Deficits in phoneme awareness do not arise from failures in rapid auditory processing

MICHAEL STUDDERT-KENNEDY
Haskins Laboratories, New Haven, Connecticut, USA

Abstract. Many studies have found that phonological deficits in poor readers are associated with deficits in speech perception. Two hypotheses have been proposed concerning the nature of the latter: a speech-specific and a general auditory hypothesis. The main topic of the paper is the general auditory hypothesis and its special form as proposed by Tallal and her colleagues (1993). The paper reviews the evidence for these hypotheses and finds it to be either purely correlational or flawed by misinterpretation of results and/or lack of necessary experimental controls. Moreover, a recent control study, the first of its kind, found no support for Tallal's special form of the general auditory hypothesis. The paper concludes that deficits in speech perception often observed in impaired readers are phonetic (speech-specific), not auditory, in origin.

Keywords: Phoneme awareness, Non speech controls, Rapid auditory processing, Slowed speech

Introduction

In 1973 Isabelle Liberman proposed that the initial obstacle in learning to read lay in becoming aware that words can be analyzed into sequences of phonemes. The difficulty arose, she argued, because the gestures that compose consonants and vowels are intricately overlapped in the articulatory syllable, so that phonemes cannot be recovered by simple linear segmentation of the acoustic signal: Speech is not an acoustic alphabet (A. M. Liberman et al. 1967). Later work by Isabelle Liberman, Donald Shankweiler, and their colleagues at Haskins laboratories, and by researchers at many other laboratories, has established that phonological deficits are present in virtually every poor reader, and that degree of phoneme awareness is indeed the best single predictor of reading success (for review, see Brady & Shankweiler 1991).

Why is phoneme awareness difficult for some children to achieve? One account, developed at Haskins Laboratories over the past two decades, proposes that the deficit stems from poor speech perception. Poor speech perception gives rise both to 'fuzzy' or 'underspecified' lexical (and so phonological) representations and to weak verbal short term memory. These

in turn give rise to deficits in syntactic awareness and in comprehension in listening and/or reading.

Two hypotheses have been proposed concerning the nature of the deficit in speech perception: A general auditory and a speech-specific hypothesis. The speech-specific hypothesis (Liberman 1998) proposes that the deficit is in the phonetic transform from analog neural response pattern to digital lexical/phonological representation. Evidence from poor readers consistent with the speech-specific hypothesis includes: low perceptual performance on speech, but not on non-speech, under demanding conditions; low short-term memory for words, but not for non-verbal sounds or pictures; similar patterns of error in verbal short-term memory, whether words are heard or read, suggesting a deficit in the phonological representation common to vision and audition rather than independent deficits in both modes of input (for references supporting the statements of the preceding two paragraphs, see Mody, Studdert-Kennedy & Brady 1997).

The general auditory hypothesis, and its special form as propounded by Tallal and her colleagues (e.g., Tallal et al. 1993), is the main topic of what follows. This hypothesis holds that the deficit is not specific to speech, but lies in the neural transform from acoustic signal to auditory representation. Evidence consistent with the general auditory hypothesis comes from language-and/or reading-impaired subjects who tend statistically to be impaired in certain non-speech auditory tasks, including: temporal order judgments (TOJ) for rapidly presented complex tones differing in fundamental frequency (e.g., Reed 1989; Tallal 1980) and for rapidly presented acoustic clicks (Kinsbourne et al. 1991), choice reaction time to pure tones differing in frequency (Nicholson & Fawcett 1994), sensitivity to rate and depth of acoustic frequency modulation (Stein & McAnally 1995), and auditory localization (Hari & Kiesila 1996). Performance on several of these tasks has been found to be significantly correlated with performance on various linguistic tasks relevant to reading.

