Vol. 19, No. 3, pp. 75–116
August 2012

PERSPECTIVES on Language Learning and Education

AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION

SPECIAL INTEREST GROUP 1

In This Issue

A New Perspective on Developmental Language Problems: Perceptual Organization Deficits

Susan Nittrouer

Department of Otolaryngology, The Ohio State University
Columbus, OH
Disclosure: Susan Nittrouer has no financial or nonfinancial relationships related to the content of this article.

Children with a variety of language-related problems, including dyslexia, experience difficulty processing the acoustic speech signal, leading to proposals of diagnostic entities known as auditory processing deficits*. Although descriptions of these deficits vary across accounts, most hinge on the idea that problems arise at the level of detecting and/or discriminating sensory inputs. In this article, the author re-examines that idea and proposes that the difficulty more likely arises in how those sensations get organized into service for auditory comprehension of language.*

Speech-language pathologists working in the schools have widely varying caseloads, including children who are struggling to learn language, especially reading. Objectives for intervention with these children typically involve trying to hone their sensitivity to phonemic structure. Although specific strategies for meeting this objective have changed over the years, the principles underlying that objective have remained consistent. This article examines the history of those principles and suggests how methods of intervention would be modified if a new perspective were to be embraced. The focus of this article is narrowly on auditory processing of speech signals, but that focus really serves as a proxy for broader issues. By exploring ideas and controversies surrounding auditory processing and related disorders, we are able to examine perspectives about speech perception and psycholinguistic processing and the role these perspectives play in shaping our interventions.

A Focus on Segments

For decades, views about the structure of spoken language have been based on the phoneme and the notion that phonemic segments exist physically in the speech signal, even though the cues to those segments overlap in time (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). According to this perspective, listeners extricate those cues from the ongoing speech stream and use them to recover the associated phonemes, a process termed *decoding*. Informal inspection of our own speech perception suggests that the perspective is correct because most of us have a strong impression that what we are doing when we listen to speech is recovering a string of separate phonemes. Empirical evidence for the reality of phonemes has been provided through examination of the kinds of production errors people make: Like segments are transposed with each other (i.e., consonant for consonant; vowel for vowel), while syllabic place is preserved (Shattuck-Hufnagel, 1983). These consistencies of error suggest that speakers have well-defined phonemic representations.

The poor phonological skills of individuals with language problems similarly suggest that phonemes must be accessible in the speech signal. Adults and children with dyslexia exhibit severe difficulties in recognizing and manipulating phonemes, both overtly and implicitly, and those difficulties seem to underlie reading problems (e.g., Boada & Pennington, 2006; Liberman & Shankweiler, 1985; Pratt & Brady, 1988; Scarborough, 1990; Wagner, 1986). Clearly, then, having access to phonemic segments is important to the reading process, at least for alphabetic orthographies. Of course, the very fact that alphabetic orthographies exist at all means that phonemic segments must be recoverable from the speech stream; otherwise, all orthographies would be syllabaries or logographies.

However, as strong as these sources of support are for the reality of the phoneme, they do not necessarily mean that language users always recover phonemic segments from the speech signal and use them for psycholinguistic processes such as lexical access. For purposes other than reading, the evidence is actually mixed concerning whether or not phonemic segments are obligatorily invoked. One example of a psycholinguistic process other than reading where recovery of phonemic units has been viewed as essential is the short-term storage and recall of words. In memory experiments, participants may hear a string of words and are asked to recall them in order. An advantage is typically found in short-term recall for phonologically non-confusable items, such as non-rhyming words, over confusable items, such as rhyming words (e.g., Conrad & Hull, 1964; Nittrouer & Miller, 1999; Salame & Baddeley, 1986). That kind of finding supports models of perception in which phonemes are recovered upon hearing a speech signal and used to store words in a short-term memory buffer (e.g., Baddeley, 1992). But even that evidence for the role of phonemes in psycholinguistic processing is not incontrovertible.

