From Ear to Cortex: A Perspective on What Clinicians Need to Understand About Speech Perception and Language Processing

Susan Nittrouer

Utah State University, Logan

ne day not long after my 3-year-old daughter started wearing hearing aids, I went to pick her up at her day care center and was called into the office. The director informed me that the staff was concerned because the hearing aids did not seem to be working. For a moment, I was confused. I checked the

ABSTRACT: Phoneme-sized phonetic segments are often defined as the most basic unit of language organization. Two common inferences made from this description are that there are clear correlates to phonetic segments in the acoustic speech stream, and that humans have access to these segments from birth. In fact, well-replicated studies have shown that the acoustic signal of speech, lacks invariant physical correlates to phonetic segments, and that the ability to recognize segmental structure is not present from the start of language learning. Instead, the young child must learn how to process the complex, generally continuous acoustic speech signal so that phonetic structure can be derived.

This paper describes and reviews experiments that have revealed developmental changes in speech perception that accompany improvements in access to phonetic structure. In addition, this paper explains how these perceptual changes appear to be related to other aspects of language development, such as syntactic abilities and reading: Finally, evidence is provided that these critical developmental changes result from adequate language experience in naturalistic contexts, and accordingly suggests that intervention strategies for children with language learning problems should focus on enhancing language experience in natural contexts.

KEY WORDS: speech perception, phonological processing, phonological awareness, language processing hearing aids every morning, I thought to myself, and knew they were always in perfect working order. So I asked for clarification. The director replied that my daughter did not seem to be hearing, even with the hearing aids. Suddenly, recognition of the true problem struck me. The day care staff thought that my daughter would put on the hearing aids and understand everything they said. I explained that it takes a while for children with hearing loss to develop the language they missed before the loss was identified. I informed the director that the real source of my daughter's apparent lack of hearing was her continued language delay. The signal was reaching her, but she could not yet make sense of it.

A year later (during which time she was in intensive intervention), my daughter contracted bilateral ear infections and could not wear her hearing aids for a week. One day during that period, I went to pick her up and was approached by an excited day care provider exclaiming that she was happy to see that my daughter's hearing loss was improving. Now, so it seemed to this provider, my daughter could hear, even without her aids. Of course, her auditory thresholds had not changed; her language abilities had improved. My daughter was now able to make sense of the message being conveyed, even with a restricted acoustic signal.

What this story illustrates is how strongly language underlies other communicative and perceptual processes, affecting even the very reception of the acoustic speech signal. Native language background and general language proficiency affect how, and how well, we apprehend phonological structure from the acoustic speech signal. From our own experiences, many of us know the frustration of traveling to a foreign country and trying to master at least a few key phrases in another language, such as

Hello, Thank you, and Where's the bathroom? After what we consider a splendid attempt, a native speaker may correct our pronunciation. Often, the problem is that we feel certain we said exactly what that person is saying. We cannot discriminate our attempt from his correction. Or, we may feel that everyone is speaking too quickly, and so we cannot extricate the string of phonetic segments. Why is this? Certainly we share the same general auditory capacities as speakers of other languages. Why does the amount and kind of language experience we receive in childhood affect our abilities to recognize phonetic segments? These are the topics to be explored in this paper.

THE TASK FACING THE CHILD: LEARNING TO EXTRICATE AND USE PHONOLOGICAL STRUCTURE

The term "phonological processing," as used here, refers to the combined processes of accessing phonological structure and then using that structure for further language processing. Converging evidence from research in language development, reading, and memory all highlight the pivotal role that the abilities to access and use phonological structure have in language processing. Probably most apparent is the need to be able to recognize phonological units (specifically, phoneme-sized phonetic segments) in order to learn to read: alphabetic labels are attached to these units in many orthographies. The notion that the source of the problems faced by poor readers is an impaired ability to recognize and manipulate phonological structure is well supported and readily accepted today. One of the best single volumes documenting this phenomenon remains a collection of papers offered in tribute to the work of Isabelle Liberman, who is largely credited with discovering this connection (Brady & Shankweiler, 1991; see also Wagner & Torgesen, 1987, for a thorough review). However, phonological processing is critical to other aspects of language processing as well.

Competent language users store words in their mental lexicons using phonological codes. Evidence from a variety of sources indicates that phonologically similar words are grouped together in the lexicon in what are sometimes called "neighborhoods," and that the lexicon is searched using phonological probes during lexical retrieval (e.g., Charles-Luce & Luce, 1990; Kohn et al., 1987; Peter & Turvey, 1994; Van Orden, 1987; Van Orden, Johnston, & Hale, 1988). Although investigators differ concerning what they believe to be the exact nature of the lexical retrieval process, it is clear that the phonological organization of the lexicon helps to make lexical access fast and efficient (e.g., Damian & Martin, 1999; Marslen-Wilson, 1990). Individuals who fail to use a phonological code for lexical storage have difficulty accessing words in their own lexicons (e.g., Denkla & Rudel, 1976; Katz, 1986).

Competent language users also make use of phonological codes for storing incoming signals in verbal working memory (Baddeley, 1966; Baddeley, Thomson, & Buchanan, 1975; Campbell & Dodd, 1980; Spoehr & Corin, 1978). To

do this, two kinds of phonological skills are required. First, the listener must be able to recognize phonological structure in the language that he or she hears; second, the listener must be able to put words into memory using that code, and do so quickly enough to keep up with the incoming message. Individuals with poor phonological processing abilities, such as poor readers, consequently would be expected to have poor retention of word strings. Several investigators have found evidence to support this prediction (e.g., Hall, Wilson, Humphreys, Tinzmann, & Bowyer, 1983; Mann & Liberman, 1984; Nittrouer & Miller, 1999; Spring & Perry, 1983). This constraint on verbal working memory interferes with comprehension of sentences with complex syntax because such sentences are usually long (Bar-Shalom, Crain, & Shankweiler, 1993; Byrne, 1981; Smith, Mann, & Shankweiler, 1986).

Finally, the ability to apprehend phonological structure from the signal facilitates speech perception in noisy backgrounds. Children with language impairments have more difficulty extracting speech from noise than children without language impairments (e.g., Cornelissen, Hansen, Bradley, & Stein, 1996; Cunningham, Nicol, Zecker, Bradlow, & Kraus, 2001), but the common wisdom is that the problems of listening in noise are, at least partly, the source of the language difficulties. In fact, research suggests that the direction of causality is opposite to that suggestion. That is, it is precisely the phonological processing problem that makes it hard for these children to extricate the speech signal from the noisy background. Evidence of this is provided by the finding that children with poor phonological processing abilities show no deficit in recognition for nonspeech sounds presented in noise (Brady, Shankweiler, & Mann, 1983). Specifically, Brady et al. demonstrated that age-matched third graders with good or poor reading abilities (and so, with good or poor phonological processing abilities) were similarly competent at recognizing nonspeech sounds embedded in noise. However, when the task switched to word recognition in noise, the performance of the children with good reading abilities improved, whereas the performance of the children with poor reading abilities remained the same as it had been for the nonspeech signals.

WHY THE TASK OF ACCESSING PHONETIC STRUCTURE IS SO COMPLEX

This paper would not need to be written, and research in the areas of language processing, speech perception, and reading would be considerably simpler, if it were the case that phonetic units existed in the acoustic speech signal. In that case, recognizing phonetic units for the purpose of language processing would merely consist of picking them out of the signal, one at a time, in sequence. Consequently, the learning requirements would be minimal, if any. According to that scenario, it is conceivable that a disorder would exist in which some individuals were unable to perform this simple task—that they would somehow lack the ability to recognize these easily retrievable entities.