In interpreting such correlations, three cautions should be observed. First, correlation is not causation. A causal role for a particular auditory deficit in defective speech perception can be established only by demonstrating equivalent deficits in perceiving both speech and an acoustically matched non-speech control. Second, none of the auditory tasks listed above has any bearing on speech perception: There are no phonetic contrasts for which any of these tasks might serve as an appropriate non-speech control. Third, performance on TOJ and other perceptual tasks may vary with a general cognitive capacity (or its precursors in infancy), such as attentiveness or intelligence, or with a host of other factors (Locke 1998). If such variables are neither experimentally controlled nor statistically partialed out, correla-

tions are uninterpretable. Accordingly, whatever the utility of, say, 'auditory temporal processing' thresholds (Benasich 1998; Benasich & Tallal 1996) for predicting language deficits, their explanatory worth is nil in the absence of necessary controls.

I turn now to a brief history of the special form of the auditory hypothesis developed by Tallal and her colleagues (e.g., Tallal et al. 1993), a history without which the hypothesis cannot be fairly assessed (for a fuller account, see Mody et al. 1997; Studdert-Kennedy & Mody 1995). Tallal's hypothesis includes two key propositions: (1) specialization of the left cerebral hemisphere for speech perception is grounded in a prior specialization of that hemisphere for 'rapid auditory temporal processing', and (2) phonological deficits in some dysphasic children, some aphasic adults, and some impaired readers, or dyslexics, stem from deficits in 'rapid auditory temporal processing', such that they cannot easily perceive the rapid acoustic changes at the onset of stop-vowel syllables. I begin with the second proposition.

The rapid auditory processing hypothesis

Three basic studies from which all the later work springs (Tallal & Piercy 1973, 1974, 1975) compared dysphasic children with normal controls on temporal order judgment (TOJ) and discrimination for pairs of synthetic stimuli with 'long' (428 msec) and 'short' (8–305 msec) interstimulus intervals (ISI). The stimuli included 'short' (75 msec) and 'long' (250 msec) complex tones, differing in fundamental frequency, 'short' (43 msec) and 'long' (250 msec) steady-state vowels, and stop-vowel syllables (/ba,da/, 250 msec) with 'short' (43 msec) or 'long' (95 msec) first, second and third formant (F1, F2, F3) onset transitions. Dysphasic children performed significantly worse than normals on short tones, short vowels and short transition consonants at short ISIs, but not at long ISIs, nor on the corresponding long stimuli. From the similar results for short steady-state vowels and short transition consonants, and from improved performance on long transitions, the authors concluded that 'it is the brevity, not the transitional character . . . of synthesized stop consonants which results in the impaired perception of our dysphasic children' (Tallal & Piercy 1975: 73). The dysphasic children also performed equally well on TOJ and discrimination, from which the authors concluded that apparent deficits in auditory sequencing could be due to difficulty in discriminating and identifying stimuli rapidly rather than in temporal order judgment itself (Tallal & Piercy 1973: 396). Two later studies of reading-impaired children reached the same conclusion (Reed 1989; Tallal 1980).

Unfortunately, Tallal & Newcombe (1978) abandoned the conclusion that it was the brevity, not the transitional properties of stop-vowel syllable onsets that caused difficulty, and proposed, without any new evidence or argument, that the children of the earlier studies had suffered from 'a defect in temporal acoustic processing' (p. 22) and had had difficulty with speech sounds that incorporate rapidly changing acoustic spectra (p. 13). The changed interpretation evidently rested on two unexplained decisions: (1) to reinterpret improved performance with lengthened transitions as due to a decrease in rate of spectral change rather than to an increase in duration, and (2) to view a change in fundamental frequency across two discrete, rapidly successive complex tones as somehow equivalent to a rapid continuous sweep in timbre across a syllable onset (for fuller discussion, see Studdert-Kennedy & Mody 1995).

The equation of discrete tone sequences with continuous spectral sweeps cannot be justified either acoustically or perceptually. And the reason for improved discrimination (if any) with lengthened transitions is an experimental question: Is it duration, rate, or simply increased phonetic contrast? In fact, several studies designed to test the effect of lengthening transitions on /ba/-/da/ identification and/or discrimination have found no improvement in adult aphasics (Blumstein, et al. 1984; Riedel & Studdert-Kennedy 1985), dyslexic children (McAnally et al. 1997) or learning disabled children (Bradlow et al. 1999).