According to traditional accounts, acoustic structure in the signal serves to specify phonemic structure, but nothing more. Once a string of phonemes has been recovered, acoustic information is discarded, according to traditional accounts. However, such models have been challenged by a few scientists who point to evidence that acoustic information unrelated to phonemes, such as speaker identity, influences how efficiently words are stored in memory buffers and subsequently recovered (e.g., Coleman, 2002; Port, 2007). For example, storage and retrieval of word lists is better for single-talker lists than for multi-talker lists, even when words in both lists are readily recognized (e.g., Martin, Mullennix, Pisoni, & Summers, 1989; Palmeri, Goldinger, & Pisoni, 1993). Consequently, it cannot be the case that phonemes are recovered from the signal and alone used for all subsequent processing. Other kinds of acoustic structure must also be retained, and that structure is apparently recruited into psycholinguistic service.

Nonetheless, traditional accounts about the centrality of phonemes in psycholinguistic processing have shaped the way that speech perception research has been conducted since its inception. For the most part, investigators have searched for the acoustic properties used by listeners in deciding the identity of phonemes. These acoustic properties, termed *cues*, are brief and spectrally discrete in most accounts of speech perception (Repp, 1983). Most experiments exploring possible acoustic cues and their role in how listeners recognize phonemes are conducted in essentially the same way: by manipulating one of those cues along a continuum in steps of equal size, embedding each acoustic bit that results from that process into a base syllable in which all other acoustic structure is held constant, and measuring the effect of that cue in a binary choice, phoneme labeling task. These are the classic categorical perception experiments, which show that sharp labeling functions are derived when stimuli from acoustic continua are heard—at least, those are the findings for typical adults. For listeners with specific kinds of language problems, such as dyslexia, those labeling functions are often shallower than normal (e.g., Godfrey, Syrdal-Lasky, Millay, & Knox, 1981; Nittrouer, 1999; Werker & Tees, 1987). Because traditional models of speech perception suggest that the recovery of phonemes derives automatically from detecting those acoustic cues, evidence of shallower-than-normal labeling functions can only mean a lack of sensitivity to those cues, according to these models.

In turn, poor sensitivity to acoustic cues is blamed for the psycholinguistic processing problems of children with certain language deficits, including dyslexia. This line of reasoning is the basis of the diagnostic entity known as *auditory processing disorders*. Associated auditory interventions to treat those disorders often focus on improving sensitivity to acoustic, usually nonspeech, signal structure, in hopes that it will improve the ability to recover phonemes.

Is the acoustic-phonemic link automatic?

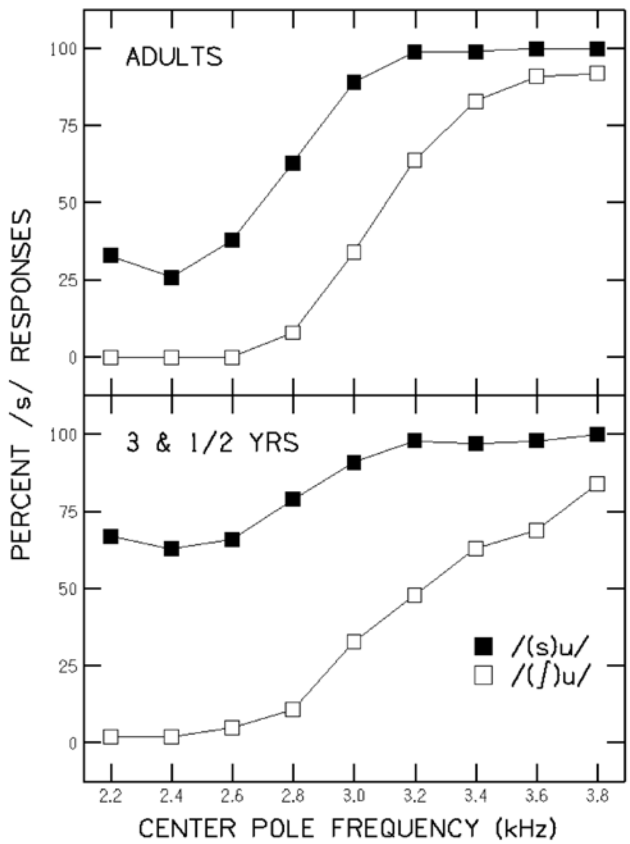
In contrast to the traditional account outlined above, studies with various populations of listeners show that the ability to discriminate acoustic cues does not automatically translate into the ability to recognize phonemes. Some of this evidence comes from studies of cross-linguistic speech perception. As early as 1975, Miyawaki and colleagues showed that native Japanese speakers failed to use syllable-initial third-formant (F3) glides to label stimuli along a *rock* to *lock* continuum, even though those same subjects were perfectly able to discriminate the F3 glides when presented in isolation. Based on that finding, it was concluded that recovery of phonemic structure is not ensured by sensitivity to acoustic cues. It appears that listeners can selectively attend or not attend to specific cues, something that suggests a top-down influence on how acoustic structure is handled perceptually.