However, such a hypothetical disorder would be much more straightforward in nature (and so easier to repair) than the phonological processing and related language problems that are actually found in children (and adults) with language-related disorders.

In fact, phonetic segments do not exist as neatly packaged and sequenced units. The acoustic signal of speech shows no distinctive boundaries that might mark where one segment ends and another begins, and the acoustic properties that can be associated with any particular segment are spread over fairly broad temporal regions. This situation is greater than what is normally termed "coarticulatory effects." That term suggests that there are primitive forms of phonetic units, and these forms are modified by phonetic context. But, there are no phonetic primitives.

In spite of this lack of correspondence between physical units and linguistic segments, most investigators interested in speech perception during the second half of the 20th century spent their time looking for acoustic correlates of phonetic segments. These correlates, most often termed "cues," were brief bits of the signal, believed to indicate phonetic identity. To be sure, differences of opinion existed concerning what the nature of these cues would ultimately turn out to be. Some investigators' believed that settings of broad-spectrum properties (i.e., stop bursts that spanned the entire range of frequencies important to speech) would be found to signal each linguistic element, and these settings would be the same across all phonetic contexts (e.g., Blumstein & Stevens, 1980; Kewley-Port, 1983; Stevens & Blumstein, 1978). Others postulated that specific spectral regions (i.e., individual formants or formant transitions) would be found that signal specific phonetic elements, and that the settings of these spectral regions would be found to vary depending on the settings of other acoustic properties in the vicinity (e.g., Liberman, 1957; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Mann & Repp, 1980; Studdert-Kennedy, 1983). Unfortunately, both lines of investigation have largely failed to explain how it is that listeners' derive phonetic structure from the acoustic, speech signal.

Just before the turn of the century, many speech researchers turned away from the view of speech perception that is inherent to the experimental approach described above (i.e., of searching for acoustic cues to phonetic identity). Instead, the notion has emerged that no one definable acoustic property or set of properties is associated with each phonetic segment. Rather, the perceiver organizes information from across the temporal and spectral domains of the signal using strategies specific to one's native language. With the right organizational strategy, phonetic structure emerges. As explained already, the task facing the young child is to develop the language-specific perceptual skills needed to organize the acoustic speech signal so that phonetic structure can be derived.

In spite of the readily apparent mismatch between the physical and linguistic organization of the signal, the notion has been perpetuated that human infants are born able to analyze the speech signal as is necessary to derive phonetic structure. This position arises from studies reporting that acoustically different stimuli are discriminated by some

ŧ

infants when, and only when, those differences support different phonetic descriptions (e.g., Eilers, Gavin, & Oller, 1982; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Kuhl, 1979; Moffitt, 1971). However, there have been reports challenging this position (Eimas & Miller, 1980; Lasky, Syrdal-Lasky, & Klein, 1975; Nittrouer, 2001). Regardless of how that work on infant speech perception sorts itself out, other relevant results remain irrefutable: The abilities to access phonetic structure consciously or to use that structure in language processing are skills that emerge over the first decade or so of life. For example, data collected using a variety of methods demonstrate that children only gradually acquire access to the various levels of linguistic structure, with access to phonetic structure the last level to be attained (Fox & Routh, 1975; Liberman, Shankweiler, Fischer, & Carter, 1974; Treiman & Zukowski, 1991; Walley, Smith, & Jusczyk, 1986). Evidence also exists showing that the lexicon only gradually acquires the . phonetic organization described earlier (Charles-Luce & Luce, 1990), and that the ability to store words in working memory using a phonological code develops through childhood (Nittrouer & Miller, 1999; Treiman, 1995). In sum, a preponderance of evidence indicates that children must learn to organize the incoming speech signal in order to derive phonetic structure, and to use that structure for further linguistic processing.

GENERAL PROCEDURES FOR STUDYING SPEECH PERCEPTION

When speech perception research was based on the notion that isolated acoustic correlates of phonetic categories (i.e., cues) could be identified, the basic procedure was to vary one acoustic property along a continuum and measure changes in phonetic judgments. However, as the underlying model of speech perception research has changed, procedures have had to change. Currently, the speech perception experiments that inform us involve the manipulation of multiple acoustic properties. Usually, one property is manipulated along a continuum, going from a setting appropriate for one phonetic category to a setting appropriate for another phonetic category, such that the two categories form a minimal pair. A second acoustic property is manipulated dichotomously, set solidly for one or the other phonetic category. The resulting stimuli (i.e., every level of the continuously varied property paired with each level of the dichotomously varied property) are played for listeners multiple times. The listener's task is to make a binary choice labeling decision. Results are plotted for both settings of the dichotomously varied property as the proportion of one category label (the y axis) given at each level of the continuously varied property (the x axis). Results are interpreted according to the steepness of the functions (i.e., the steeper the functions, the more weight that was assigned to the continuously varied property represented on the x axis) and the separation in functions (i.e., the greater the separation, the more weight that was assigned to the dichotomously varied property). Consequently, steep

functions without much separation between them indicate phonetic decisions that were based largely on the continuously varied property, rather than on the dichotomously varied property. Conversely, shallow functions that are widely separated indicate that the listener weighted the dichotomously varied property greatly. (Figures 1 and 3 provide examples of these response patterns.)

EVIDENCE OF LANGUAGE-SPECIFIC PERCEPTUAL STRATEGIES

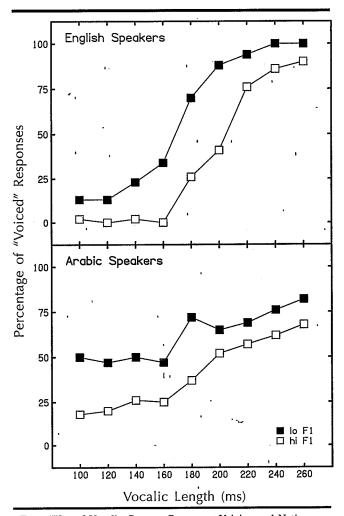
Crowther and Mann (1994) provided an example of the experimental approach described above. They investigated language-specific differences in the weighting of two acoustic properties that affect judgments of voicing for syllable-final stops in English: duration of the vocalic portion of the syllable and frequency of the first formant (F1) at voicing offset. In many languages, including English, the duration of the vocalic portion of a syllable differs depending on whether a stop at the end of that' syllable is voiced or voiceless (e.g., Chen, 1970). In all languages, F1 is higher in frequency at voicing offset before voiceless than before voiced final stops, owing to the simple articulatory fact that F1 will be higher if voicing is ended before the vocal tract achieves closure than it will be if voicing continues into closure. Accordingly, Crowther and Mann constructed synthetic versions of pot and pod that varied in vocalic duration (i.e., the continuously varied property) and in F1-offset frequency (i.e.; the dichotomously varied property).

These stimuli were played for adults whose native language was either English or Arabic. Arabic was selected because it does not exhibit a voicing-related vocalic length distinction (Flege & Port, 1981). Mean functions for each group are shown in Figure 1. The functions for native Arabic listeners are much shallower than those of the native English listeners, and so Crowther and Mann concluded that the Arabic listeners did not weight vocalic length as much as the English listeners in their decisions of syllable-final stop voicing. There is no apparent difference across groups in the separation between functions, and so the authors concluded that listeners in both groups weighted F1 frequency at voicing offset similarly. Presumably, the English listeners were no more sensitive to changes in temporal information than the Arabic listeners. Consequently, this study serves as an example of how general language background shapes the way that acoustic properties are weighted in phonetic decisions.