Here I should add that no experiment, since Tallal and Newcombe first formulated the rapid auditory processing hypothesis, has ever shown that the difficulty some children or adults may have in discriminating rapidly presented pairs of stop-vowel syllables is due to the inability of their auditory systems to analyze, integrate or otherwise perceive the rapid spectral changes on which the contrast depends. The reason is simply that, until the work of Mody et al. (1997), no experiment had ever run the controls necessary to determine whether the difficulty lies in how the auditory system analyzes the acoustic information, in how the language system uses it to form a phonetic representation, or simply in discriminating two phonetically similar syllables at rapid rates of presentation.

Similar strictures apply with equal force to the training studies reported by Merzenich et al. (1998). From this and other published reports (Merzenich et al. 1996; Tallal et al. 1996) it appears that no attempts were made: (1) to determine the precise nature of the children's speech perception deficits, if any, (2) to separate the effects of adaptive training in 'temporal processing' from the effects of intensive exposure to modified natural speech in computer games, and (3) to separate the effect of overall prolongation of speech by 50%

from the effect of differentially slowing and amplifying transitional elements (Tallal et al. 1996: 81).

Whatever improvement the children of these studies displayed may therefore have arisen from intensive exposure to natural speech, slowed not specifically in its formant transitions, but in the overall rate at which words were delivered. Slowed speech may be easier to perceive than normal speech not because the auditory system does not then have to contend with rapid acoustic changes, but because the reduced rate allows more time for the disabled language system to form phonetic representations. The question of which interpretation is correct can, and should, be decided by systematic controlled experiments (cf. Lacerda & Lindblom 1998; Studdert-Kennedy et al. 1995).

Hemispheric specialization for speech perception

Let me now turn briefly to Tallal's assumption that hemispheric specialization for speech perception is grounded in a prior specialization for 'rapid auditory temporal processing'. The assumption rests on the finding of Schwartz & Tallal (1980) that the right-ear advantage (REA) for stop-vowel syllables, dichotically presented to 30 normal adults, was significantly reduced from about 10% for syllables with short (40 msec) onset transitions to about 6% for syllables with long (80 msec) transitions. The authors interpreted this drop as evidence that the left hemisphere is specialized for processing rapid acoustic changes.

Two conditions are necessary, however, for an ear advantage in dichotic listening: (1) hemispheric specialization for the stimulus materials, and (2) fuller access to the specialized hemisphere from the contralateral than from the ipsilateral ear. Many studies have demonstrated that the REA for speech varies with factors that may affect access to the left hemisphere, such as signal/noise ratio, signal energy or duration, difficulty of the phonetic contrast, and so on (Studdert-Kennedy & Shankweiler 1981). Since the REA for the long transitions in the Schwartz & Tallal study (1980) was still appreciable, and perhaps significant, the drop may simply have reflected fuller access of the left ear signal to the language hemisphere, due to increased energy in the transitions, rather than a difference in hemispheric specialization.

I emphasize the ambiguity of this outcome for three reasons. First, two other studies, cited by Schwartz and Tallal in support of their interpretation, also had ambiguous outcomes: Efron (1963) and Swisher & Hirsh (1972) observed non-speech TOJ deficits in left-lesioned aphasics. Both studies explicitly left open, however, the question of whether the TOJ deficits were

a cause or a consequence of aphasia. Second, a later publication by Tallal and her colleagues adapted the data of Schwartz & Tallal (1980, Table 1) for presentation as a bar graph, but increased the total left ear score on long transitions for the entire group by about 400 responses, and thus eliminated the possibly significant (6%) right ear advantage for that condition (Tallal et al. 1993, Figure 7; reprinted 1995). Interested readers should therefore consult the table in the original paper, not its later misprinting.