The idea that listeners selectively attend to (i.e., weight) acoustic cues in the signal was expanded and supported when Nittrouer and Studdert-Kennedy (1987) first showed that children and adults weight differently the cues used in fricative labeling. In that study, an acoustic continuum of synthetic fricative noises was developed to span the range from [ʃ] to [s]. The noises from that continuum were combined with vocalic portions taken from a male speaker saying consonant-vowel syllables starting with [ʃ] or [s], which meant that the vocalic portions had formant transitions appropriate for either one or the other fricative. Until that report, most speech scientists believed that the spectral structure of the fricative noise itself (i.e., how high or low the noise was in frequency) was the primary cue to fricative identity (e.g., Harris, 1958), meaning that the greatest perceptual weight was given to the spectral structure of those noises.

However, it was also known that adults show small shifts in where they place their labeling functions depending on whether formant transitions signal syllable-initial [ʃ] or [s] (Mann & Repp, 1980). The prediction going into that 1987 experiment by Nittrouer and Studdert-Kennedy was that children would either weight similarly or, if anything, weight more strongly than adults that primary cue of fricative-noise spectrum and show less influence of the formant transitions. These latter cues were described as secondary, and the presumption was that children need to learn how they co-vary with the fricative-noise spectra before they can recruit them into perceptual service.

As it turned out, results were opposite to predictions. Children weighted fricative-noise spectra less and formant transitions more than adults did. These outcomes are illustrated in Figure 1. Shallower labeling functions for children than for adults indicate that children assigned less weight to fricative-noise spectra, the cue represented on the *x*-axis. Greater separation between functions means that children weighted formant transitions more strongly, because the two labeling functions are from stimuli with different formant transitions. In subsequent experiments, it was shown that this difference in weighting of cues to fricative identity could not be explained by sensitivity to those cues because adults and children showed similar sensitivity (Nittrouer, 1996a). That finding mirrored trends for second-language speakers. Still other experiments reported that the weighting of formant transitions diminishes while the weighting of fricative-noise spectra increases throughout childhood. The sum of all these results led to the idea that the observed developmental shift in weighting strategies is related to increasing awareness of phonemic structure (Mayo, Scobbie, Hewlett, & Waters, 2003; Nittrouer, 1992, 1996b). If that is true, then children lagging in their development of sensitivity to phonemic structure should show immature weighting strategies compared to age-matched peers. That prediction was tested and supported for children defined by their poor phonemic awareness (Nittrouer, 1996b, 1999), as well as for children diagnosed with dyslexia (Boada & Pennington, 2006; Johnson, Pennington, Lowenstein, & Nittrouer, 2011). These results form support for the position that the problems exhibited by children with some kinds of language deficits might stem from how they are processing the signal perceptually, rather than from how sensitive they are to acoustic structure.

Figure 1. Labeling functions for adults (top) and 3-year-olds (bottom). The symbol in parentheses indicates the fricative context from which syllable-initial formants were derived.



Perceptual Organization of Sensory Information Explains Cognitive Representations

So far, we have seen that traditional accounts of speech perception propose listeners harvest acoustic cues from the signal, and from those cues directly recover strings of phonemic segments. According to those accounts, no other perceptual effect was presumed to influence the relation between sensory information and cognitive representation. However, that view of speech perception is out of step with most accounts of perceptual processing. Early in the twentieth century, the Gestalt psychologist Max Wertheimer mused about how numerous separate sensations merged into a few integrated forms as he stood looking out his window (Wertheimer, 1923). Although hundreds of individual colors and nuances of brightness reached his eye, he saw only a house, a tree, and the sky. With some effort, he went on to report, he discovered that sensory structure from physically disconnected objects could merge to form new objects. For example, a bit of the window sash inside the house and a tree branch outside combined in his visual field to form an N. This last effect emphasizes the notion that perceivers shape the way sensory information is organized.

