J Am Acad Audiol 10: 304–318 (1999)

Acoustic-Phonetic Approach toward Understanding Neural Processes and Speech Perception

Dawn Burton Koch*
Therese J. McGee*
Ann R. Bradlow*
Nina Kraus*†

Abstract

This review paper describes an "acoustic-phonetic" experimental approach aimed at understanding normal and abnormal speech perception processes from both a behavioral and an electrophysiologic perspective. First, we consider the relevant acoustic characteristics of speech and identify a set of acoustic-phonetic classes that represent the parameters most important for making an acoustic signal sound like speech. Second, we review what is known about the neurophysiologic representation of acoustic-phonetic speech parameters in animal and human subjects. Third, we describe how an acoustic-phonetic approach has been useful in understanding the biologic basis of some auditory learning problems in children and in characterizing the behavioral and neurophysiologic changes resulting from speech-sound training. Finally, we discuss these findings and how they may expand the diagnostic and rehabilitative repertoire of practicing audiologists.

Key Words: Acoustic phonetics, evoked potentials, neurophysiology, speech perception

Abbreviations: F_0 = fundamental frequency; F_1, F_2, etc. = first formant, second formant, etc; fMRI = functional magnetic resonance imaging; JND = just noticeable difference; MMN = mismatch negativity

Traditionally, audiology training and practice have been centered on understanding the normal and abnormal function of the outer, middle, and inner ear and on diagnosing and treating hearing impairments resulting from damaged peripheral hearing mechanisms. Most audiologists have a cursory knowledge of the anatomy of the central auditory system and a vague understanding of the physiology and function of the central auditory pathways.

That situation is understandable. The physiology of the middle ear, the cochlea, and the auditory nerve has been the primary focus of auditory research, and a great deal is understood about these structures and how they transduce acoustic signals into neural activity (Pickles, 1988). Much less is known about the central auditory system. Its complexity—its multiple pathways, nuclei, primary and secondary cortical areas, and elaborate network of intertract and interhemispheric connections—defies easy study. Most of the information we know about central auditory physiology has been obtained from animal subjects, and, for the most part, only simple acoustic stimuli have been used in those experiments.

Our knowledge of how the human brain hears is limited and based primarily upon psychophysical and behavioral tests of auditory perception. Moreover, clinical tests of central auditory function have been directed toward finding a particular "site of lesion," rather than upon understanding the normal or abnormal physiologic processes that contribute to the per-

---

*Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois; †Departments of Neurobiology, Physiology, Otolaryngology, and Northwestern University, Evanston, Illinois

Reprint requests: Dawn Burton Koch, Department of Communication Sciences and Disorders, Northwestern University, 2299 N. Campus Dr., Evanston, IL 60208

ception or misperception of auditory signals, especially speech (Musiek and Lamb, 1994). To quote Phillips (1995),

> with the present state of knowledge, awareness of lesion locus is useful in that it may provide new evidence on the basic science question of structure-function relations in the nervous system; it may, however, be of little direct clinical utility until such time as means are available to intervene at the neurologic level. What may be most important clinically is the specification of the perceptual or other skills that are compromised in the impaired listener, since it is that knowledge that, in the short term, will guide management strategies, if not therapeutic ones.

Consequently, it may be fruitful at this juncture in auditory neuroscience and audiology to probe the relationship between a listener's perception of acoustic signals and the neurophysiologic representation of those same signals in the same person. In particular, it would be of great utility to link the perception of speech to underlying central physiologic processes, especially since audiologists often encounter individuals whose peripheral hearing function appears normal but who complain about having difficulty understanding speech (e.g., individuals with central auditory processing disorders).

In our laboratory, we have taken an "acoustic-phonetic" approach aimed at understanding normal and abnormal speech perception processes from both a behavioral and electrophysiologic perspective. As a starting point, we postulate that the inherently complex speech signal can be decomposed into a finite and well-defined set of acoustic features that are most important for making an acoustic signal sound like speech. Specifically, rather than using highly simplified signals, we use stimuli that closely model the acoustic features that phoneticians have identified as critical for speech perception. It is in this context that we term our approach "acoustic phonetic." Then, we assess the ability of normal and impaired listeners to hear those elemental speech sounds. Concurrently, we probe the brain's representation of those speech sounds by measuring cortical evoked potentials in the same listeners and by directly measuring speech-elicited neural responses in the thalamus and cortex of animals. By using this approach, we have begun to uncover some relationships between the perception of speech and the physiologic processes that may contribute to that perception.

In this review, we highlight some of the concepts, experiments, and results associated with an "acoustic-phonetic" approach. The paper is organized into four sections. First, we consider briefly the relevant acoustic characteristics of speech and identify a set of acoustic-phonetic parameters that are applicable in this experimental approach. Second, we review what is known about the neurophysiologic representation of speech in animal and human subjects. Third, we describe how an acoustic-phonetic approach has been useful in understanding the biologic basis of some auditory learning problems in children and in characterizing the behavioral and neurophysiologic changes resulting from speech-sound training. Finally, we discuss these findings and how they may expand the diagnostic and rehabilitative repertoire of practicing audiologists.

ACOUSTIC-PHONETIC LOOK AT THE SPEECH SIGNAL

The speech signal is an extremely challenging signal to work with because of its acoustic complexity, especially compared to simple pure tones, clicks, and noise bursts. Speech is also a particularly interesting signal because it is the primary medium of linguistic communication and therefore has an inherent ecologic validity that highly simplified signals do not have. In this section, we provide a general, though brief, framework and set of organizing principles that can be used toward understanding the speech signal.

Our basic approach assumes that the speech signal can be described, to a large extent, by a finite set of acoustic-phonetic parameters that play a critical role in imparting linguistically relevant structure to the acoustic medium. These basic acoustic-phonetic properties characterize all speech sounds and form the physical basis for the sound structure of human language. By defining and classifying these critical speech dimensions, it becomes possible to manipulate them in experimental and clinical tests and thereby to reveal some of the ways in which the ear and brain represent these acoustic-phonetic properties. In the discussion that follows, we have drawn heavily from a variety of sources, most notably from Pickett (1980), Stevens (1980), Ladefoged (1982), Rosen (1992), and Johnson (1997).

"Global" Acoustic Characteristics

Certain "global" acoustic attributes characterize all samples of running speech regardless of language, talker, or communicative situation. They distinguish the spectrotemporal structure of speech from all other acoustic signals. In the temporal domain, a basic property of the speech signal is a series of peaks and valleys in the amplitude envelope. These amplitude variations result from overall energy differences between sounds produced with an open vocal tract, such as vowels, and sounds produced with more constricted vocal tracts, such as consonants. These amplitude fluctuations are heard as variations in loudness and are closely tied to the syllable-based rhythmic structure of speech.

In the spectral domain, speech is characterized by rapidly changing amplitude peaks and valleys across the frequency spectrum. Those fast spectral variations arise from the filter characteristics of the vocal tract, which enhances the energy in certain frequency regions and attenuates the energy in others. The frequency locations of these amplitude peaks (the formants) and valleys give the various speech sounds their characteristic "timbre" or "color." Because the vocal tract filter changes constantly in running speech as the tongue and lips move to articulate the various speech sounds, the frequencies of the peaks and valleys in the spectrum change constantly. Some of these spectrum changes are very rapid and can occur within 20 to 25 msec, as in the movement from a stop consonant to a following vowel. Other spectrum changes are more gradual and occur over longer time periods (i.e., 40–50 msec), as in the case of formant movements from a glide, such as /w/ or /j/, into a following vowel.

In addition to these general temporal and spectral attributes, all of the individual speech sounds that characterize spoken English can be grouped further into classes with distinctive spectrotemporal characteristics. These classes can be described in terms of their specific amplitude envelope and spectral characteristics. In many respects, this bidimensional scheme for describing speech-sound acoustics maps directly onto the traditional articulatory dimensions of manner and place of articulation.