My third reason for emphasizing the ambiguity is that a recent dichotic study has resolved it. Best & Avery (1999) tested native Zulu speakers and native English speakers on Zulu click consonants, which conform perfectly to the requirements of Tallal's hypothesis: they are brief (<50 msec), spectrally complex, and have rapid formant transitions. Yet they are heard as non-speech noises, extraneous to the flow of speech, by listeners unfamiliar with click languages. The authors report significant REAs in native Zulu speakers, but not in native English speakers, for both click-vowel syllables and isolated clicks excised from speech. Evidently, perception of both stop and click consonants depends on a left hemisphere mechanism specialized for handling linguistic information.

Independent tests of the auditory hypothesis

The first published attempt to test the rapid auditory processing hypothesis experimentally came from Mody et al. (1997) (see also Studdert-Kennedy, Mody & Brady 2000). Twenty second-grade reading-impaired subjects were selected for, among other things, their errors on /ba/-/da/ TOJ. In Experiment 1a they were tested on discrimination and TOJ of synthetic /ba/-/da/ at short interstimulus intervals (ISIs): Errors increased monotonically as ISI decreased on both discrimination and TOJ, with no significant difference between tasks. In Experiment 1b, the same procedures were followed with syllable pairs, /ba/-sa/and /da/-/∫a/. The poor readers now made essentially no errors on either task. Thus, they performed perfectly under time pressure when they could clearly discriminate and identify the paired syllables. Evidently, /ba/ and /da/ were difficult because they are very similar.

Is the similarity of /ba/ and /da/ auditory or phonetic? In Experiment 2, subjects discriminated non-speech control stimuli, consisting of two sine waves with durations and frequency trajectories identical to those of the center frequencies of F2 and F3 that carried the /ba/-/da/ contrast in Experiment 1, but not heard as speech. Contrary to the results for /ba/-/da/, non-speech performance was completely unaffected by decreases in ISI. From these results we conclude that the poor readers' difficulties with /ba/-/da/ were due neither to a deficit in TOJ, nor to a general deficit in rate of

auditory processing, nor to a deficit in processing brief patterns of rapidly changing acoustic information, but to difficulty in discriminating and iden- tifying phonetically similar (although phonologically contrastive) syllables under time pressure.

Since Mody et al. (1997) was published, at least four other papers have reported failures to confirm aspects of Tallal's hypothesis (Bradlow et al. 1999; Bishop et al. 1999; McAnally et al. 1997; Nittrouer 1999). Thus, not one of five independent studies has found evidence in children with either specific language impairment or specific reading impairment for the auditory deficit that Tallal and her colleagues propose as the basis of these children's disorders.

Conclusions

1. No experimental study has yet supported either of the main assumptions of the rapid auditory processing hypothesis. We have no evidence that the left cerebral hemisphere is specialized for rapid auditory processing, and difficulties in discriminating /ba/-da/ at rapid rates of presentation, when observed, are phonetic not auditory in origin.
2. No experimental study has yet demonstrated a mechanism by which any known auditory deficit in the reading impaired might disrupt either speech perception or phoneme awareness. Repeatedly observed corre- lations between performances on certain non-speech auditory tasks and various linguistic tasks may reflect the dependence of both types of task on general cognitive capacities, such as intelligence and attentiveness, or on other common factors.
3. Future tests of the general auditory hypothesis should abandon the corre- lation coefficient, a notoriously unreliable measure for small samples, and often difficult to interpret. Such tests should adopt instead the rigorous analytic techniques of experimental psychology: systematic manipulation of acoustic variables in speech and in matched non-speech controls.

