Speech Perception Deficits in Poor Readers:

A Reply to Denenberg's Critique

Michael Studdert-Kennedy, Maria Mody, and Susan Brady

Abstract

We reply to Denenberg's (1999) recent critique of our work (Mody, Studdert-Kennedy, & Brady, 1997). Denenberg mounted two main lines of criticism, one concerning characteristics of the population sampled for the experimental group, and the other a statistical critique, concerning (a) violation of parametric assumptions for use of the *F* distribution and (b) our supposed acceptance of the null hypothesis of no differences between experimental and control groups. We show that the first criticism stemmed from a misunderstanding of the experimental hypothesis and that the second can be answered by both parametric and nonparametric comparisons across conditions within the experimental group, without reference to the control group. Thus, our original conclusion stands: The difficulty with rapid /ba/–/da/ discrimination that some children with reading impairment may experience does not stem from difficulty in discriminating the rapid spectral transitions at stop-vowel syllable onsets.

We have critically reviewed all the relevant work by Tallal, either in the introduction to the article under discussion or elsewhere (Studdert-Kennedy & Mody, 1995). Here it should suffice to note the immediate antecedents to our study. In her only study of children with specific reading impairment, Tallal (1980) extended to this population use of the rapid tone discrimination procedure she had previously used on children with specific language impairment (Tallal & Piercy, 1973, 1974, 1975). She reported that only 9 (45%) of her 20 participants with reading impairment made more errors than the worst of 12 nondisabled controls, but she also reported a highly significant rank-order correlation coefficient (rho = .81, $p < .001$) for the whole group, between rapid tone discrimination and nonsense word reading.

In an attempt to account for this correlation, Tallal speculated that children with reading impairment might suffer from the same combination of deficits as the dysphasic children of her earlier studies (Tallal & Piercy, 1973, 1974, 1975), namely, difficulty in discriminating between both rapidly presented tones and rapidly presented stop-vowel syllables (/ba/–/da/); the latter difficulty in speech perception would then underlie the children's phonological deficit in reading. Tallal further speculated that both speech deficits and nonspeech deficits would reflect impairment in the same "basic perceptual mechanism" (p. 196) for analyzing rapidly changing acoustic information, such as the discrete pitch change in brief, rapidly presented tone pairs and the spectral sweep in the onset transitions of stop-vowel syllables.

The predicted combination of speech and nonspeech deficits in some children with specific reading impairment was indeed confirmed by Reed (1989), the only researcher to have tested such children with the Tallal paradigm on both tones and stop-vowel syllables. Reed did not confirm, however, that these two deficits arose from failure of the same perceptual mechanism, nor has any subsequent experiment tested the hypothesis that children with reading impairment who have difficulty in discriminating /ba/ from /da/ at rapid presentation rates also have difficulty in discriminating rapidly presented nonspeech spectral sweeps corresponding to stop-vowel syllable onset transitions. That was the main purpose of our study.

We turn now to Denenberg's (1999) critique. We omit comment on many minor errors, such as his notion that the lack of significant group differences on sine wave identification and discrimination training constituted an "experimental failure" (p. 381, middle column). In fact, of course, this outcome was exactly what we had expected. But because the training portions of the study were entirely secondary and had little bearing on our overall conclusions, we concentrate on Denenberg's two main lines of criticism.

Participants and Reading Impairment

The main purpose of our study was not, as Denenberg implies, "to understand students with reading impair-

ment" (p. 383). Our intent, explicitly stated in the article (Mody et al., 1997, p. 207), was to test Tallal's hypothesis that the difficulties some poor readers may experience in /ba/–/da/ temporal order judgment (TOJ) and discrimination at rapid presentation rates stem from difficulties in discriminating the rapid acoustic changes at syllable onsets. For this, we obviously required a sample of poor readers who had difficulty with /ba/–/da/ TOJ at short interstimulus intervals (ISIs) and a control sample of good readers who did not (Experiment 1a). We could then test the poor readers in two ways. First, we could ask whether their difficulties with TOJ at short ISIs persisted for a more easily discriminable pair of syllables (i.e., stop-vowel vs. fricative-vowel); if the difficulties disappeared, we could conclude that they had been due not to TOJ itself but to being required to discriminate acoustically and phonetically similar syllables at short ISIs (Experiment 1b). Second, we could ask whether the children's difficulties with /ba/–/da/ discrimination persisted when the stimuli were nonspeech control patterns acoustically matched to the onset transitions of /ba/ and /da/; if the difficulties disappeared, we could conclude that they were phonetic—that is, specific to speech—and were not due to the general auditory deficit hypothesized by Tallal (Experiment 2).

For control purposes we also needed an appropriately matched sample of children without reading impairment and without difficulties in rapid /ba/–/da/ TOJ. The main function of the control group was to provide a normal baseline for performance on the nonspeech task, on which, due to the unfamiliar quality of the sounds, errors were likely to be significantly higher for both groups than on speech. As we shall see, all the main conclusions of the article can be reached by within-group comparisons across conditions on the poor readers without reference to the control group.