DEVELOPMENTAL CHANGES I: ATTENDING TO DETAIL IN THE ACOUSTIC SIGNAL

The perceptual strategies of the English and Arabic listeners in the study just described were presumably similar when these listeners were newborns, and so the language-related differences must reflect developmental

Figure 1. Labeling functions for English and Arabic speakers for pot/pod stimuli.^a

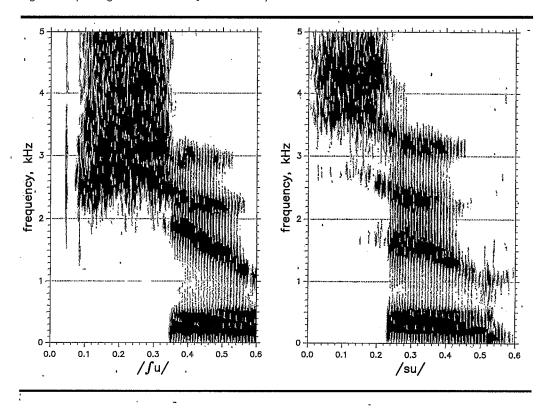


^a From "Use of Vocalic Cues to Consonant Voicing and Native Language Backgrounds: The Influence of Experimental Design," by C. S. Crowther and V. A. Mann, 1994, Perception & Psychophysics, 55, p. 518. Copyright 1994 by the Psychonomic Society. Reprinted with permission.

changes that occurred as those listeners gained experience with their native languages. What exactly changes as children spend time listening to the language of those around them, and how are those changes related to the ability to recover phonetic structure from the speech signal?

To answer those questions, several experiments have been run, many using fricative-vowel (FV) stimuli (e.g., Nittrouer, 1992; Nittrouer & Studdert-Kennedy, 1987). In these experiments, stimuli are modeled after natural tokens of /ʃ/-vowel and /s/-vowel syllables. Spectrograms for natural tokens of /ʃu/ and /su/ are shown in Figure 2. As can be seen, the aperiodic fricative noise associated with /ʃ/ generally has its lower frequency limit at approximately 2.2 kHz, whereas the noise associated with /s/ has its lower frequency limit at roughly 3.6 kHz. At the same time, when the syllable starts with /ʃ/ rather than with /s/, the second formant (F2) is higher, and the third formant (F3) is lower, at voicing onset.

Figure 2. Spectrograms of natural /ʃu/ and /su/ syllables.



For the perception experiments, nine synthetic fricative noises (150 ms long) were generated that had energy centered at frequencies ranging from 2.2 kHz (appropriate for /ʃ/) to 3.8 kHz (appropriate for /s/). Thus, the spectrum of the fricative noise served as the continuously varied property. The dichotomously manipulated property was the formant transitions, which were appropriate for either syllable-initial /ʃ/ or /s/. In most experiments, two vowels were used, one rounded (such as /u/) and one nonrounded (such as $/\alpha$). Thus, 36 stimuli are created (9 noises \times 2 formant-transition conditions × 2 vowel contexts). However, each vowel context was presented separately so that the labeling decision was a binary choice (e.g., shoe or Sue). Listeners indicated their responses by pointing to one of two pictures, which represented the response alternatives. Stimuli were presented 10 times each, in blocks of 18. In general, 12 to 14 listeners of each age group served as participants.

Results were plotted as the percentage of /s/ responses, at each level of the fricative noise, for each kind of vocalic portion separately. Figure 3 displays typical results, in this case for 3-, 5-, and 7-year-olds, as well as for adults. These results happen to be for stimuli made with synthetic vocalic portions (Nittrouer & Miller, 1997), but the same results are found when natural vocalic portions are used (see Figure 7 later). Two trends, both statistically significant, are apparent. First, the functions become steeper with increasing age, suggesting that more weight is being given to the fricative-noise spectrum in decisions of fricative place as listeners get older. Second, the separation in functions diminishes, suggesting that less weight is being

placed on formant transitions. These trends suggest that young children pay particular attention to the formant transitions of the speech signal. Such a strategy would both facilitate early attempts at linguistic parsing as well as indicate the articulatory gestures required to produce one's native language.

Formant transitions form the backbone of the syllable, so to speak, tying together syllable nucleus with syllable margins. Thus, paying attention to formant transitions could provide a way for the child without much language to begin parsing the incoming signal. Formant transitions also provide information about the place of vocal tract constriction at either end of the transition. So, by paying attention to these transitions, the child can also glean information about where constrictions should be made in order to produce the words being heard, although not information about precise constriction shapes. Presumably, then, pressure to recognize syllable-internal structure (most likely from a burgeoning lexicon) and to produce articulatory gestures more precisely (in order to be better understood) leads children to begin to discover spectral details on either side of those formant transitions.

Of course, there could be other explanations for the agerelated trends exhibited in Figure 3. For example, children might have more difficulty than adults performing the task used to collect data. Perhaps children cannot attend to the task long enough to provide reliable responses. Perhaps they have difficulty labeling somewhat ambiguous stimuli, which are often used in perceptual experiments. Perhaps they have trouble using nonsense labels, which are often used in these experiments as well. Consequently, it is

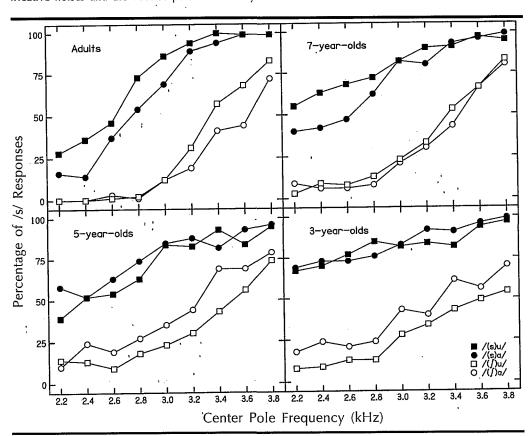


Figure 3. Labeling functions for adults and children for /ʃ/-vowel and /s/-vowel stimuli.^a Both the fricative noises and the vocalic portions were synthetic.

helpful to compare perceptual results for children and adults using the same task, but with stimuli for which no age-related changes would be expected. As just such a test, children and adults were asked to label /f/-vowel and /0/vowel syllables using the same task as that described above for /[/-vowel and /s/-vowel syllables. Harris (1958) showed that adults base their decisions of fricative place largely on the fricative-noise spectrum for /ʃ/ versus /s/, as seen in the experiment described above. However, for decisions of /f/ versus /θ/. Harris reported that adults pay particular attention to formant transitions. This difference in perceptual strategy is likely due to the simple fact that /f/ and / θ /· spectra are more similar to each other than are /s/ and /[/ spectra. If the assertion is accurate that observed differences across age groups in labeling results for /ʃ/-vowel and /s/-vowel were due to developmental changes in perceptual weighting strategies, then similar labeling results for adults and children would be expected for /f/-vowel and $/\theta$ -vowel. In other words, there should be some phonetic decisions for which listeners do not modify their strategies as they gain experience with a native language because the original strategies are just right. .