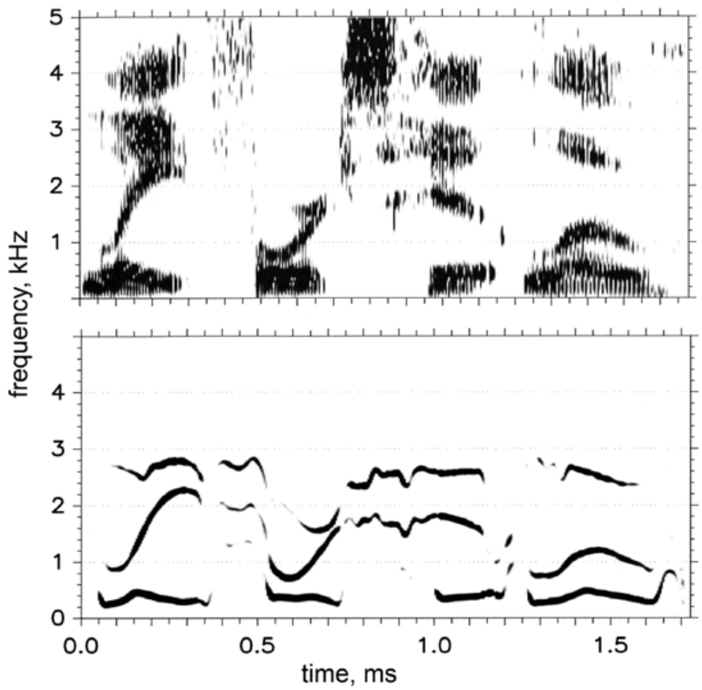
Demonstrations of how perceivers shape the way sensory information is organized are abundant for visual signals. Figure 2 shows a drawing that can be organized to reveal either a woman looking away, with only her left ear in view, or a woman in profile, so that the ear becomes an eye. In either case, the sensory information reaching the retina is exactly the same; how that information is organized determines which image is obtained. And, with some effort, one can alternate between the two images in rapid succession, creating bistability of the visual form.

Figure 2. Ambiguous picture illustrating how perceptual organization works in vision.

From *American Journal of Psychology.* Copyright 1930 by the Board of Trustees of the University of Illinois. Used with permission of the University of Illinois Press.

When it comes to speech signals, evidence that listeners perceptually organize the sensory input can similarly be proffered, but it requires some signal processing. Natural speech has several qualities that makes it unique among acoustic signals, most of which stem from the fact that speech production can be viewed as arising from a largely independent source and filter (Fant, 1960). As a result, the sound source for much of the signal is pulsed, due to the opening and closing of the glottis, and there is clear harmonic structure. The vocal tract filter shapes those harmonics, amplifying some and attenuating others, to create the formants we are accustomed to seeing on spectrograms, like the one shown on the top of Figure 3. In all cases, those formants have broad bandwidths. If natural speech is not processed in some way, recognition is too accurate to permit the investigation of perceptual organizing strategies. Thus, the goal of signal processing in experiments designed to examine perceptual organization of speech is to diminish the speech-like qualities of the signal, while maintaining some pre-specified aspect of signal structure. In 1981, several scientists found they were able to accomplish this goal by replacing the first three formants with sine waves replicating the center frequencies of those formants, and eliminating all other cues. These signals lack harmonic structure and because sine waves are only 1 Hz wide, the formants have no bandwidth. The signal structure maintained by this processing is the time-varying nature of the formants. This kind of signal is shown on the bottom of Figure 3. Although it may require a little practice, listeners are typically able to organize these three time-varying sine waves in a manner that allows them to recover lexical information. For a demonstration, readers can go to www.speechdevelopment.org/OSSO, click on the link titled, "Normal Development," and scroll to the bottom of the page.

Figure 3. Spectrograms of the sentence, "Late forks hit low." Natural speech is represented on top; sine wave speech on the bottom.



In 2001, a group of investigators asked if listeners are able to hold on to their auditory impressions of these sine waves, even when they are harnessed to a speech perception task (Remez, Pardo, Piorkowski, & Rubin). As the Miyawaki et al. (1975) experiment demonstrated, listeners are able to report on the auditory qualities of an isolated formant, such as whether it is relatively high or low in pitch. Once that formant is integrated into a complete speech signal, however, perceptual access to its acoustic qualities is lost. For sine wave replicas of words, on the other hand, Remez et al. (2001) discovered that listeners can concurrently attend to both the lexical and auditory qualities of the signals. This finding in the domain of auditory speech perception is directly analogous to what happens when we rapidly alternate between the two images found in Figure 2. The finding with sine wave speech demonstrates that listeners perceptually organize signals in ways to facilitate their abilities to recover linguistically relevant forms, or objects.