Vowels, Glides, Liquids, and Nasals

Vowels, glides, liquids (/r/ and /l/), and nasals are a class of sounds called sonorants. In their production, air is allowed to pass freely through the oral or nasal passage. Consequently, sonorants can be distinguished from nonsonorants by peaks in the amplitude envelope. The frequency spectra of sonorants have rich harmonic structure with clear formants. Sonorants can be distinguished from each other based on their relative intensity. Vowels and glides are the most intense, and liquids and nasals are somewhat less intense. In addition, individual sounds within each sonorant class can be identified based on their formant frequencies.

The values of vowel formant frequencies vary with tongue position. In general, the first formant (F_1) is inversely related to tongue height. For example, F_1 frequency is relatively low for /i/ and /u/, the two vowels produced with a high tongue position. In contrast, F_1 frequency is relatively high for /a/, the vowel produced with a low tongue position. In general, the second and third formants (F_2 and F_3) are related to tongue position in the front-back dimension. For example, F_2 frequency is relatively high for /i/, which is produced with a fronted tongue position. In contrast, F_2 frequency is relatively low for /u/, which is produced with a back tongue position. Thus, for /i/, a vowel produced with a high and fronted tongue position, F_1 is low and F_2 is high in frequency, yielding a wide separation between F_2 and F_1. In contrast, for /a/, a vowel produced with a low and back tongue position, F_1 is high and F_2 is low in frequency. In fact, for this vowel, F_2 and F_1 are very close together, forming one broadband high-energy region.

The glides of English—/w/ (as in *why*) and /j/ (as in *yes*)—have vowel-like formants but with a slightly lower intensity than the vowels. Specifically, /w/ has /u/-like formants (low F_1 and low F_2) and /j/ has /i/-like formants (low F_1 and high F_2). The liquids, /r/ and /l/, also have clear formant structure and often show an abrupt drop in intensity. Nasal sounds also have clear formant structure, although they are usually less intense than vowels. Nasals also typically have a strong low-frequency murmur.

Fricatives, Stops, and Affricates

Fricatives, stops, and affricates are termed obstruents because they are produced with an "obstruction" in the vocal tract. Fricatives are produced with a very narrow constriction somewhere in the vocal tract, resulting in turbulent airflow, which then is filtered through the part of the vocal tract that is anterior to the point of constriction as the air passes from the lungs through the mouth. Fricative amplitude envelope characteristics are characterized by aperiodic energy. The frequency spectra of fricatives generally are characterized by broadband noise with a spectral peak that varies in frequency with the location of the constriction in the vocal tract. Fricatives that are produced at the lips and teeth, such as English /f, v, θ, ð/, have essentially no vocal tract filtering and therefore have diffuse spectra with energy spread over a wide frequency range. In contrast, fricatives that are produced further back in the vocal tract, such as English /s, z, ʃ, ʒ/, have compact spectra with energy peaks in a higher frequency region.

Stop consonants are temporally complex because they have two separate portions that are clearly identifiable in the acoustic signal. First, corresponding to the stop closure, is a silent gap. Then, corresponding to the stop release, is a brief burst of aperiodic energy. For English syllable-initial voiceless stops, (/p, t, k/), this release burst is followed by a period of heavy /h/-like aspiration, which is followed by the onset of periodicity for the following vowel. For voiced stops (/b, d, g/) and syllable-internal voiceless stops, the release burst is followed immediately by the beginning of periodicity and formant structure for the following vowel. This difference in stop release characteristics, the voice onset time, is the primary cue to stop voicing.

Spectrally, individual stops are identified primarily by spectral characteristics in the brief period between release of the closure and the onset of energy for the vowel. The critical spectral characteristics occur over a very brief period of 20 to 40 msec as the tongue moves from a position that is appropriate for the stop place of articulation into a position that is appropriate for the following vowel. Thus, the spectral characteristics of this brief transition period are constantly changing. For bilabial stops (/p, b/), both the second and third formants in this transition period are relatively low in frequency. For alveolar stops (/t, d/), the second and third formants are comparatively high in frequency. For velar stops (/k, g/), the second formant is relatively high and the third formant is close in frequency to the second formant. The specific frequency values vary substantially depending on the vowel that follows. Nonetheless, the relative pattern of formant frequencies remains constant. Affricates, such as /tʃ/ (as in *chop*) and /dʒ/ (as in *judge*), have a stop-like closure portion and a fricative-like release portion.

As an illustration of some of these acoustic-phonetic features, Figure 1 shows a time waveform and time-aligned spectrogram of the word "speech" spoken by a male speaker of American English. The salient vowel and consonant boundaries have been marked by vertical lines and labeled with International Phonetic Alphabet symbols. In the time waveform (upper panel), the initial /s/ is seen clearly by the aperiodic noisy portion that precedes the silent period for the /p/ closure. The very-low amplitude /p/ closure is followed by a brief burst of energy corresponding to the stop release, which then is followed by the relatively high-amplitude, periodic vowel portion. Finally, the silent closure period and the fricated release for the word-final affricate /tʃ/ is evident. Note that the information carried by the silent periods for the stop and affricate closures is as integral to the acoustic-phonetic "message" as the information carried by the energy-filled periods. In the spectrogram (lower panel), the initial /s/ fricative noise with a high energy peak (represented by a dark bar in the figure) is centered around 6 kHz. In the vowel portion following the /p/ closure and release, the clear formant structure is represented by the dark bands of high energy. As is typical for the high front vowel /i/, F_1 is quite low in frequency and F_2 is quite high—above 2 kHz. Finally, following the /t/ closure portion of the final affricate, the aperiodic energy for the release of the final affricate /tʃ/ is seen. Here, the energy peak is considerably lower (around 4 kHz) than that of the initial /s/.

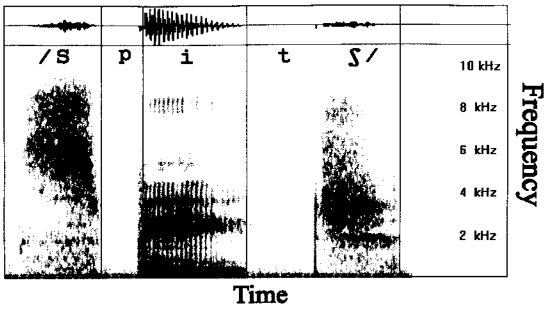


Figure 1 Time waveform (upper panel) and time-aligned spectrogram (lower panel) of the word "speech" spoken by a male talker of American English. The vowel and consonant boundaries are marked by vertical bars and labeled with International Phonetic Alphabet symbols. In the frequency-by-time plot (spectrogram), changes in intensity across the frequency range are represented by gray-scale variations. Darker bands represent high-energy regions and lighter bands represent low-energy regions.

In summary, the spectrotemporal characteristics of the acoustic speech signal are distinctive and carry information that is used by the listener to distinguish the various speech sounds.

Importantly, one can identify a set of acoustic-phonetic features that characterize the speech-sound classes that play key roles in conveying meaningful linguistic information. Those acoustic-phonetic features, in turn, can be controlled and manipulated experimentally to study the neural representation or the perception of important speech-sound classes.

ACOUSTIC-PHONETIC LOOK AT THE NEURAL REPRESENTATION OF SPEECH

The complexity of the central nervous system inherently means that acoustically complex stimuli are represented along multiple dimensions and that patterns of responses across populations of neurons likely are key to the representation of speech signals. Moreover, that representation changes from the periphery to the cortex. Interestingly, acoustic-phonetic patterns that are important in speech may be represented distinctly in the neurophysiologic responses of the auditory system. The data described below exemplify how a number of speech-relevant acoustic dimensions are represented from the auditory nerve to the cortex.