To test that prediction, a study was conducted using natural /f/ and / θ / noises (the dichotomous property) and

synthetic vocalic portions that had formant transitions varying along a nine-step continuum from those appropriate for a preceding /f/ to those appropriate for a preceding θ (Nittrouer, 2002). This arrangement of acoustic properties is opposite to that of the /ʃ/-/s/ studies, but was necessary precisely because /f/ and / θ / spectra are so similar: There simply is not enough acoustic difference between the two noises to make a continuum. With this modification in arrangement, however, interpretation of results shifts. Now, the steepness of the functions will indicate the extent to which phonetic judgments were based on formant transitions, and separation in functions will indicate the extent to which phonetic judgments were based on the fricative-noise spectra. Figure 4 displays results for children (ages 4, 6, and 8 years of age) and adults, and shows that adults and children performed more similarly with these stimuli than with the /ʃ/ -vowel and /s/-vowel stimuli. Importantly, this result bolsters the interpretation from the /ʃ/-vowel and /s/-vowel experiments that the age-dependent results were due to different weighting strategies for those stimuli and not the result of attention or task demands. Across these experiments, what is found is that children's strategies are the same regardless of the fricative decision to be made—they rely on formant transitions to a great extent. It is adults' perceptual

^a From "Predicting Developmental Shifts in Perceptual Weighting Schemes," by S. Nittrouer and M. E. Miller, 1997, Journal of the Acoustical Society of America, 101, p. 2262. Copyright 1997 by the Acoustical Society of America. Reprinted with permission.

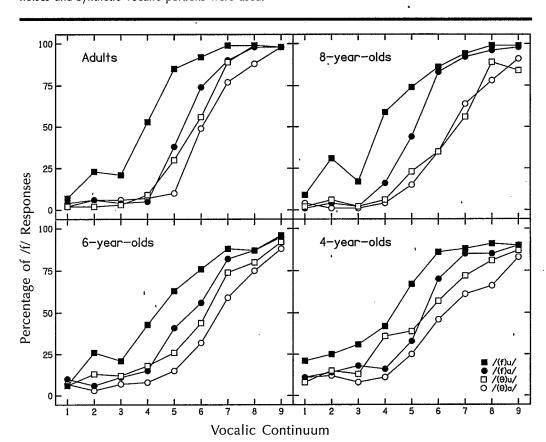


Figure 4. Labeling functions for adults and children for /f/-vowel and /θ/-vowel stimuli.^a Natural noises and synthetic vocalic portions were used.

strategies that differ depending on the decision involved. Effectively, adults seem to have learned what information is most informative for the decision to be made.

DEVELOPMENTAL CHANGES II: LEARNING TO BE FLEXIBLE

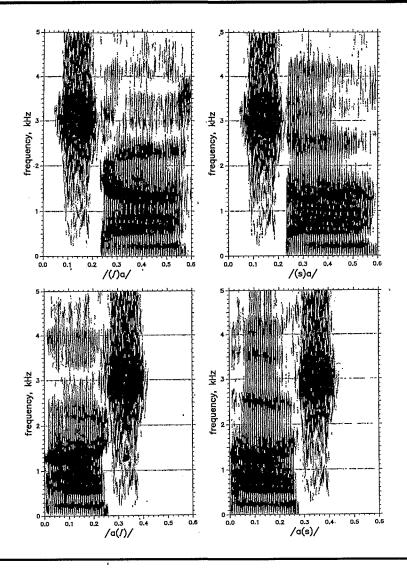
Not only do adults use different weighting strategies for different phonetic decisions, they sometimes use different strategies for the same phonetic decision when different syllable structures are involved. Figures 5 and 6, from Nittrouer, Miller, Crowther, and Manhart (2000), illustrate this point. These figures show syllables that begin or end with the synthetic noise from the middle of an /ʃ/-to-/s/continuum (i.e., centered at 3.0 kHz). The vocalic portions are natural, taken from a speaker saying /ʃ/-vowel, /s/-vowel, vowel-/ʃ/, or vowel-/s/. Figure 5 displays stimuli with the vowel /a/, with FV syllables shown in the top half and vowel-fricative (VF) syllables shown in the bottom half. Formant transitions clearly signal place of fricative constriction when the fricative is in the syllable-initial position. As was seen in Figure 2, F2 is higher and F3 is

lower at voicing onset when the syllable starts with /ʃ/ rather than /s/. However, these differences in formant frequencies at voicing edge are not seen when the fricative is at the end of the syllable. In particular, F2 is similar in frequency at voicing offset preceding both /f/ and /s/, and these syllable-final transitions appear to be a mirror image of the syllable-initial F2 transition for preceding /ʃ/. Although not quite as remarkable, F3 distinguishes between fricatives more poorly for fricatives at the ends of syllables, as well. Figure 6 shows stimuli with the vowel /u/, and similar trends are found. For syllable-initial fricatives, formant transitions clearly signal fricative place. For syllable-final fricatives, they do not. In this case, syllablefinal transitions (again, particularly F2) for syllables with both /[/ and /s/ resemble those of syllable-initial /s/. That is, F2 is relatively low at the edge of voicing. So looking across syllable structures, formant transitions are found to vary with fricative place-of-articulation for FV syllables, but to vary with vowel identity for VF syllables.

Both FV and VF syllables of the type shown in 'Figures 5 and 6 were presented to children (5 and 7 years of age) and adults for fricative labeling. Results for the FV stimuli, for adults and 5-year-olds, are shown on the left of Figure 7. As in Figure 3, children's functions appear well

^a From "Learning to Perceive Speech: How Fricative Perception Changes, and How It Stays the Same," by S. Nittrouer, 2002, Journal of the Acoustical Society of America, 112, p. 715. Copyright 2002 by the Acoustical Society of America. Reprinted with permission.

Figure 5. Sample spectrograms of fricative-/a/ and /a/-fricative stimuli.^a Synthetic noises and natural vocalic portions were used. The specific stimuli shown consist of the fricative noise from the center of the continuum (with a center frequency of 3.0 kHz), and vocalic portions taken from a male speaker saying /ʃa/, /sa/, /aʃ/, and /as/. Although the natural fricative noise was removed to make these stimuli, the fricative label in parentheses indicates what that fricative context had been.



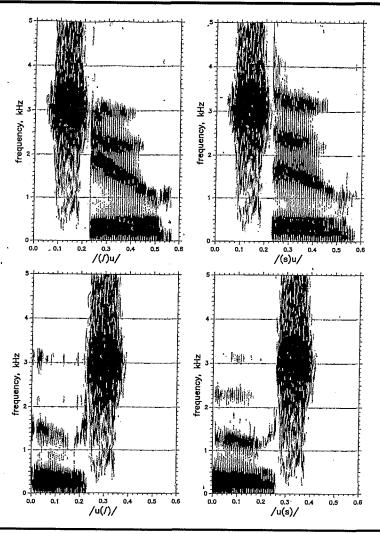
^a From "The Effect of Segmental Order on Fricative Labeling by Children and Adults," by S. Nittrouer, M. E. Miller, C. S. Crowther, and M. J. Manhart, 2000, Perception & Psychophysics, 62, p. 272. Copyright 2000 by the Psychonomic Society. Reprinted with permission.

separated depending on whether formant transitions were appropriate for a preceding /s/ (filled symbols) or /ʃ/ (open symbols), indicating a strong reliance on formant transitions for fricative decisions.¹ The right-hand side of Figure 7 shows results for the VF stimuli, and reveals different patterns. For adults, functions are much closer together, indicating that they paid even less attention to formant transitions for these syllables than for the FV syllables. Clearly this strategy is appropriate given that information regarding fricative place is not to be found in the formant transitions for VF syllables. For 5-year-olds, on the other hand, their functions are well separated, but this time the

separation depends on whether the vowel was /a/ or /u/. More /s/ responses were given to /u/ stimuli (squares) than to /a/ stimuli (circles). This pattern mirrors the acoustic

The difference in labeling functions between adults and 5-year-olds may not appear as great in Figure 7 as in Figure 3. This is because of the tough perceptual demands of the experiment for which results are displayed in Figure 7 (Nittrouer et al., 2000). In that study, children needed to provide data not only for these natural FV and VF stimuli, but also for the stimuli played with the bits reversed. As a result, only 5-year-olds with relatively mature perceptual skills made it through this experiment. Nonetheless, the differences between 5-year-olds and adults observed in earlier studies were also found in this one, just attenuated somewhat in magnitude.