How does this idea help explain problems of children with auditory processing deficits?

As with perceptual weighting strategies for discrete acoustic cues to phonemic identity, there is evidence that these more global strategies of perceptual organization vary across languages. To extend our visual analogy, this is equivalent to suggesting that efficient processing of one language, say English, favors a perceptual strategy that evokes the image of the woman looking over her shoulder, while processing of a different language favors the strategy that evokes the woman in profile. The veracity of this account has been demonstrated primarily with a signal processing strategy other than sine waves, namely amplitude envelopes. These latter signals are created by dividing the speech spectrum into some number of channels (usually between four and sixteen), and recovering only amplitude structure (i.e., the envelopes) from each channel. That set of envelopes is then superimposed on noise bands filtered to match the channels into which the speech spectrum had been divided. The resulting signals lack the qualities of typical speech, as sine wave replicas do, and it takes some effort upon first hearing them to organize them according to speech-governed strategies. To test this oneself, the reader can return to the website listed above, where the sine wave signals were found, and listen to a sample of amplitude-envelope speech.

In 2009, Nittrouer, Lowenstein, and Packer reported that non-native, albeit fluent, speakers of English were poorer at recognizing these signals than were native, adult speakers of English. A study by Fu, Zeng, Shannon, and Soli (1998) had previously demonstrated that speakers of the same language background as the non-native speakers in Nittrouer et al. (2009) perform better when listening to amplitude envelopes of speech in their first language; therefore, this outcome cannot be attributable to the non-native speakers' simply having difficulty with amplitude-envelope speech. Seven-year-old participants in the Nittrouer et al. study who were native speakers of English performed similarly to the non-native speaking adults. Similar results have been reported elsewhere for children (e.g., Eisenberg, Shannon, Schaefer Martinez, Wygonski, & Boothroyd, 2000; Nittrouer & Lowenstein, 2010), so it seems appropriate to suggest that the perceptual strategies used to organize speech-like signals in order to recover linguistic structure are acquired over the course of language acquisition. As with perceptual weighting strategies for acoustic cues, this developmental outcome raised the possibility that children who lag in language development may be delayed in the acquisition of appropriate organizational strategies. To test that hypothesis, Johnson et al. (2011) presented sentences processed to preserve only amplitude envelopes to children with histories of dyslexia, speech sound disorder, or both, as well as to children with typical language development. Children in all three disordered groups were poorer than those developing typically at recognizing the processed sentences, supporting the hypothesis that they had less effective perceptual organization strategies for speech signals. Thus, it seems that the underlying processing difficulty encountered by children with a variety of language deficits rests with being able to organize the sensory information reaching them.

Experiments with processed signals not only inform us about how listeners need to organize the sensory information they receive, they also instruct us about the sorts of representations derived as a result. These experiments with sine waves or amplitude envelopes commonly use sentence-length materials. Recognition accuracy drops precipitously when phoneme recognition is the goal (Eisenberg et al., 2000; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995), suggesting that these strategies of perceptual organization operate over signal stretches longer than a phoneme or two. Procedures and outcomes of these experiments reinforce the notion that listeners do not necessarily recover phonemic representations when listening to speech. In much of our daily lives, the units of analysis are likely lexical items in sentence-length context. Of course, recognizing words in sentences requires that certain cognitive and psycholinguistic operations be invoked, as well. For example, short-term recall permits the storage of sentence-length materials, upon which syntactic analyses can be done. In turn, a listener's knowledge of syntactic structure can be used to facilitate recognition in top-down fashion. An important aspect of the Johnson et al. (2011) study was that the authors were able to demonstrate that differences in recognition for the amplitude-envelope signals across the groups (typical language, reading, speech-sound, or both problems) could not be attributed to differences in other, related operations. The only supportable interpretation of results was that children with dyslexia and/or speech sound disorder have immature strategies for perceptual organization of speech-like signals.