Postscript: An alternative hypothesis

If segmentation of words into their phonological components is an emer- gent consequence of lexical growth, as several authors have proposed (e.g., Lindblom 1992; Lindblom et al. 1984; Studdert-Kennedy 1987, 1989) we may hypothesize that a smaller than usual lexicon will result in defective ('fuzzy'/'weak') phonological representations, and so defective phoneme awareness. Recently, Nittrouer (1996) reported delayed development both of

aspects of normal speech perception and of phoneme awareness in children whose exposure to spoken language had been less than the middle class norm due either to low socioeconomic status or to a history of otitis media. If the deficits of the language-impaired children of Tallal, Merzenich and their colleagues were likewise due to inadequate exposure to spoken language, any apparent improvement in their language skills, after training with computer-modified speech, might simply have followed from their sudden, intensive immersion in the speech they had missed, now slowed to a cognitively manageable rate. An experimental test of such a hypothesis would be an appropriate next step in the Tallal-Merzenich research program.

Acknowledgments

References

Benasich, A. (1998). Temporal integration as an early predictor of speech and language development. In: C. von Euler, I. Lundberg, I. & R. Llinas (eds.), *Basic mechanisms in cognition and language* (pp. 123–142). New York: Elsevier.

Benasich, A. & Tallal, P. (1996). Auditory temporal processing thresholds, habituation, and recognition memory over the 1st year, *Infant Behavior and Development* 19: 339–357.

Best, C.T. & Avery, R.A. (1999). Left hemisphere advantage for click consonants is determined by linguistic significance and experience, *Psychological Science* 10: 65–70.

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., Tartter, V.C., Nigro, G. & Statlender, S. (1984). Acoustic cues for the perception of place of articulation in aphasia, *Brain and Language* 22: 128–149.

Bradlow, A.R., Kraus, N., Nicol, T.G., McGee, T.J., Cunningham, J., Zecker, S.G. & Carrell, T.D. (1999). Effects of lengthened formant transition duration on discrimination and neural representation of synthetic CV syllables by normal and learning-disabled children, *Journal of the Acoustical Society of America* 106: 2086–2096.

Brady, S.A. & Shankweiler, D.P. (eds.) (1991). *Phonological processes in literacy*. Hillsdale, NJ: Erlbaum.

Efron, R. (1963). Temporal perception, aphasia and déjà-vu, *Brain* 86: 403–424.

Hari, R. & Kiesila, P. (1996). Deficit of temporal auditory processing in dyslexic adults, *Neuroscience Letters* 205: 138–140.

Kinsbourne, M., Rufo, D.T., Gamzu, E., Palmer, R.L. & Berliner, A.K. (1991). Neuropsychological deficits in adults with dyslexia, *Developmental Medicine and Child Neurology* 3: 763–775.

Lacerda, F. & Lindblom, B. (1998). Some remarks on Tallal's transform in the light of emergent phonology. In: C. von Euler, I. Lundberg & R. Llinas (eds.), *Basic mechanisms in cognition and language* (pp. 197–222). New York: Elsevier.

Liberman, A.M. (1998). The biology of speech: Two theories. In: C. von Euler, I. Lundberg, & R. Llinas (eds.), *Basic mechanisms in cognition and language* (pp. 223–229). New York: Elsevier.

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. (1973). Segmentation of the spoken word, *Bulletin of the Orton Society* 23: 65–77.

Lindblom, B. (1992). Phonological units as adaptive emergents of lexical development. In: C.A. Ferguson, L. Menn & E. Stoel-Gammon (eds.), *Phonological development* (pp. 131–163). Timonium, MD: York Press.

Lindblom, B., MacNeilage, P. & Studdert-Kennedy, M. (1984). Self-organizing processes and the explanation of phonological universals. In: B. Butterworth, B. Comrie & O. Dahl (eds.), *Explanations of language universals* (pp. 181–203). New York: Mouton.

Locke, J. (1998). Early developmental delay: Problems for the theorist. In: C. von Euler, I. Lundberg & R. Llinas (eds.), *Basic mechanisms in cognition and language* (pp. 231–242). New York: Elsevier.

McAnally, K.I., Hansen, P.C., Cornelissen, P.L. & Stein, J.F. (1997). Effect of time and frequency manipulation on syllable perception in developmental dyslexics, *Journal of Speech, Language, and Hearing Research* 40: 912–924.

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.