Auditory Nerve

Acoustic elements of speech signals clearly are well represented in the response of the normal auditory (VIIIth) nerve. One way this representation has been observed is by looking at the neural firing rate, or average number of spikes per second, that occurs in response to a speech stimulus. Across the array of auditory neurons, the spectral characteristics of the formant structure are represented by increased firing rates of low and medium spontaneous rate fibers of the appropriate characteristic frequency (Voigt et al, 1982; Miller and Sachs, 1983; Sinex and Geisler, 1983). Nonlinear response characteristics enhance this effect, making the peaks of the formants more prominent within the spectral array (Sinex and Geisler, 1984). This pattern of responses appears to be important at low-to-moderate intensity levels in a quiet background. At higher intensities and in noise, neural firing rates saturate across the neural population and spectral peaks are not apparent in the overall firing rate.

It is the timing pattern of the neural impulses and the synchronization of firing across the neural population that appear to be important at higher intensities and in a noisy background. Pauses in the acoustic waveform—like those present in stop consonants—are well represented as decreases in the firing rate of appropriate duration (Sinex and McDonald, 1988; Sinex, 1993; Sinex and Narayan, 1994). For voiced sounds like vowels, the formant structure is represented in the timing of neural spikes. Specifically, spikes are synchronized to the periods of the frequencies of the largest harmonic components of each formant peak across the population of neurons.

Unvoiced aperiodic elements of the speech signal, such as fricatives or (stop) bursts, do not contain harmonics and consequently do not elicit a synchronous response for the duration of the stimulus across the population of auditory nerve fibers. However, fricatives and bursts do elicit particularly strong synchronized onset responses (Delgutte and Kiang, 1984b). Interestingly, stop bursts and fricatives tend to be of short duration, and very rarely does the duration of a fricative hold important phonemic information. Thus, the well-synchronized neural representation of stimulus onset dovetails with perceptual utility. Notably, these timing and synchronization patterns are especially evident at higher stimulus levels and are resistant to the presence of background noise (Kiang and Moxon, 1974; Young and Sachs, 1979; Sachs et al, 1983; Delgutte and Kiang, 1984c; Deng and Geisler, 1987; Geisler and Gamble, 1989). Thus, at stimulus levels at which the overall neural firing rate is saturated, the timing and synchrony of neural impulses serves to encode important elements of the speech signal.

From these data, it is clear that the synchronized pattern of responses across a neuronal population is relevant to speech representation even at the periphery (Young and Sachs, 1979; Delgutte and Kiang, 1984a; Carney and Geisler, 1986; Palmer, 1990). Therefore, a number of clinically relevant inferences can be drawn. Cochlear damage results, of course, in a loss of spectral representation. In addition, damage to the nonlinear mechanisms in the cochlea results in a loss of formant enhancement. Damage that affects timing and synchronization in the auditory nerve also has a deleterious effect on speech perception, especially in noisy environments (e.g., inner hair cell damage [Harrison, 1998], acoustic neuroma, auditory neuropathy [Kraus et al, 1993]). It has been shown that, while restoring audibility is important, improving the signal-to-noise ratio also dramatically improves speech perception for persons with peripheral sensorineural hearing impairment. This result would be expected based upon how the auditory nerve represents speech signals (Turner and Holte, 1987; Turner and Robb, 1987; Geisler, 1989; Zeng and Turner, 1990; Viemeister and Plack, 1993). Thus, sensorineural impairment affects not only auditory sensitivity but also alters mechanisms crucial to sorting the speech signal from background noise.

Brain Stem/Thalamus

At the brainstem and thalamic levels, much less data are available on the encoding of speech information. Experiments from sites in the brain stem have addressed the coding of amplitude modulation, which may contribute to the perception of envelope, pitch, and fundamental frequency. These experiments have shown good synchrony to amplitude modulations across the neural population (Rhode, 1995). Furthermore, the encoding of amplitude modulation appears to be excellent in background noise (Frisina et al, 1994; Rhode and Greenberg, 1994).

The thalamus is the most peripheral location at which behavioral hemispheric asymmetries have been reported in response to speech stimuli. Studies of Parkinson's patients have noted that left thalamic lesions can result in deficits of speech and language (Samra et al, 1969).

More recently, King et al (1997, 1999) used synthetic speech syllables to elicit aggregate responses from the medial geniculate in the guinea pig thalamus. These animals showed greater asymmetries in response amplitude to the speech stimuli than to pure tones. Ten of 12 animals had consistently larger amplitude responses from the left than the right thalamus in response to speech stimuli. In the other two animals, the right thalamus was "dominant." An asymmetric aggregate response to speech indicates either a greater degree of synchrony or a larger population firing in synchrony on the dominant side. Notably, this asymmetry may be a lower level manifestation of the long-documented cortical asymmetry associated with speech and language processing.

Auditory Cortex

Evidence on cortical processing of complex stimuli is available from a number of areas, including clinical observations, behavioral results, and single- and multiple-neuron studies. From these reports, it is clear that the auditory cortex is critical to the encoding and perception of complex acoustic signals like speech. More specifically, it appears that the primary cortex has at least three functions. It serves as a gateway to association and cognitive areas, responds to complicated properties of the stimulus, and provides synchronized responses across the neural population to information-bearing elements of a complex signal.

Clinical case reports indicate that different locations in the cortex serve different purposes because cortical lesions can impair one function, yet leave another intact. For example, musical perception may be affected while speech perception remains normal. Perception of melody may be disrupted, while rhythm perception remains intact (Peretz et al, 1994). Moreover, perception of words can be impaired to the point of "word deafness," while pitch and melodic perceptions are unaffected (Takahashi et al, 1992; Klein et al, 1995). Phillips and Farmer (1990) reviewed a series of reports in which cortical damage affected the processing of rapid acoustic transitions that characterize consonants but did not affect the perception of slowly changing signals. Similarly, Micelli (1982) reported a case of cortical damage in which only the discrimination of stop consonants was affected.

Functional differentiation in the cortex also has been demonstrated in evoked potential studies. Sharma et al have described a patient who underwent a left-hemisphere frontal-temporal-parietal craniotomy for removal of an arterial venous malformation (Sharma et al, 1994; Kraus et al, 1995a). Behaviorally, his pitch discrimination was normal, but the discrimination of a consonant contrast (/da/-/ga/) was severely impaired. Mismatch negativity responses (MMNs)—neurophysiologic indicators that the brain has detected a stimulus difference—were consistent with the patient's behavioral performance. MMNs were evident when syllables differed in pitch (/da/ high versus /da/ low) but not when there was a phonemic difference (/da/ versus /ga/) (Fig. 2).