Putting it all in Perspective

Two ideas have been addressed in this article. First, traditional models of speech perception were re-examined. These models propose that listeners recover strings of phonemes from the acoustic speech signal through automatic translations of acoustic cues into those phonemes. In place of such accounts, I proposed here that more flexible models of speech perception be entertained, ones suggesting that in much of our daily communication explicit phonemic structure is not required, and so not recovered. Rather, that level of structure is accessed only for some psycholinguistic tasks, such as reading.

The second notion that was challenged was that auditory sensitivity to acoustic cues in the speech signal is all that accounts for speech recognition abilities. If that were true, we would be able to assess an individual's speech recognition abilities simply by measuring auditory sensitivity. Instead, top-down perceptual processes are recruited into action. At the level of phonemic decision-making, listeners weight the variety of cues available for use. Across listeners, these weighting strategies are known to be influenced by listener age and language experience.

At a level above phoneme perception, listeners appear to apply organizational strategies to the sensory inputs reaching their nervous systems. For nonspeech acoustic signals, the principles that account for these organizational strategies are similar to those that account for organization of visual signals. The principles described by Wertheimer (1923) to describe how sensory input to the visual system gets organized have been appropriated and applied to auditory perception in a model known widely as Auditory Scene Analysis (Bregman, 1990). These principles include phenomena such as *common fate* (spectral components that start and stop at the same time are grouped together) and *proximity* (spectral components close in frequency are grouped together). However, these principles do not adequately account for how speech signals are organized. In a series of carefully designed experiments, Remez, Rubin, Berns, Pardo, and Lang (1994) were able to dismiss these principles one by one as explanation for how listeners organize speech-like signals. Instead, the very fact that a signal can be recognized as having been generated by a human vocal tract seems more critical to evoking the kind of perceptual organization applied to speech.

In sum, the major shift in perspective concerning what the deficit could be for children with language or reading problems is that they may have adequate sensitivity to the sensory information in the acoustic speech signal, but may fail to organize that information as needed for recovering linguistic structure, including but not limited to phonemes. These problems in perceptual organization would appear as problems in sensitivity to acoustic structure because of the methods we have traditionally used in experiments and in the clinic, and the interpretations we are predisposed to apply to the resulting data. For example, in categorical perception experiments, labeling functions are shallower for listeners when the acoustic cue manipulated across the continuum (represented on the *x*-axis) is one that those particular listeners do not weight particularly strongly. It is not that those listeners lack sensitivity to small changes in that acoustic property; rather, they may not attend to that property as strongly as other listeners when it is assembled into a phonetic object.

Implications for Clinical Treatment

If children with some language or related deficits indeed have difficulties organizing sensory inputs properly in order to recover linguistic structure in an efficient manner, the question arises as to what should be done clinically. Certainly, activities that help focus children's attention on phonemic, lexical, and syntactic structure are useful, but attention should be spread across these levels of structure and not focused singly on phonemes. Training should be provided on psycholinguistic tasks other than simple recognition and/or production. For example, practice should emphasize skills such as short-term memory of linguistically relevant units. Accommodations that ensure that an optimal signal is available to these children, even in less-than-optimal environments such as classrooms, are also helpful. Presumably, once a child discovers the strategies needed to organize speech signals effectively in optimal conditions, transferring those strategies to less-than-optimal conditions will be easier.

One kind of activity that is likely not useful for children with language problems rooted in deficits of perceptual organization involves discrimination tasks with nonspeech acoustic signals. Remez et al. (1994) artfully demonstrated that the organizational strategies that serve speech perception are not based on general auditory processing. Rather, speech-specific strategies are invoked precisely when the listener recognizes that the signal was generated by a human speaker. Consequently, any clinical intervention that can facilitate children learning how the signal emanating from a speaker is structured will be useful. These methods would involve letting children have access to audiovisual speech signals, focusing intervention on long signal stretches, and providing children with ample opportunity to produce the kinds of speech we want them to learn to process more efficiently. This last suggestion emphasizes what has come to be generally recognized as a strong perception-production link.

Summary

The ideas discussed here suggest that speech-language pathologists have an important role to play in helping children struggling to learn language who are diagnosed with auditory processing deficits. Although these children might have difficulty processing acoustic speech signals, that difficulty likely does not arise from deficits we typically associate with lesions in the auditory system. Instead, the problems faced by these children appear to be due to difficulty organizing these signals perceptually in the most efficient manner to evoke linguistic form.