Figure 6. Spectrograms of fricative-/u/ and /u/-fricative stimuli.a See Figure 5 for details.



^a From "The Effect of Segmental Order on Fricative Labeling by Children and Adults," by S. Nittrouer, M. E. Miller, C. S. Crowther, and M. J. Manhart, 2000, Perception & Psychophysics, 62, p. 273. Copyright 2000 by the Psychonomic Society. Reprinted with permission.

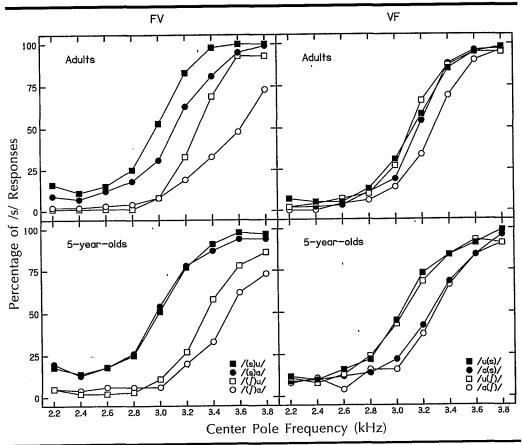
structure of the stimuli in that formant transitions for /u/-fricative syllables resembled those of /su/, whereas those for /a/-fricative syllables resembled those of /ʃa/. Thus, the 5-year-olds were apparently still weighting formant transitions in their decisions of fricative identity, even though those transitions did not provide information about fricative place. These children still had something to learn about speech perception.

DEVELOPMENTAL CHANGES III: AWARENESS OF PHONETIC STRUCTURE

Thus far, two developmental changes in speech perception have been described: a developmental shift in the weighting of some acoustic properties that contribute to phonetic decisions, and an increase in the flexibility of

these weighting strategies so that they vary across utterances with different phonetic structures. The suggestion was made that these developmental changes in speech perception are related to developmental changes in access to phonological (especially phonetic) structure. Several experiments have been conducted to test this hypothesis. Here, the term "phonological awareness abilities" refers specifically to how skillfully an individual can recognize phonological structure in the speech signal. Although the term can refer to phonological structure at any level, phonetic structure is of interest here. Under typical conditions, this process takes place without conscious awareness, and so measuring these abilities becomes tricky. Just about any method that might be devised to evaluate how well an individual can access phonological structure requires the use of some task that involves bringing these normally automatic processes to the level of awareness, and the very ability to do that may vary across individuals.

Figure 7. Labeling functions for adults and 5-year-olds for fricative-vowel and vowel-fricative stimuli. Stimuli consisted of synthetic noises and natural vocalic portions, as shown in Figures 5 and 6.



^a Adapted from "The Effect of Segmental Order on Fricative Labeling by Children and Adults," by S. Nittrouer, M. E. Miller, C. S. Crowther, and M. J. Manhart, 2000, Perception & Psychophysics, 62. Copyright 2000 by the Psychonomic Society. Adapted with permission.

Therefore, care must be taken to ensure that it is not variation in these metalinguistic abilities that is actually being evaluated. One way to do that is by taking advantage of the fact that all children (whether normally developing or delayed) are constantly acquiring new skills. By evaluating a range of related skills, it is often possible to find some for which the experimental group performs similarly to the control group, and some for which group differences are found. A finding of similarity can be used to support the position that the groups do not differ on the more general ability of being able to bring to conscious awareness structure that is not usually at that level of awareness. The finding of some group differences can then be used to support the position that the two groups genuinely do differ on those specific phonological awareness skills examined.

In one such study (Nittrouer, 1999), 110 children (ages 8 to 10 years) were tested on a variety of tasks, including phonological awareness tasks and fricative labeling for EV syllables. All children participating passed a hearing screening and had normal articulation and nonverbal cognitive abilities. As part of the protocol, these children were given a brief test of word recognition (i.e., the reading subtest of the Wide Range Achievement Test—

Revised; Jastak & Wilkinson, 1984). Children who scored better than one standard deviation below the mean were considered to have normal reading abilities for their ages; children who scored at or more than one standard deviation below the mean were considered to be poor readers. Although the independent measure of interest in this study was not reading ability per se, dividing the groups this way provided two groups of children differing in phonological awareness abilities, as will be seen.

To evaluate the abilities of these children to access specifically phonetic structure, three tasks were used that varied in when they would be expected to appear developmentally. The first was one in which children had to decide which word, out of three, began with the same initial consonant as a target word (i.e., the "initial-consonant-the-same" [ICTS] task). This task was considered the easiest of the three because the child only had to recognize the phonetic segment at the beginning of the word. The second task, a phoneme-deletion task, required the child to provide the real word that would result if a specified segment was removed from a nonsense word (e.g., If you say /plost/ without the /l/, you get /post/.). This task was considered more difficult than the first because the child not only had

to access the phonetic structure of an item, but he or she also had to determine what it would be if one segment were removed. The third task, a pig Latin task, was considered the most difficult because the child had to remove a segment from one part of the word, synthesize a new syllable with that segment, and attach it to the end of the new word (e.g., the pig Latin version of the word candy is andy-cay).

Results are shown in Table 1. Both groups performed similarly on the ICTS task, indicating that even the children with poor reading skills were able to recognize the segment at the start of words and could bring this information to conscious awareness when asked to do so. For the phonemedeletion task, children with poor reading abilities had more trouble than children with normal reading abilities. Similarly, poor readers had more difficulty than normal readers with the pig Latin task. Precisely because no group difference was found for the first task, it can be concluded that these latter group differences reveal true delays in the development of phonological awareness abilities on the part of the children with poor reading abilities.

The fricative-labeling task described earlier for FV syllables (using natural vocalic portions) was also administered to these two groups of children. Figure 8 shows labeling functions and reveals that functions were shallower and more widely separated as a function of formant transitions for the children with poor reading abilities than for those with normal reading abilities. Other developmental studies (some of which have been reviewed here) have shown that this pattern is typical of children younger than 8 years of age, who are presumably developing language normally. Thus, this study provides support for the suggestion that learning how to weight the acoustic properties of speech is related to learning to recognize phonetic structure.

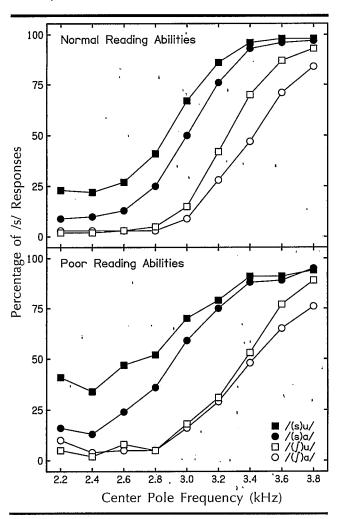
Language Experience Underlies Developmental Changes in Perceptual Strategies and Phonological Awareness

The hypothesis has been proposed that appropriate and adequate experience with one's native language promotes these related developmental changes. Another investigation (Nittrouer, 1996) specifically tested that hypothesis. In this study, several groups of 8-year-olds presumed to have deficits in the amount and kind of experiences they had

Table 1. Mean number of items correct on three phonological awareness tasks for 8- to 10-year-olds with normal or poor reading abilities. Standard deviations are provided in parentheses. Total number of items on each task was 24 for initial consonant the same (ICTS), 32 for phoneme deletion, and 48 for pig Latin (Nittrouer, 1999).