Differentiation in the representation of complex stimuli also has been demonstrated in the guinea pig thalamus and cortex. Kraus et al (1994a, b) recorded MMNs from the cortical surface and from the thalamus in response to a tonal contrast and two consonant contrasts (/ba/-/wa/, /da/-/ga/). Although all three contrasts elicited surface MMNs, only the tonal contrast elicited a robust thalamic MMN. A small thalamic MMN was evident to /ba/-/wa/ and no thalamic MMN response was observed to /da/-/ga/. In other words, the tonal difference was well rep-

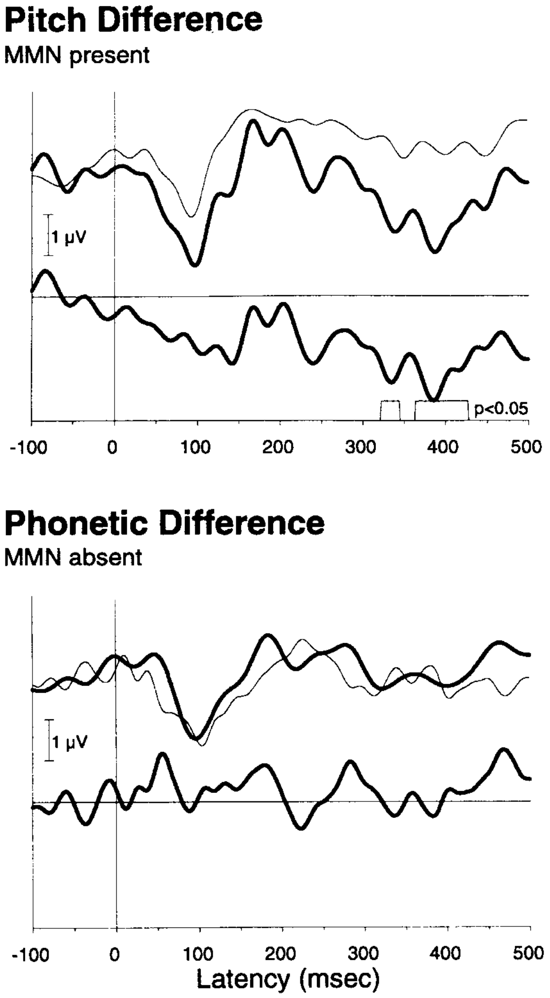


Figure 2 Neurophysiologic representation of two acoustic contrasts in a patient with a left frontal-temporal-parietal lesion. In each plot, the top thin line is the response to a standard stimulus, either a high-pitched /da/ (top panel) or a /ga/ (bottom panel). The top thick line in both plots is the response to a low-pitched /da/. The mismatch response (MMN) is seen in the difference wave (bottom thick line) as a deflection below the zero line. The boxes below indicate the latency ranges over which a significant MMN occurred ($p < .05$). Mismatch activity was evident in response to the pitch contrast (high-pitched /da/ vs low-pitched /da/) but was absent in response to the phonetic contrast (/da/ vs /ga/) (modified from Kraus et al, 1995a).

resented at the thalamic level, but the consonant differences, particularly /da/-/ga/, were well represented only at the cortical level.

In addition to clinical cases and evoked potential results, single-unit studies also indicate that the cortex is specialized for processing complex signals. For example, Rauschecker et al (Rauschecker et al, 1995; Rauschecker, 1997) have shown that cortical cells prefer complex stimuli to simple stimuli. In monkey nonprimary auditory cortex, the responses of single units were greatly facilitated when harmonics were added, one by one, to a tonal stimulus.

The response of ferret cortical neurons to sophisticated properties of complex signals has been described in a series of reports by Shamma et al. In their studies, they presented synthesized stimuli that varied systematically in formant spacing, formant peak, and formant transition. They showed that individual cells are sensitive to particular formant spacings and particular rates of formant transition (Kowalski et al, 1995, 1996a, b; Shamma and Versnel, 1995; Depireux et al, 1998). Furthermore, a cell's response to formant spacing and formant peak frequency was a good predictor of the cell's response to an actual speech stimulus (which is composed of multiple formant spacings, as well as multiple rates of transition in the various formants). In fact, predictions based on these more complex stimulus properties were more accurate than predictions based on the cell's response to simple stimuli, such as tones. It is our hypothesis that, although all of the acoustic elements of speech are represented even in the auditory nerve, it is essential to the language processor that cells at some level respond specifically to complex properties of speech stimuli, further refining the input to brain regions where speech is given meaning.

Temporal response patterns in the primary auditory cortex also are fundamental to the encoding of speech stimuli. Phillips (1993) and Phillips and Hall (1990) argue that the temporal precision of neural firing is important at the cortical level. Single units precisely encode the timing of the onset of a single event even though they may not be able to entrain to periodic events in stimuli of longer duration. For example, in monkey auditory cortex, Wang et al (1995) have described a timing pattern that emphasizes the onsets of discrete stimulus events but does not represent rapid changes such as fast periodicities.

These data contradict arguments that the auditory cortex has poor temporal processing ability. Those arguments are based on the response patterns of subcortical and cortical neurons to clicks. Cortical neurons show much poorer entrainment to fast repetition rates than do neurons from the thalamus and brain stem. Cortical neurons respond to click rates up to only 10 or 20 per second, indicating that pitch and fast periodicities are not represented in the responses of those neurons (de Ribaupierre et al, 1972; Creutzfeldt et al, 1980; Rouiller et al, 1981; Schreiner and Langner, 1988; Eggermont, 1991).

However, as Phillips argues clearly, the temporal precision of firing, rather than the ability to entrain to periodicity, may be the critical contribution of cortical neurons to the processing of complex stimuli. Recent data further support that contention. Heil and Irvine (1997) compared the temporal precision of the first spike of the responses of auditory-nerve and cortical neurons in cats. They demonstrated that cortical neurons not only show a high degree of precision but actually can have better precision than auditory-nerve fibers. An earlier study in the electric fish showed that temporal processing can become more precise at higher levels within a neural system (Carr et al, 1986). Heil and Irvine's data demonstrate a similar pattern in the mammalian auditory system.

The temporal precision and complex response patterns of a single neuron are only part of the cortical story. Eggermont (1990) and Eggermont and Smith (1996) have noted that synchronization and correlation across the neural array likely are more important than the response pattern of an individual neuron. Neurons not only must respond precisely, they must also respond together in synchrony. It may be that only part of the response of one neuron is significant, that is, the part that is synchronous with fellow neurons so that a cohesive pattern of neural activity is formed across the array. Eggermont proposes that "the brain can be viewed as a democracy of nerve cells, and as in any democracy, it will be governed by cooperation and competition." Thus, this democracy assigns relative importance to an acoustic element within a stimulus by the degree to which the neural population synchronizes to it.

Eggermont (1997) has shown that synchronization of responses occurs across relatively wide areas of the cortex. Similarly, Rauschecker (1998) has proposed that the presence of multiple tonotopic maps implies multiple representations of the stimulus, which is consistent with the notion that a complex stimulus will elicit a synchronized representation in multiple areas. Furthermore, the occurrence of synchronization across multiple cortical areas combined with relative changes in firing rate appears to signal the difference between steady-state and dynamic stimuli (Eggermont, 1997).

Interestingly, clinical measures like evoked potentials tap into these critical synchronous aspects of neural encoding. Several animal studies have shown that synchronization to voice onset time is so salient that evoked responses recorded from the cortical surface show a double onset response. One peak corresponds to the initial burst of the syllable, while the second peak corresponds to the onset of voicing. Even the categorical perception of the transition from voiced to unvoiced (as in /ba/ vs /pa/) is reflected in the evoked potential (Kurtzberg et al, 1988; Steinschneider et al, Eggermont, 1995; McGee et al, 1996; 1994; Koch et al, 1997).

For other acoustic elements of speech, evoked potential timing patterns indicate that synchronized activity emphasizes perceptually important acoustic features. For example, a comparison of cortical responses to /ba/ and /wa/ shows differences that coincide in time with the difference in vowel onset. Evoked potential responses to /da/ and /ga/ show timing differences that occur at stimulus onset, where the stimuli are different spectrally (McGee et al, 1996). These evoked potential data indicate that speech-relevant acoustic parameters are represented distinctly by the cortex, and that those response differences are evident even on the scalp surface.

Plasticity is another fundamental property of the central auditory system, with important implications for speech perception. Auditory system plasticity can result from reorganization following sensory deprivation, and as a consequence of perceptual experience and learning (Irvine and Rajan, 1996; Rajan and Irvine, 1998). For example, reorganization of the cortex is evident in single neuron responses following cochlear lesions in animals (Robertson and Irvine, 1989; Merzenich et al, 1991; Harrison et al, 1993). Similarly, functional reorganization of cortical activity, as measured by evoked potentials or fMRI, occurs after sensory deprivation in humans (Neville et al, 1983, 1998).