Acknowledgments

This work was supported by Grant No. R01 DC00633 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health. The author thanks Amanda Caldwell, Joanna Lowenstein, and Rebecca McCauley for helpful suggestions on an earlier version.

References

Baddeley, A. (1992). Working memory. *Science, 255*, 556–559.

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology, 95*, 153–193.

Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: The MIT Press.

Coleman, J. (2002). Phonetic representations in the mental lexicon. In J. Durand & B. Laks (Eds.), *Phonetics, phonology, and cognition* (pp. 96–130). Oxford, UK: Oxford University Press.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55*, 429–432.

Eisenberg, L. S., Shannon, R. V., Schaefer Martinez, A., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *Journal of the Acoustical Society of America, 107*, 2704–2710.

Fant, G. (1960). *Acoustic theory of speech production*. The Hague, Netherlands: Mouton.

Fu, Q., Zeng, F. G., Shannon, R. V., & Soli, S. D. (1998). Importance of tonal envelope cues in Chinese speech recognition. *Journal of the Acoustical Society of America, 104*, 505–510.

Godfrey, J. J., Syrdal-Lasky, A. K., Millay, K. K., & Knox, C. M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology, 32*, 401–424.

Harris, K. S. (1958). Cues for the discrimination of American English fricatives in spoken syllables. *Language and Speech, 1*, 1–7.

Johnson, E. P., Pennington, B. F., Lowenstein, J. H., & Nittrouer, S. (2011). Sensitivity to structure in the speech signal by children with speech sound disorder and reading disability. *Journal of Communication Disorders, 44*, 294–314.

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. (1985). Phonology and the problems of learning to read and write. *Remedial and Special Education, 6*, 8–17.

Mann, V. A., & Repp, B. H. (1980). Influence of vocalic context on perception of the [zh]-[s] distinction. *Perception & Psychophysics, 28*, 213–228.

Martin, C. S., Mullennix, J. W., Pisoni, D. B., & Summers, W. V. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 676–684.

Mayo, C., Scobbie, J. M., Hewlett, N., & Waters, D. (2003). The influence of phonemic awareness development on acoustic cue weighting strategies in children's speech perception. *Journal of Speech, Language, and Hearing Research, 46*, 1184–1196.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18*, 331–340.

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. (1996a). Discriminability and perceptual weighting of some acoustic cues to speech perception by three-year-olds. *Journal of Speech and Hearing Research, 39*, 278–297.

Nittrouer, S. (1996b). 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., & Lowenstein, J. H. (2010). Learning to perceptually organize speech signals in native fashion. *Journal of the Acoustical Society of America, 127*, 1624–1635.

Nittrouer, S., Lowenstein, J. H., & Packer, R. (2009). Children discover the spectral skeletons in their native language before the amplitude envelopes. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1245–1253.

Nittrouer, S., & Miller, M. E. (1999). The development of phonemic coding strategies for serial recall. *Applied Psycholinguistics, 20*, 563–588.

Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research, 30*, 319–329.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 309–328.

Port, R. (2007). How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology, 25*, 143–170.

Pratt, A. C., & Brady, S. (1988). Relation of phonological awareness to reading disability in children and adults. *Journal of Educational Psychology, 80*, 319–323.

Remez, R. E., Pardo, J. S., Piorkowski, R. L., & Rubin, P. E. (2001). On the bistability of sine wave analogues of speech. *Psychological Science, 12*, 24–29.

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.

Repp, B. H. (1983). Trading relations among acoustic cues in speech perception are largely a result of phonetic categorization. *Speech Communication, 2*, 341–362.

Salame, P., & Baddeley, A. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society, 24*, 263–265.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development, 61*, 1728–1743.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*, 303–304.

Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 109–136). New York, NY: Springer-Verlag.

Wagner, R. K. (1986). Phonological processing abilities and reading: Implications for disabled readers. *Journal of Learning Disabilities, 19*, 623–630.

Werker, J. F., & Tees, R. C. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology, 41*, 48–61.

Wertheimer, M. (1923). Laws of organization in perceptual forms. Originally published in *Psychologische Forschung*. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (4th ed., pp. 71–88). New York, NY: The Humanities Press.