Merzenich, M.M., Miller, S., Jenkins, W.M., Saunders, G., Protopapas, A., Peterson, B. & Tallal, P. (1998). Amelioration of the acoustic and speech reception deficits underlying language-based learning impairments. In: C. von Euler, I. Lundberg & R. Llinas (eds.), *Basic mechanisms in cognition and language* (pp. 143–172). New York: Elsevier.

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.

Nicholson, R.I. & Fawcett, A.J. (1994). Reaction times and dyslexia, *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 47A: 29–48.

Nittrouer, S. (1996). The relation between speech perception and phoneme awareness: Evidence from low-SES 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.

Reed, M.A. (1989). Speech perception and the discrimination of brief auditory cues in reading disabled children, *Journal of Experimental Child Psychology* 48: 270–292.

Riedel, K. & Studdert-Kennedy, M. (1985). Extending formant transitions may not improve aphasics' perception of stop consonant place of articulation, *Brain and Language* 24: 223–232.

Schwartz, J. & Tallal, P. (1980). Rate of acoustic change may underlie hemispheric specialization for speech perception, *Science* 207: 1380–1381.

Stein, J.F. & McAnally, K. (1995). Auditory temporal processing in developmental dyslexics, *Irish Journal of Psychology* 16: 220–228.

Studdert-Kennedy, M. (1987). The phoneme as a perceptuomotor structure. In: A. Allport, D. MacKay, W. Prinz & E. Scheerer (eds.), *Language perception and production* (pp. 67–84). New York: Academic Press.

Studdert-Kennedy, M. (1989). The early development of phonology. In: C. von Euler, H. Forssberg & H. Lagercrantz (eds.), *Neurobiology of early infant behavior* (pp. 287–301). New York: Stockton.

Studdert-Kennedy, M. & Mody, M. (1995). Auditory temporal perception deficits in the reading-impaired: A critical review of the evidence, *Psychonomic Bulletin and Review* 2: 508–514.

Studdert-Kennedy, M. & Shankweiler, D. (1981). Hemispheric specialization for language processes, *Science* 211: 960–961.

Studdert-Kennedy, M., Mody, M. & Brady, S. (2000). Speech perception deficits in poor readers: A reply to Denenberg's critique, *Journal of Learning Disabilities* 33(4).

Studdert-Kennedy, M., Liberman, A.M., Brady, S.A., Fowler, A.E., Mody, M. & Shankweiler, D.P. (1995). Lengthened formant transitions are irrelevant to the improvement of speech and language impairments, *Haskins Laboratories Status Report on Speech Research* SR 119/120: 35–38.

Swisher, L. & Hirsh, I.J. (1972). Brain damage and the ordering of two temporally successive stimuli, *Neuropsychologia* 10: 137–152.

Tallal, P. (1980). Auditory temporal perception, phonics and reading disabilities in children, *Brain and Language* 9: 182–198.

Tallal, P. & Newcombe, F. (1978). Impairment of auditory perception and language comprehension in dysphasia, *Brain and Language* 5: 13–24.

Tallal, P. & Piercy, M. (1973). Developmental aphasia: Impaired rate of nonverbal processing as a function of sensory modality, *Neuropsychologia* 11: 389–398.

Tallal, P. & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception, *Neuropsychologia* 12: 83–93.

Tallal, P. & Piercy, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants, *Neuropsychologia* 13: 69–74.

Tallal, P., Miller, S. & Fitch, R.H. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. In: P. Tallal, A.M. Galaburda, R.R. Llinas & C. von Euler (eds.), *Temporal information processing in the nervous system. Annals of the New York Academy of Sciences*, Vol. 82 (pp. 27–47). New York: New York Academy of Sciences. Paper reprinted in *Irish Journal of Psychology* (1995), 16: 194–219.

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

Address for correspondence: Dr M. Studdert-Kennedy, Haskins Laboratories, 270 Crown Street, New Haven CT 06511, USA
Phone: (203) 865-6163, ext. 278; Fax: (203) 865-8963; E-mail: msk@haskins.yale.edu