In normal animals, single-unit response characteristics and firing patterns also show plasticity as a consequence of experience and perceptual learning (Buchwald et al, 1966; Halas et al, 1970; Olds et al, 1972; Kraus and Disterhoft, 1982; Weinberger et al, 1984; Recanzone et al, 1993). In humans, the biologic representation of speech sounds changes with listening training (Kraus et al, 1995b). Notably, Tremblay et al (1997, 1998) also have demonstrated that neurophysiologic change occurred to other speech sounds that were similar acoustically to the trained stimuli but to which the subjects received no training, and that preconscious neurophysiologic reorganization is evident before learning is manifested behaviorally.

Taken together, clinical, evoked potential, and single-unit studies point to the fact that the cortex plays a critical role in the processing of complex signals. It appears that the primary auditory cortex has at least three functions. First, it serves as a gateway to association and cognitive areas of the brain. Second, it processes more sophisticated properties of the stimulus, and, third, it provides synchronized responses across the neural population to important, information-bearing elements of a complex signal. Moreover, the plasticity of the auditory cortex allows it to adapt and change to best represent important information.

In summary, speech-relevant acoustic-phonetic patterns are represented clearly and differentially in different locations along the auditory pathway. If we understand the specific, information-bearing characteristics of those locations, can we begin to explore the nature of hearing problems that do not originate in the cochlea or VIIIth nerve?

ACOUSTIC-PHONETIC LOOK AT SPEECH PERCEPTION

Because we can define characteristic acoustic correlates of major speech-sound classes, we can manipulate these acoustic parameters for the purposes of clinical and experimental tests of speech-sound perception in a variety of clinical populations. In other words, we can focus specifically on the perception of acoustic parameters that are known to bear critical speech information, and thereby begin to pinpoint the exact nature and origin of auditory deficits that may affect communication.

Our experimental approach has been to attempt to link the neurophysiologic representation of speech sounds in the brain with behavioral speech perception. Auditory evoked potentials are one of the few means available for looking at speech-elicited neural activity in people. By measuring behavioral perception of well-controlled speech stimuli—along with non-attention-related evoked potentials—we can begin to establish relationships between perceptual abilities and underlying central physiologic processes.

Specifically, we have been evaluating a listener's ability to discriminate two spectrotemporal contrasts: /ba/-/wa/ and /da/-/ga/. Two continua have been synthesized so that just noticeable differences (JNDs) can be determined. One continuum—/ba/ to /wa/—represents primarily a timing difference, while the other continuum—/da/ to /ga/—represents primarily a transient spectral difference. Specifically, in the /ba/-/wa/ continuum, the duration of the formant transition portion varies from 10 (/ba/) to 40 msec (/wa/). In the /da/-/ga/ continuum, the onset frequency of the third formant varies from 2580 Hz (/da/) to 2180 Hz (/ga/), while the formant transition duration is fixed at 40 msec. The selection of those contrasts was based upon behavioral research in humans, which demonstrates that perception of these speech sounds may be disrupted differentially in adults with central auditory lesions or children with auditory processing problems (Elliott and Hammer, 1988; Elliott et al, 1989; Phillips and Farmer, 1990; Sharma et al, 1994; Kraus et al, 1995a; Tallal et al, 1996).

Some interesting results have emerged. First, the ability to discriminate stimuli along the /ba/-/wa/ and /da/-/ga/ continua varies depending upon the type of listener (Fig. 3). In a study comparing normal children and children with documented learning problems, a subset of the children with learning deficits could discriminate the /ba/-/wa/ contrast with the same precision as normal children. However, these same children had greater JNDs than normal children when tested with the /da/-/ga/ continuum. Moreover, their impaired discrimination ability along the /da/-/ga/ continuum was reflected in the absence of an MMN elicited by a /da/-/ga/ contrast (Fig. 4). The MMN elicited by a /ba/-/wa/ contrast was normal, thereby reflecting their normal behavioral discrimination of that contrast (Kraus et al, 1996). Importantly, these data indicate that the basis for some speech discrimination deficits may lie in deficits of neurophysiologic encoding along the auditory pathway.

Preliminary behavioral results comparing normal-hearing young adults and older adults (> 55 years of age) show a similar pattern. The ability to discriminate /ba/-/wa/ is the same in the two groups, while the older adults have more difficulty discriminating the /da/-/ga/ contrast than the young subjects, even though their hearing sensitivity is normal (see Fig. 3) (Koch et al, 1999).

If, in fact, the ability to discriminate relevant acoustic speech contrasts is impaired, or the neurophysiologic representation is faulty, might the ability be restored? Two strategies have been explored for this purpose, both aimed at capitalizing on the plasticity that exists in the central auditory system (Salvi et al, 1996; Palmer et al, 1998). First, some research suggests that manipulating the input signal by

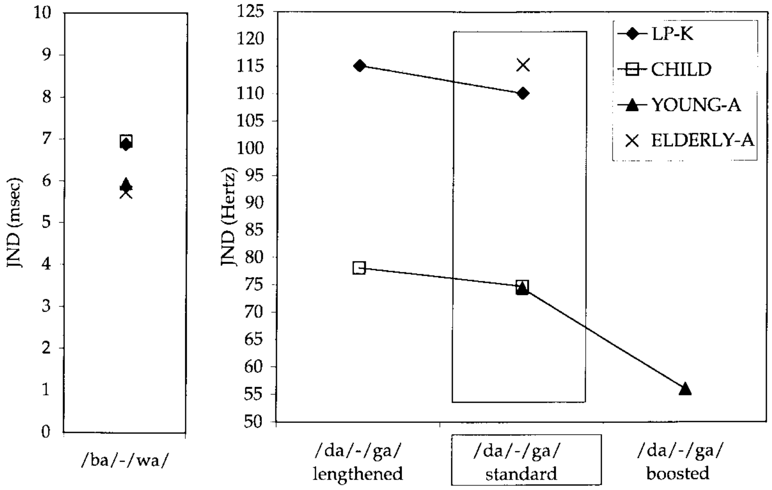


Figure 3 Just noticeable differences for four synthetic speech continua for normal children (child, n = 65), children with auditory-based learning problems (LP-k, n = 37), normal young adults (young-a, n = 17), and a group of normal-hearing women over 55 years of age (elderly-a, n = 8). The /ba/-/wa/ continuum varied in the duration of the formant transition. The /da/-/ga/-standard continuum varied in the onset frequency of the third formant with a 40-msec formant transition duration. The /da/-/ga/-boosted continuum was the same as the /da/-/ga/40 continuum but the amplitude of the formant transition was increased. The /da/-/ga/-lengthened continuum was the same as the /da/-/ga/40 continuum but the duration of the formant transition was lengthened to 80 msec.

enhancing selected acoustic features can be beneficial (Merzenich et al, 1996; Tallal et al, 1996). In our laboratory, we have attempted to enhance discrimination along the /da/-/ga/ continuum by lengthening the formant transition duration or by boosting the amplitude of the formant transition relative to the following vowel (see Fig. 3). For children with learning problems, lengthening the transition did not improve their impaired behavioral discrimination of the /da/-/ga/ contrast. Notably, however, that stimulus manipulation enhanced their neurophysiologic responses compared to the responses elicited by the shorter transition duration stimuli (Bradlow et al, 1998, 1999), which may be a precursor to behavioral improvement (see below). Other recent data from our laboratory indicate that boosting the transition amplitude does enhance behavioral discrimination along the /da/-/ga/ continuum in normal adults. These results, although preliminary, suggest that a cue-enhancement strategy may prove beneficial to clinical populations.