•	Normal readers	Poor readers
•	•	
ICTS	22.7 (2.0)	22.3 (2.2)
Phoneme deletion	23.6 (7.2)	13.8 (7.5)
Pig Latin	26.6 (15.9)	10.2 (12.4)

Figure 8. Labeling functions for 8- to 10-year-olds with normal and poor reading abilities for /ʃ/-vowel and /s/-vowel stimuli. Stimuli consisted of synthetic noises and natural vocalic portions.



^a From "Do Temporal Processing Deficits Cause Phonological Processing Problems?" by S. Nittrouer, 1999, Journal of Speech, Language, and Hearing Research, 42, p. 936. Copyright 1999 by the American Speech-Language-Hearing Association. Reprinted with permission.

had with their native language served as participants, as well as a control group. One group consisted of children living in conditions of low socioeconomic status (low-SES). Work by others provided a basis for believing that parental language input would be found to differ for these children as compared to their mid-SES peers. The language of low-SES parents has been reported to include more directives (which require no verbal response from the child) than inquiries (which do require the child to form a verbal response) (Hess & Shipman, 1965; Schachter, 1979).

When children in low-SES homes do talk, it has been observed that their efforts are less likely to get a response from their parents than the efforts of mid-SES children (Schachter & Strage, 1982). Of importance, these differences in parental language input depending on SES have been observed across ethnic and cultural groups and

geographic regions (Hess & Shipman, 1965; Laosa, 1980). For the work reported here, parents of the children in the study were asked to help their children make a Tinkertoy model from a picture, and each parent-child dyad was videotaped for 10 min. Using an interval scoring procedure (with 10-s observation intervals and 2-s recording intervals), graduate students who were blind to SES status scored parental language behaviors. Results replicated the findings of others: Children in the low-SES dyads received fewer parental inquiries than children in the mid-SES group (percentage of all behaviors during the 10-min session = 7% for the low-SES parents, 19% for the mid-SES parents) and more parental directives (16% for the low-SES parents, 10% for the mid-SES parents). The difference between groups for inquiries indicates that just in these 10-min recording sessions, children in the mid-SES group had almost three times as many opportunities to form and produce a verbal response as children in the low-SES group.

For these same groups of children, both phonological awareness abilities and perceptual weighting strategies for speech were examined. Looking first at phonological awareness abilities, only the phoneme-deletion and pig Latin tasks described earlier were administered. Results are shown in Table 2. On both tasks, children in the low-SES group performed more poorly than children in the mid-SES group. Looking next at perceptual weighting strategies for speech, the same task with the same stimuli as that used in Nittrouer (1999) was used with these children. Figure 9 shows labeling functions and indicates that these functions were shallower and more widely separated, depending on formant transitions, for the low-SES group than for the mid-SES group. Thus, this study provides support for the position that it is early language experience (rather than any other developmental effect) that is responsible for changes in the way children process the acoustic signal of speech, and so for changes in children's abilities to extract phonological structure.

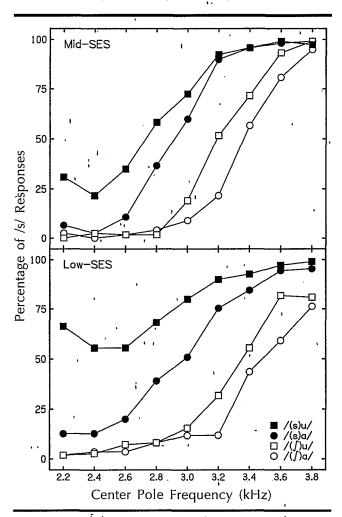
How This Approach Differs From Others

The suggestion made above is that language experience in naturalistic contexts allows children to acquire language-specific ways of organizing the acoustic speech signal, and it is through these organizational frameworks that phonetic structure emerges. Thus, this approach places the focus of phonological access and processing squarely on perceptual processes specific to one's native language. It is *not* being suggested that speech perception proceeds differently from

Table 2. Mean number of items correct on two phonological awareness tasks for 8-year-olds from mid or low socioeconomic status (SES) backgrounds. Standard deviations are provided in parentheses (Nittrouer, 1996).

	Mid-SES	Low-SES
Phoneme deletion Pig Latin	21.9 (1.8) 19.9 (2.5)	8.7 (1.8) 0.8 (0.6)

Figure 9. Labeling functions for 8-year-olds from mid-socioeconomic status (SES) and low-SES backgrounds for /ʃ/-vowel and /s/-vowel stimuli. Stimuli were the same as those for which labeling results are shown in Figure 8.



^a Adapted from "The Relation Between Speech Perception and Phonemic Awareness: Evidence From Low-SES Children and Children With Chronic OM," by S. Nittrouer, 1996, Journal of Speech and Hearing Research, 39. Copyright 1996 by the American Speech-Language-Hearing Association. Adapted with permission.

the perception of other distal stimuli. Skilled perceivers of ' any stimulus know what information to pay attention to and how to organize the various sources of information so they can extract the representation they seek. For example, the goal of wine or tea tasting is to recognize discrete flavors in the substance, and so skilled wine and tea tasters have learned what to pay attention to and how to organize that information to recognize those flavors. That information is available to all of us, but most of us have not honed our perceptual skills so precisely as to allow recognition of those flavors. And that is exactly the suggestion being made here for speech perception: The child must learn what information to extract from the acoustic speech signal and how to organize that information in order to access phonetic structure. Thus, the perception of speech is not special. It is exactly like skilled perception of any other stimulus.

Others take a different approach, placing the focus of phonological access on general auditory processing, either explicitly or implicitly. In particular, these approaches suggest that children who are having difficulty learning to access phonetic structure suffer from problems in auditory processing of a nonlinguistic nature (e.g., Chermak & Musiek, 1997; Friel-Patti, 1999; Jerger & Musiek, 2000). It follows from this suggestion that nonlinguistic processing mechanisms must be responsible for phonetic access. Theoretical perspectives of this sort mandate a view of speech perception in which phonetic segments are represented isomorphically in the acoustic speech signal such that the listener only needs to pluck them off. Auditory processing deficits, it is proposed, hinder a child's ability to glean these segments from the signal (e.g., Merzenich et al., 1996). In fact, evidence has been garnered showing that children with language impairments experience greater backward, masking than children without these impairments (Wright et al., 1997). However, Rosen and Manganari (2001) showed that the backward masking deficits of children with language impairments could not explain their speech perception difficulties. Deficits in backward masking could only. interfere with speech perception if phonetic segments were represented isomorphically in the acoustic signal, such that a later arriving segment would mask an earlier arriving segment. Rosen and Manganari's study showed that the children with language impairments (and deficits in backwards masking) were equally poor at discriminating consonants in consonant-vowel and vowel-consonant syllables. The reason for this lack of a demonstrable connection undoubtedly has to do with the lack of correspondence between acoustic structure and linguistic structure already noted.

Another, slightly different proposal is that children with, language impairments are particularly poor at processing rapidly arriving information (e.g., Tallal & Piercy, 1973; 1974). Again, this suggestion could only explain deficits in accessing phonetic segments if phonetic segments were represented isomorphically in the signal, so that rate of arrival was a concern. In any event, other investigators have failed to garner support for this suggestion (e.g., Bishop, Carlyon, Deeks, & Bishop, 1999; Mody, Studdert-Kennedy, & Brady, 1997; Nittrouer, 1999). In general, models of phonetic access (and related disabilities) based on the notion that phonetic segments (or alternatively, features that can be used to construct phonetic segments) are clearly and sequentially instantiated in the acoustic speech signal cannot be supported. Of course, there are physical constraints that can affect how much of the speech signal is available to the listener, with peripheral hearing loss being the most obvious. However, these conditions generally result in an experiential deficit as well, which can affect a child's opportunities to develop appropriate perceptual strategies for speech.

