Exp. Child Psychol. 48, 270–292.
Sachs, M.B., Young, E.D., 1979. Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. J. Acoust. Soc. Am. 66, 470–479.
Sachs, M.B., Voigt, H.F., Young, E.D., 1983. Auditory nerve representation of vowels in background noise. J. Neurophysiol. 50, 27–45.
Shamma, S.A., 1985. Speech processing in the auditory system. I: The representation of speech sounds in the responses of the auditory nerve. J. Acoust. Soc. Am. 78, 1612–1621.
Siegel, S., Castellan, N.J., 1988. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York.
Silkes, S.M., Geisler, C.D., 1991. Responses of 'lower-spontaneous-rate' auditory-nerve fibers to speech syllables presented in noise. I: General characteristics. J. Acoust. Soc. Am. 90, 3122–3139.
Sinex, D.G., Geisler, C.D., 1983. Responses of auditory-nerve fibers to consonant-vowel syllables. J. Acoust. Soc. Am. 73, 602–615.
Smith, D.I., Kraus, N., 1987. Effects of chloral hydrate, pentobarbital, ketamine, and curare on the auditory middle latency response. Am. J. Otolaryngol. 8, 241–248.
Spenner, H., Urbas, J.V., 1986. Masking and the perception of stop consonants: psychoacoustical and electrophysiological experiments. Exp. Brain Res. 62, 607–617.
Stark, R.E., Heinz, J.M., 1996. Vowel perception in children with and without language impairment. J. Speech Hear. Res. 39, 860–869.
Steinschneider, M., Arezzo, J., Vaughan, H.G., Jr., 1982. Speech evoked activity in the auditory radiations and cortex of the awake monkey. Brain Res. 252, 353–365.
Tallal, P., Piercy, M., 1974. Developmental aphasia: rate of auditory

processing and selective impairment of consonant perception. Neuropsychologia 12, 83–93.

Tallal, P., Miller, S.L., Bedi, G., Byma, G., Wang, X., Nagarajan, S.S., Schreiner, C., Jenkins, W.M., Merzenich, M.M., 1996. Language comprehension in language-learning impaired children improved with acoustically modified speech. Science 271, 81–84.

Walton, J.P., Frisina, R.D., Ison, J.R., O'Neill, W.E., 1997. Neural correlates of behavioral gap detection in the inferior colliculus of the young CBA mouse. J. Comp. Physiol. A 2, 161–176.

Wiesenfeld, K., Moss, F., 1995. Stochastic resonance and the benefits of noise: from ice ages to crayfish an SQUIDs. Nature 373, 33–36.

Wright, B.A., Lombardino, L.J., King, W.M., Puranik, C.S., Leonard, C.M., Merzenich, M.M., 1997. Deficits in auditory temporal and spectral resolution in language-impaired children. Nature 387, 176–178.

Young, E.D., Sachs, M.B., 1979. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. J. Acoust. Soc. Am. 66, 1381–1403.