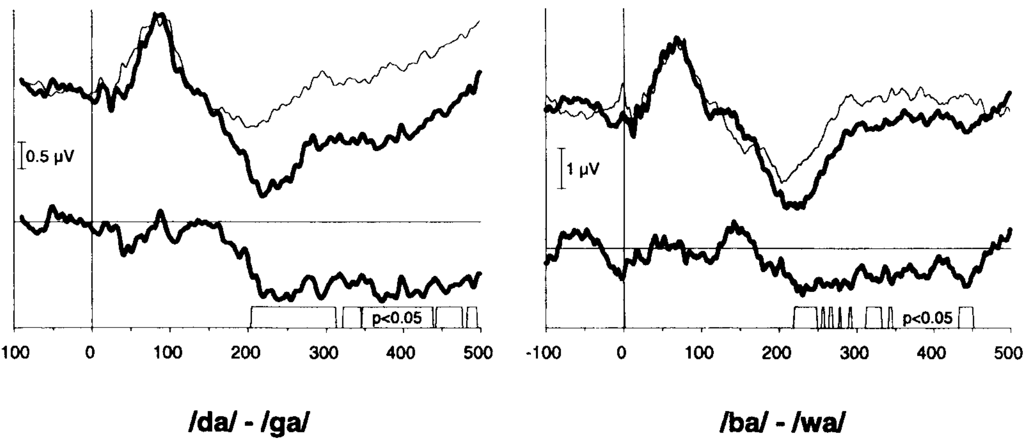
Second, other studies indicate that training a listener to hear difficult contrasts may lead to speech discrimination improvement that generalizes to other speech communication situations (Pisoni et al, 1982; McClaskey et al, 1983; Bradlow et al, 1997). Furthermore, several recent experiments in our laboratory demonstrate that neurophysiologic changes occur with speech-sound training over a short time period (a few weeks), and that those changes can be apparent before behavioral improvement is evident (Kraus et al, 1995b; Tremblay et al, 1997, 1998). In other words, those data suggest that auditory training alters the neural activity that underlies the representation of speech before those changes are incorporated into conscious behavior.

In summary, an acoustic-phonetic experimental approach allows us to begin to link single-neuron physiology in animals, speech-elicited neural responses in people, and speech perception. Importantly, speech-specific evoked potentials and behavioral perception can be measured concurrently in human listeners. Furthermore, these data indicate that the representation of speech sounds is plastic and that rehabilitative strategies involving enhancing speech contrasts or focused listening training may prove efficacious in some populations that exhibit speech perception deficits of central origin.

ACOUSTIC-PHONETIC APPROACHES IN CLINICAL AUDIOLOGY: FUTURE DIRECTIONS

For audiologists, understanding the nature and origins of hearing loss is critical to designing appropriate habilitation or rehabilitation programs. The acoustic-phonetic perspective described here provides a framework for understanding some of the physiologic and psy-

A. Good /da/-/ga/ perceivers.



B. Poor /da/-/ga/ perceivers.

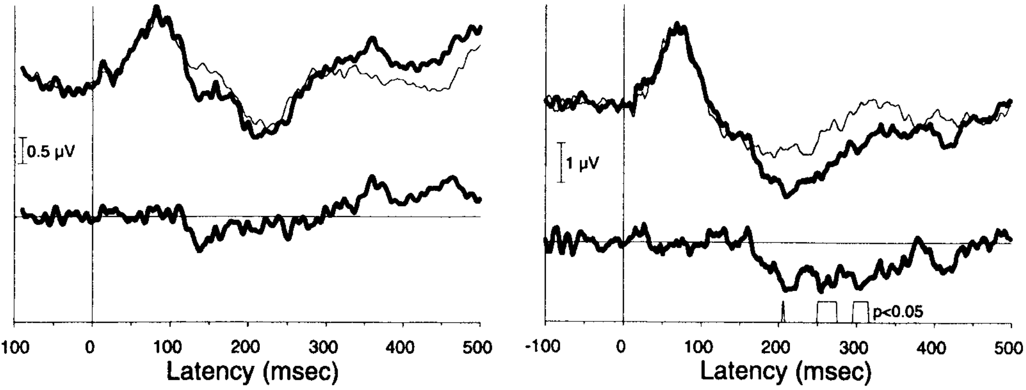


Figure 4 Grand average MMN responses elicited by a /ba/-/wa/ and a /da/-/ga/ contrast in *A*, children who were "good" /da/-/ga/ perceivers and *B*, children who were "poor" /da/-/ga/ perceivers. Responses from an F_z electrode location are shown. In the first column, the top thin line is the response to a /da/ stimulus when it was presented alone. The top thick line is the response to a /da/ stimulus when it signaled an acoustic change in a sequence of /ga/ stimuli. In the second column, the top thin line is the response to a /wa/ stimulus when it was presented alone. The top thick line is the response to a /wa/ stimulus when it signaled an acoustic change in a sequence of /ba/ stimuli. In both columns, the MMN is seen in the difference wave (lower thick line) as a deflection below the zero line. The boxes below indicate the latency ranges over which a significant MMN occurred ($p < .05$). Both groups showed an MMN response to the /ba/-/wa/ contrast. However, only the "good" /da/-/ga/ perceivers evidenced an MMN response to the /da/-/ga/ contrast (adapted from Kraus et al, 1996).

choacoustic processes contributing to speech perception that are separate from peripheral hearing impairment. As we continue these studies, we can begin to answer questions like:

1. What classes of speech sounds may be less discriminable in individuals with auditory perceptual deficits? Are those classes different for different clinical populations? Are they different for different ages?
2. Is the perception of some speech-sound classes particularly vulnerable in adverse listening conditions? Under what circumstances does the perception decrease?
3. Can an acoustic-phonetic test paradigm become clinically applicable? Would the behavioral and neurophysiologic information obtained from individuals lead to an improved diagnosis of hearing impairment? Put differently, could we distinguish between peripheral hearing problems, preconscious central sensory deficits, and cognitive or language deficits?
4. Would behavioral and neurophysiologic test results facilitate treatment, either through focused auditory training or augmentive devices such as hearing aids? Could we monitor the biologic impact of rehabilitation

before behavioral changes are evident, that is, can we determine whether an individual auditory system is amenable to training?

The answers to these questions will open up a new avenue of practice for audiologists and other professionals who are concerned with hearing impairment. We are beginning to understand how the brain hears and how the central auditory system contributes to both normal and impaired communication. As we gain further knowledge about these processes, we expect to develop and implement rehabilitation strategies focused specifically on individual auditory perceptual difficulties.

Acknowledgment. The authors greatly appreciate the scientific contributions of their colleagues Thomas Carrell, Jenna Cunningham, Cynthia King, Trent Nicol, Anu Sharma, Kelly Tremblay, and Steven Zecker. This work was supported by the National Institute on Deafness and Other Communicative Disorders (DC01906, DC01510, DC03762), Foundation for Hearing and Speech Research, Hugh S. Knowles Center, National Organization for Hearing Research.

REFERENCES

Bradlow AR, Pisoni DB, Akahane-Yamada R, Tohkura Y. (1997). Training Japanese listeners to identify English /r/ and /l/: some effects of perceptual learning on speech production. *J Acoust Soc Am* 101:2299–2310.

Bradlow AR, Kraus N, Nicol T, McGee T, Cunningham J, Carrell TD. (1998). Speech-sound perception in normal and learning-disabled children: effect of lengthened CV transition duration. *J Acoust Soc Am* 103:2985.

Bradlow AR, Kraus N, Nicol T, McGee T, Cunningham J, Carrell TD. (1999). Effect of lengthened transition duration on discrimination and neural representation of synthetic CV syllables by normal and learning-disabled children. Submitted.

Buchwald J, Halas E, Schramm S. (1966). Changes in cortical and subcortical unit activity during behavioral conditioning. *Physiol Behav* 1:11–22.