IMPLICATIONS FOR ASSESSMENT AND INTERVENTION

The tightly linked set of language skills that emerged for my daughter during that first year of intervention and hearing aid use arose largely from her general language experiences. The special program that she attended ensured that hardly a minute of her 6-hour school day went by without an opportunity to hear a language model or attempt an utterance herself. The exclamation of the day care provider that her hearing loss had improved after that first year of hearing aid use and intervention attests to the pervasive effects of such language enrichment. The studies reviewed in this paper have provided more specific descriptions of exactly why it is that early language experience affects one aspect of a child's language abilities-phonological processing. Speech perception, often considered a peripheral process, is fundamentally influenced by a child's experience with a native language. The child must learn what acoustic properties to pay particular attention to, which ones to ignore, and how to adjust those weighting strategies as a function of linguistic structure. The development of these perceptual strategies is in turn associated with the development of the ability to access phonological structure in the speech signal, which is the first step in phonological processing. Phonological processing abilities are at the heart of many other language skills, such as verbal working memory, lexical retrieval, and comprehending speech in noisy environments.

Appreciating the intricate links between various aspects of language processing should help us design better diagnostic tools and strategies for intervention. In recent years, the need for detailed descriptions of children's language problems (for such purposes as justifying placement in special programs and writing detailed individualized educational plans) has led to the development of new measurement tools that evaluate very specific skills. However, while it may be appropriate and even desirable to assess discrete speech and language skills, this practice has unfortunately contributed to an approach to intervention that focuses on remediating skills in isolation. There has been a proliferation of programs that would have children practicing, in a rote manner, isolated skills, with the expectation that the child will then be able to combine these skills to achieve language competency. This approach cannot work. It is analogous to learning the separate skills required for skiing (i.e., balancing, poling, moving hips and knees) in a gym, and then being expected to combine these skills on the mountain. Anyone who has taken a skiing lesson knows that students are skiing from the start. Instruction is ongoing and contextual. Similarly, language intervention requires a knowledgeable clinician engaging the child in natural language activities, but with an awareness of the language level of that child.

Thinking specifically of perceptual weighting strategies for speech, there is no way to facilitate the developmental changes documented here other than through the general enrichment of the language environment. Upon recalling the program that facilitated my own daughter's language development, several strategies stand out as having been used routinely. These are listed below. Regardless of which of these activities was used, however, the staff was constantly modeling syntactic and grammatical structures appropriate for the objectives set for each child, and drawing the children's attention to various levels of

linguistic structure (i.e., phrases, words, syllables, onsets and rhymes, or phonetic units), as appropriate for each child's language abilities.

- Personal narratives: Frequently, a staff member (e.g., teacher, speech-language pathologist, aide) modeled a personal narrative for the children, relating a story from his or her own life. Then, each child was encouraged to provide a narrative.
- Stories from pictures: Children picked pictures from magazines, catalogs, or the school's supply of photographs (the staff frequently photographed activities around the school). Each child was then encouraged to generate a story about that picture, which a staff member wrote below the picture. Finally, the child would be asked to retell the story a day or two later.
- Book reading: Naturally, books were read frequently to the children.
- "Plan, do, review" activities: These activities could involve anything, including walks in the park, cooking, or art projects.
- Role playing or imaginative play: As with all other activities, this sort of play merely served as a forum for working on individual language objectives.

Although my daughter has a peripheral impairment that degrades the acoustic signal available to her, these strategies promoted language development. Investigators studying speech perception by listeners with normal hearing have shown that the acoustic signal can be tremendously degraded by using either sine wave analogues of speech (e.g., Remez, Rubin, Berns, Pardo, & Lang, 1994) or modulated-noise analogues (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995), and listeners still recover phonetic structure. However, the analogues must mimic the signal of the listener's native language. In other words, the most critical aspect of speech perception is organizing the signal as appropriate for one's native language, rather than getting every detail of the signal. Listeners with auditory deficits that constrain their access to portions of the . acoustic speech signal can learn how to organize the (degraded) signal that they do receive in order to access phonetic structure, just as the listeners in the Remez et al. and Shannon et al. studies were able to do, but it requires even more experience than the typical child receives.

In summary, there is no substitute for a good clinician. Rote practice on isolated skills, detached from meaningful, natural contexts, cannot work. Language intervention must be structured to facilitate the child's honing of the perceptual skills needed to process the acoustic speech signal such that phonetic structure can emerge.

ACKNOWLEDGMENTS

This work was supported by Research Grant R01-DC-00633 from the National Institute on Deafness and Other Communication Disorders to the author, and by Grant P60-00982 from the National Institute on Deafness and Other Communication Disorders to Boys

Town National Research Hospital. The constructive comments of Maureen Higgins, Doug Keefe, and Carol J. Strong on earlier versions of this manuscript are gratefully acknowledged.

EREVISE OF BUILDING

REFERENCES

- Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. Quarterly Journal of Experimental Psychology, 18, 362-365.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575-589.
- Bar-Shalom, E. G., Crain, S., & Shankweiler, D. P. (1993). A comparison of comprehension and production abilities of good and poor readers. *Applied Psycholinguistics*, 14, 197–227.
- Bishop, D. V. M., Carlyon, R. P., Deeks, J. M., & Bishop, S. J. (1999). Auditory temporal processing impairment: Neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language, and Hearing Research*, 42, 1295-1310.
- Blumstein, S. E., & Stevens, K. N. (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America*, 67, 648-662.
- Brady, S. A., & Shankweiler, D. P. (Eds.). (1991). Phonological processes in literacy. Hillsdale, NJ: Lawrence Earlbaum.
- Brady, S., Shankweiler, D., & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology*, 35, 345-367.
- Byrne, B. (1981). Deficient syntactic control in poor readers: Is a weak phonetic memory code responsible? *Applied Psycholinguistics*, 2, 201–212.
- Campbell, R., & Dodd, B. (1980). Hearing by eye. Quarterly Journal of Experimental Psychology, 32, 85-99.
- Charles-Luce, J., & Luce, P. A. (1990). Similarity neighborhoods of words in young children's lexicons. *Journal of Child Language*, 17, 205-215.
- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22, 129–159.
- Chermak, G. D., & Musiek, F. (1997). Central auditory processing disorders. San Diego, CA: Singular.
- Cornelissen, P. L., Hansen, P. C., Bradley, L., & Stein, J. F. (1996). Analysis of perceptual confusions between nine sets of consonant-vowel sounds in normal and dyslexic adults. *Cognition*, 59, 275-306.
- Crowther, C. S., & Mann, V. A. (1994). Use of vocalic cues to consonant voicing and native language background: The influence of experimental design. *Perception & Psychophysics*, 55, 513-525.
- Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. Clinical Neurophysiology, 112, 758-767.
- Damian, M. R., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 345-361.
- Denkla, M. B., & Rudel, R. G. (1976). Naming of object-drawings by dyslexic and other learning disabled children. Brain and Language, 3, 1-15.

- Eilers, R. E., Gavin, W., & Oller, D. K. (1982). Cross-linguistic perception in infancy: Early effects of linguistic experience.