Carney LH, Geisler CD. (1986). A temporal analysis of auditory-nerve fiber responses to spoken stop consonant-vowel syllables. *J Acoust Soc Am* 79:1896–1914.

Carr CE, Heiligenberg W, Rose GJ. (1986). A time-comparison circuit in the electric fish midbrain. I. Behavior and physiology. *J Neurosci* 6:107–119.

Creutzfeldt O, Hellweg F-C, Schreiner C. (1980). Thalamocortical transformation of responses to complex auditory stimuli. *Exp Brain Res* 39:87–104.

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

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

Delgutte B, Kiang NYS. (1984c). Speech coding in the auditory nerve: V. Vowels in background noise. *J Acoust Soc Am* 75:908–918.

Deng L, Geisler CD. (1987). Responses of auditory-nerve fibers to nasal consonant-vowel syllables. *J Acoust Soc Am* 82:1977–1988.

Depireux DA, Simon JZ, Klein DJ, Shamma SA. (1998). Representation of dynamic complex spectra in primary auditory cortex. *Assoc Res Otolaryngol Abstr* 21:148.

de Ribaupierre F, Goldstein M, Yeni-Komshian G. (1972). Cortical coding of repetitive acoustic pulses. *Brain Res* 46: 205–225.

Eggermont J. (1990). *The Correlative Brain: Theory and Experiment in Neural Interaction. Studies of Brain Function*. Vol. 16. Berlin: Springer-Verlag.

Eggermont J. (1991). Rate and synchronization measures of periodicity coding in cat primary auditory cortex. *Hear Res* 56:153–167.

Eggermont J. (1995). Representation of a voice onset time continuum in primary auditory cortex of the cat. *J Acoust Soc Am* 98:911–920.

Eggermont J. (1997). Firing rate and firing synchrony distinguish dynamic from steady state sound. *Neuroreport* 8:2709–2713.

Eggermont J, Smith GM. (1996). Neural connectivity only accounts for a small part of neural correlation in auditory cortex. *Exp Brain Res* 110:379–391.

Elliott L, Hammer M. (1988). Longitudinal changes in auditory discrimination in normal children and children with language-learning problems. *J Speech Hear Disord* 53:467–474.

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

Frisina RD, Walton JP, Karcich KJ. (1994). Dorsal cochlear nucleus single neurons can enhance temporal processing capabilities in background noise. *Exp Brain Res* 102:160–164.

Geisler CD. (1989). The responses of models of "high-spontaneous" auditory-nerve fibers in a damaged cochlea to speech syllables in noise. *J Acoust Soc Am* 86:2192–2205.

Geisler CD, Gamble T. (1989). Responses of "high-spontaneous" auditory-nerve fibers to consonant-vowel syllables in noise. *J Acoust Soc Am* 85:1639–1652.

Halas E, Beardsley J, Sandie M. (1970). Conditioned neuronal responses at various levels in conditioning paradigms. *Electroencephalogr Clin Neurophysiol* 28:468–477.

Harrison RV. (1998). An animal model of auditory neuropathy. *Ear Hear* 19:355–361.

Harrison R, Stanton S, Ibrahim D, Nagasawa A, Mount R. (1993). Neonatal cochlear hearing loss results in developmental abnormalities of the central auditory pathways. *Acta Otolaryngol* 113:296–302.

Heil P, Irvine DR. (1997). First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *J Neurophysiol* 78:2438–2454.

Irvine DR, Rajan R. (1996) Injury- and use-related plasticity in the primary sensory cortex of adult mammals: possible relationship to perceptual learning. *Clin Exp Pharmacol Physiol* 23:939–947.

Johnson K. (1997). *Acoustic and Auditory Phonetics*. Cambridge, MA: Blackwell.

Kiang NY, Moxon EC. (1974). Tails of tuning curves of auditory-nerve fibers. *J Acoust Soc Am* 55:620–630.

King C, Nicol T, McGee T, Kraus N. (1997). Aggregate responses in the left and right thalamus and cortex to auditory stimulation. *Assoc Res Otolaryngol Abstr* 20:104.

King C, Nicol T, McGee T, Kraus N. (in press). Thalamic asymmetry is related to acoustic signal complexity. *Neurosci Lett.*

Klein SK, Kurtzberg D, Brattson A, Kreuzer JA, Stapells DR, Dunn MA, Rapin I, Vaughan HG Jr. (1995). Electrophysiologic manifestations of impaired temporal lobe auditory processing in verbal auditory agnosia. *Brain Lang* 51:383–405.

Koch D, Tremblay K, Dunn I, Dinces E, Carrell T, Kraus N. (1997). Speech-evoked N1 and mismatch neurophysiologic responses in cochlear-implant users and normal listeners. *Assoc Res Otolaryngol Abstr* 20:80.

Koch DB, Bradlow AR, Bellis TJ, Cunningham J, Kraus N. (1999). Effects of age on speech-sound representation and perception in normal-hearing women. *Assoc Res Otolaryngol Abstr* 22:286.

Kowalski N, Depireux DA, Shamma SA. (1996a). Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. *J Neurophysiol* 76:3503–3523.

Kowalski N, Depireux DA, Shamma SA. (1996b). Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. *J Neurophysiol* 76:3524–3534.

Kowalski N, Versnel H, Shamma SA. (1995). Comparison of responses in the anterior and primary fields of the ferret cortex. *J Neurophysiol* 73:1513–1523.

Kraus N, Disterhoft J. (1982). Response plasticity of single neurons in rabbit auditory association cortex during tone-signalled learning. *Brain Res* 246:205–215.

Kraus N, McGee T, Carrell T, Sharma A. (1995a). Neurophysiologic bases of speech discrimination. *Ear Hear* 16:19–37.

Kraus N, McGee T, Carrell T, King C, Tremblay K, Nicol T. (1995b) Central auditory system plasticity associated with speech discrimination training. *J Cogn Neurosci* 7:25–32.

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

Kraus N, McGee T, Ferre J, Hoeppner JA, Carrell TD, Sharma A, Nicol T. (1993). Mismatch negativity in the neurophysiologic/behavioral evaluation of auditory processing deficits: a case study. *Ear Hear* 14:223–234.

Kraus N, McGee T, Littman T, Nicol T, King C. (1994a). Nonprimary auditory thalamic representation of acoustic change. *J Neurophysiol* 72:1270–1277.

Kraus N, McGee T, King C, Littman T, Nicol T. (1994b). Discrimination of speech-like signals in auditory thalamus and cortex. *J Acoust Soc Am* 96:2758–2768.

Kurtzberg D, Stapells DR, Wallace IF. (1988). Event-related potential assessment of auditory system integrity: implications for language development. In: Vietze P, Vaughan Jr HG, eds. *Early Identification of Infants with Developmental Disabilities*. Philadelphia: Grune and Stratton, 160–180.

Ladefoged P. (1993). *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.

McClaskey C, Pisoni D, Carrell T. (1983). Transfer of training of a new linguistic contrast in voicing. *Percept Psychophys* 34:323–330.

McGee T, Kraus N, King C, Nicol T. (1996). Acoustic elements of speechlike stimuli are reflected in surface recorded responses over the guinea pig temporal lobe. *J Acoust Soc Am* 99:3606–3614.

Merzenich M, Grajski K, Jenkins W, Recanzone G, Peterson B. (1991). Functional cortical plasticity: cortical network origins of representational changes. *Cold Spring Harbor Symp Quant Biol* 55:873–887.

Merzenich MM, Jenkins WM, Johnston P, Schreiner C, Miller SL, Tallal P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* 271:77–81.

Micelli G. (1982). The processing of speech sounds in a patient with cortical auditory disorder. *Neuropsychologia* 20:5–20.

Miller MI, Sachs MB. (1983). Representation of stop consonants in the discharge patterns of auditory-nerve fibers. *J Acoust Soc Am* 74:502–517.

Musiek FE, Lamb L. (1994). Central auditory assessment: an overview. In: Katz J, ed. *Handbook of Clinical Audiology*. 4th Ed. Baltimore: Williams & Wilkins, 197–211.

Neville H, Schmidt A, Kutas M. (1983). Altered visual-evoked potentials in congenitally deaf adults. *Brain Res* 266:127–132.

Neville HJ, Bavelier D, Corina D, Rauschecker J, Karni A, Lalwani A, Braun A, Clark V, Jezzard P, Turner R. (1998). Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience. *Proc Natl Acad Sci U S A* 95:922–929.

Olds J, Disterhoft J, Segal M, Kornblith C, Hirsh R. (1972). Learning centers of rat brain mapped by measuring latencies of conditioned unit responses. *J Neurophysiol* 35:202–219.

Palmer AR. (1990). The representation of the spectra and fundamental frequencies of steady-state single- and double-vowel sounds in the temporal discharge patterns of guinea pig cochlear-nerve fibers. *J Acoust Soc Am* 88:1412–1426.

Palmer CV, Nelson CT, Lindley GA. (1998). The functionally and physiologically plastic adult auditory system. *J Acoust Soc Am* 103:1705–1721.

Peretz I, Kolinsky R, Tramo M, Labrecque R, Hublet C, Demeurisse G, Belleville S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain* 117:1283–1301.

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

Phillips DP. (1995). Central auditory processing: a view from auditory neuroscience. *Am J Otol* 16:338–352.

Phillips DP, Farmer ME. (1990). Acquired word deafness and the temporal grain of sound representation in the primary auditory cortex. *Behav Brain Res* 40:85–94.

Phillips DP, Hall SE. (1990). Response timing constraints on the cortical representation of sound time structure. *J Acoust Soc Am* 88:1403–1411.

Pickett JM. (1980). *The Sounds of Speech Communication*. Baltimore: University Park Press.

Pickles J. (1988). *An Introduction to the Physiology of Hearing*. New York: Academic Press.

Pisoni DB, Aslin RN, Perey AJ, Hennessy BL. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *J Exp Psychol Hum Percept Perform* 8:297–314.

Rajan R, Irvine DR. (1998). Neuronal responses across cortical field A1 in plasticity induced by peripheral auditory organ damage. *Audiol Neuro-Otol* 3:123–144.

Rauschecker JP. (1997). Processing of complex sounds in the auditory cortex of cat, monkey, and man. *Acta Otolaryngol Suppl* 532:34–38.

Rauschecker JP. (1998). Parallel processing in the auditory cortex of primates. *Audiol Neuro-Otol* 3:86–103.

Rauschecker JP, Tian B, Hauser M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268:111–114.

Recanzone G, Schreiner C, Merzenich M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J Neurosci* 13:87–104.

Rhode WS. (1995). Interspike intervals as a correlate of periodicity pitch in cat cochlear nucleus. *J Acoust Soc Am* 97:2414–2429.

Rhode WS, Greenberg S. (1994). Encoding of amplitude modulation in the cochlear nucleus of the cat. *J Neurophysiol* 71:1797–1825.

Robertson D, Irvine DR. (1989). Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J Comp Neurol* 282:456–471.

Rosen S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Phil Trans R Soc Lond* 336:367–373.

Rouiller E, Ribaupierre Y de, Toros-Morel A, Ribaupierre F de. (1981). Neural coding of repetitive clicks in the medial geniculate body of cat. *Hear Res* 5:81–100.

Sachs MB, Voigt HF, Young ED. (1983). Auditory nerve representation of vowels in background noise. *J Neurophysiol* 50:27–45.

Salvi RJ, Henderson D, Fiorino F, Colletti V, eds. (1996). *Auditory System Plasticity and Regeneration*. New York: Thieme Medical.

Samra K, Riklan M, Levita E, Zimmerman J, Waltz JM, Bergmann L, Cooper IS. (1969). Language and speech correlates of anatomically verified lesions in thalamic surgery for parkinsonism. *J Speech Hear Res* 12:510–540.

Schreiner CE, Langner G. (1988). Coding of temporal patterns in the central auditory nervous system. In: Edelman GM, Gail WE, Cowan WM, eds. *Auditory Function*. New York: John Wiley & Sons, 337–361.

Shamma SA, Versnel H. (1995). Ripple analysis in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary spectral profiles. *Auditory Neurosci* 1:255–270.

Sharma A, Kraus N, Carrell T, Thompson C. (1994). Physiologic bases of pitch and place of articulation perception: a case study. *J Acoust Soc Am* 95:3011.

Sinex DG. (1993). Auditory nerve fiber representation of cues to voicing in syllable-final stop consonants. *J Acoust Soc Am* 94:1351–1362.

Sinex DG, Geisler CD. (1983). Responses of auditory-nerve fibers to consonant-vowel syllables. *J Acoust Soc Am* 73:602–615.

Sinex DG, Geisler CD. (1984). Comparison of the responses of auditory nerve fibers to consonant-vowel syllables with predictions from linear models. *J Acoust Soc Am* 76:116–121.

Sinex DG, McDonald LP. (1988). Average discharge rate representation of voice onset time in the chinchilla auditory nerve. *J Acoust Soc Am* 83:1817–1827.

Sinex DG, Narayan SS. (1994). Auditory-nerve fiber representation of temporal cues to voicing in world-medial stop consonants. *J Acoust Soc Am* 95:897–903.

Steinschneider M, Arezzo JC, Schroeder C, Vaughan HG Jr. (1994). Speech evoked activity in primary auditory cortex: effects of voice onset time. *Electroencephalogr Clin Neurophysiol* 92:30–43.

Stevens K. (1980). Acoustic correlates of some phonetic categories. *J Acoust Soc Am* 68:836–842.

Takahashi N, Kawamura M, Shinotou H, Hirayama K, Kaga K, Shindo M. (1992). Pure word deafness due to left hemisphere damage. *Cortex* 28:295–303.

Tallal P, Miller SL, Bedi G, Byma G, Wang X, Nagarajan SS, Schreiner C, Jenkins WM, Merzenich MM. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 271:81–84.

Tremblay K, Kraus N, Carrell T, McGee T. (1997). Central auditory system plasticity: generalization to novel stimuli following listing training. *J Acoust Soc Am* 102:3762–3773.

Tremblay K, Kraus N, McGee TJ. (1998). The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *Neuroreport* 9:3557–3560.

Turner CW, Holte L. (1987). Discrimination of spectral-peak amplitude by normal and hearing-impaired subjects. *J Acoust Soc Am* 81:445–451.

Turner CW, Robb MP. (1987). Audibility and recognition of stop consonants in normal and hearing-impaired subjects. *J Acoust Soc Am* 81:1566–1573.

Viemeister NF, Plack CJ. (1993). Time analysis. In: Yost WA, Popper AN, Fay RR, eds. *Human Psychophysics.* New York: Springer-Verlag, 116–154.

Voigt HF, Sachs MB, Young ED. (1982). Representation of whispered vowels in discharge patterns of auditory-nerve fibers. *Hear Res* 8:49–58.

Wang X, Merzenich MM, Beitel R, Schreiner CE. (1995). Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 74:2685–2706.

Weinberger MN, Hopkins W, Diamond DM. (1984). Physiological plasticity of single neurons in auditory cortex of the cat during acquisition of the pupillary conditioned response: I. Primary field (AI). *Behav Neurosci* 98:171–188.

Young ED, Sachs MB. (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.

Zeng FG, Turner CW. (1990). Recognition of voiceless fricatives by normal and hearing-impaired subjects. *J Speech Hear Res* 33:440–449.