 Journal of Child Language, 9, 289-302.
- Eimas, P. D., & Miller, J. L. (1980). Discrimination of information for manner of articulation. *Infant Behavior and Development*, 3, 367-375.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in early infancy. Science, 171, 304-306.
- Flege, J., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. Language and Speech, 24, 125-146.
- Fox, B., & Routh, D. K. (1975). Analyzing spoken language into words, syllables, and phonemes: A developmental study. *Journal* of Psycholinguistic Research, 4, 331-342.
- Friel-Patti, S. (1999). Clinical decision-making in the assessment and intervention of central auditory processing disorders. Language, Speech, and Hearing Services in Schools, 30, 345-352.
- Hall, J. W., Wilson, K. P., Humphreys, M. S., Tinzmann, M. B., & Bowyer, P. M. (1983). Phonemic-similarity effects in good vs. poor readers. *Memory & Cognition*, 11, 520-527.
- Harris, K. S. (1958). Cues for the discrimination of American English fricatives in spoken syllables. Language and Speech, 1, 1-7.
- Hess, R. D., & Shipman, V. (1965). Early experience and the socialization of cognitive modes in children. *Child Development*, 36, 86-99.
- Jastak, S., & Wilkinson, G. S. (1984). The Wide Range Achievement Test-Revised. Wilmington, DE: Jastak Associates.
- Jerger, J., & Musiek, F. (2000). Report of the consensus conference on the diagnosis of auditory processing disorders in school-aged children. Journal of the American Academy of Audiology, 11, 467-474.
- Katz, R. B. (1986). Phonological deficiencies in children with reading disability: Evidence from an object-naming task. Cognition, 22, 225-257.
- Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 73, 322–335.
- Kohn, S. E., Wingfield, A., Menn, L., Goodglass, H., Berko Gleason, J., & Hyde, M. (1987). Lexical retrieval: The tip of the tongue phenomenon. Applied Psycholinguistics, 8, 245-266.
- Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of* the Acoustical Society of America, 66, 1668-1679.
- Laosa, L. M. (1980). Maternal teaching strategies in Chicano and Anglo-American families: The influence of culture and education on maternal behavior. Child Development, 51, 759-765.
- Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four- to six-and-a-half-month-old infants from Spanish environments. *Journal of Experimental Child Psychol*ogy, 20, 215-225.
- Liberman, A. M. (1957). Some results of research on speech perception. *Journal of the Acoustical Society of America*, 29, 117-123.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. Psychological Review, 74, 431-461.
- Liberman, I. Y., Shankweiler, D. P., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18, 201-212.

- Mann, V. A., & Liberman, I. Y. (1984). Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities*, 17, 592-599.
- Mann, V. A., & Repp, B. H. (1980). Influence of vocalic context on perception of the [S]-[s] distinction. Perception & Psychophysics, 28, 213-228.
- Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G. Altman (Ed.), Cognitive models of speech perception: Psycholinguistic and computational perspectives (pp. 148-172). Cambridge, MA: MIT Press.
- 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.
- Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: Auditory processing or phonological coding? *Journal of Experimental Child Psychology*, 64, 199-231.
- Moffitt, A. R. (1971). Consonant cue perception by twenty- to twenty-four-week-old infants. Child Development, 42, 717-731.
- Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics*, 20, 351–382.
- Nittrouer, S. (1996). The relation between speech perception and phonemic awareness: Evidence from low-SES children and children with chronic OM. *Journal of Speech and Hearing Research*, 39, 1059–1070.
- Nittrouer, S. (1999). Do temporal processing deficits cause phonological processing problems? *Journal of Speech, Language, and Hearing Research*, 42, 925-942.
- Nittrouer, S. (2001). Challenging the notion of innate phonetic boundaries. *Journal of the Acoustical Society of America*, 110, 1598-1605.
- Nittrouer, S. (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *Journal of the Acoustical Society of America 112*, 711-719.
- Nittrouer, S., & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America*, 101, 2253–2266.
- Nittrouer, S., & Miller, M. E. (1999). The development of phonemic coding strategies for serial recall. *Applied Psycholinguistics*, 20, 563-588.
- Nittrouer, S., Miller, M. E., Crowther, C. S., & Manhart, M. J. (2000). The effect of segmental order on fricative labeling by children and adults. *Perception & Psychophysics*, 62, 266-284.
- Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects on the perception of fricatives by children and adults. *Journal of Speech and Hearing Research*, 30, 319-329.
- Peter, M., & Turvey, M. T. (1994). Phonological codes are early sources of constraint in visual semantic categorization. *Perception & Psychophysics*, 55, 497-504.
- Remez, R. E., Rubin, P. E., Berns, S. M., Pardo, J. S., & Lang, J. M. (1994). On the perceptual organization of speech. Psychological Review, 101, 129-156.
- Rosen, S., & Manganari, E. (2001). Is there a relationship between speech and nonspeech auditory processing in children with dyslexia? *Journal of Speech, Language, and Hearing Research*, 44, 720–736.

- Schachter, F. F. (1979). Everyday mother talk to toddlers: Early intervention. New York: Academic Press.
- Schachter, F. F., & Strage, A. A. (1982). Adults' talk and children's language development. In S. G. Moore & C. R. Cooper (Eds.), *The young child: Reviews of research* (Vol. 3, pp. 79–96). Washington, DC: National Association for the Education of Young Children.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. Science, 270, 303-304.
- Smith, S. T., Mann, V. A., & Shankweiler, D. (1986). Spoken language comprehension by good and poor readers: A study with the Token Test. Cortex, 22, 627-632.
- Spoehr, K. T., & Corin, W. J. (1978). The stimulus suffix effect as a memory coding phenomenon. *Memory & Cognition*, 6, 583-589.
- Spring, C., & Perry, L. (1983). Naming speed and serial recall in poor and adequate readers. Contemporary Educational Psychology, 8, 141-145.
- Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 64, 1358–1368.
- Studdert-Kennedy, M. (1983). Limits on alternative auditory representations of speech. In C. W. Parkins & S. W. Anderson (Eds.), Cochlear prostheses: An international symposium (Vol. 405, pp. 33–38). New York: The New York Academy of Sciences
- Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature*, 241, 468-469.
- Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia*, 12, 83–93.

- Treiman, R. (1995). Errors in short-term memory for speech: A developmental study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1197-1206.
- Treiman, R., & Zukowski, A. (1991). Levels of phonological awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy* (pp. 67-84). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Van Orden, G. C. (1987). A Rows is a Rose: Spelling, sound, and reading. *Memory & Cognition*, 15, 181-198.
- Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. Journal of Experimental Psychology: Learning, Memory, & Cognition, 14, 371-385.
- Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101, 192–212.
- Walley, A. C., Smith, L. B., & Jusczyk, P. W. (1986). The role of phonemes and syllables in the perceived similarity of speech sounds for children. *Memory & Cognition*, 14, 220-229.
- 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.

Received March 16, 2002 Accepted July 23, 2002 DOI:10.1044/0161-1461(2002/020)

Contact author: Susan Nittrouer, Center for Persons with Disabilities, Utah State University, 6840 Old Main Hill, Logan, UT 84322-6840. E-mail: Nittrouer@cpd2.usu.edu



COPYRIGHT INFORMATION

TITLE: From Ear to Cortex: A Perspective on What Clinicians

Need to Understand About Speech Perception and

Language Processing

SOURCE: Language, Speech, and Hearing Services in Schools 33

no4 O 2002

WN: 0227401952001

(C) The American-Speech-Language-Hearing Association is the publisher of this article and holder of the copyright. Further reproduction of this article in violation of copyright is prohibited without the consent of the publisher. To contact the publisher: http://www.asha.org/.

Copyright 1982-2002 The H.W. Wilson Company. All rights reserved.