In any event, as we explicitly stated, the requirement that the poor readers should have difficulty with /ba/–/da/ TOJ and that the good readers should not,

> precluded either good readers or poor readers from being fully representative samples of the good and poor reading populations. Rather, they were samples from experimentally defined populations of good readers who make no errors on Tallal's task and of poor readers who make at least 8% errors on that task. Because the main goal of the study was to test a hypothesis concerning the cause of errors on the task, these were appropriate target populations. (Mody et al., 1997, p. 211)

We stand by this statement and believe that our conclusions can be fairly extrapolated to those populations.

Finding 20 poor readers and 20 good readers, matched for age and intelligence, who also met the /ba/–/da/ task performance criteria proved surprisingly difficult (see Mody et al., 1997, footnote 4, p. 209, and footnote 5, p. 210). Among other constraints, we were obliged to settle for poor readers who were no more than 5 months delayed, on average, compared with the full-year delay of Tallal's (1980) group (of which, we should recall, 55% performed normally on the tone task). Denenberg argues that because our poor readers "did not meet the criterion specified by Stark and Tallal (1988) for classification as 'reading-impaired' " (p. 383), we can draw no inferences to a population that includes children with more severe impairments.

> Yet, as we remarked in our article and here repeat, whether the difference in reading level is a serious concern depends . . . on how likely it is that identical difficulties with /ba/–/da/ TOJ stem from different perceptual deficits in more severely impaired than in less severely impaired readers. We do not find this likely. (p. 223)

As the immediately following sentences should have made clear to an attentive reader, we did not find it likely because

> several large-scale studies converge on the conclusion that reading ability is normally distributed with no qualitative difference between those who are simply less skilled and those who meet standard criteria as reading-disabled (e.g., Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992; Stanovich & Siegel, 1994). If this is so, the results of the present experiments can be generalized to specifically reading-impaired children who have difficulty with /ba/–/da/ TOJ, regardless of their degree of reading impairment. (Mody et al., 1997, p. 223)

Here, Denenberg finds that our "reasoning is questionable" (p. 382). But he does not, in fact, question our reasoning. He does not question the evidence that reading ability is normally distributed; nor does he question the propriety, in light of that evidence, of generalizing our conclusions to other children who have difficulty with /ba/–/da/ TOJ. All that Denenberg offers by way of criticism are two derogatory epithets and a portentous dictum. First, he finds our placing of poor readers on a continuum with more severely impaired readers to be "highly dubious" (p. 382. He does not explain why he finds it dubious.). Second, Denenberg finds that the use of ba/–/da/TOJ as the "ultimate criterion for defining a reading-impaired population is weak." (We entirely agree. This criterion was proposed by Tallal [1980], not by us. Subsequently, Tallal, Stark, and Mellits [1985] proposed the criterion not as "ultimate," to be sure, but as one of six variables by which "language impairment can be accurately identified" [Tallal, 1999, p. 222].) Finally, we read: "Knowledge in a field is not advanced by drawing subjective conclusions, such as 'We do not find this likely' " (p. 382). (Nor, we might add, is knowledge advanced by enouncing moral platitudes. Evidently, Denenberg did not understand that the studies cited in the next sentence, quoted above, supplied objective evidence to support our "subjective conclusions.")

In short, Denenberg objected to our participants because he failed to grasp

that we were engaged not in a normative study of children with reading impairment, but merely in testing Tallal's hypothesis concerning the perceptual mechanism underlying the errors of some such children on /ba/-/da/ TOJ and discrimination under time pressure.

Statistical Methods and Inferences

Denenberg raises two main objections to our statistical methods and inferences: first, that we repeatedly accepted the null hypothesis as proved rather than as not rejected, and, second, that we used analysis of variance on data sets for which the required population assumptions of normality and homogeneity of variance were not met.

We begin with two preliminary points. First, we agree that in several places, correctly noted by Denenberg (e.g., p. 380, 381), we slipped into phrasing that seems to take the lack of evidence for a difference between groups as evidence for the lack of a difference. We were, however, meticulous in framing our conclusions, as Denenberg himself acknowledges (p. 383): We summarized the outcome of our study not as disproving Tallal's hypothesis but as failing to support it where support would be expected (Mody et al., 1997, p. 224).

Our second and more substantive point is that, as remarked above, the main results of the study (Experiments 1a, 1b, and 2) stand unchanged by analyses that disregard the control group and rest their conclusions on significant differences across conditions for the poor readers alone.

Experiments 1a and 1b

In Experiment 1a the poor readers displayed a monotonic increase in errors as ISI decreased on /ba/-/da/ TOJ and discrimination. That is, for these acoustically and phonetically similar speech stimuli, when the interval between them was shortened, the poor readers made progressively more errors. By contrast, when each of these stimuli (/ba/, /da/) was paired with an item that was acoustically and phonetically very different, that is, /ba/-/sa/ and /da/-/ʃa/ (not /da/-/fa/, as Denenberg has it), the poor readers made essentially no errors at all: One TOJ error and two discrimination errors on 240 trials at the shortest ISI for /da/-/ʃa/ (Experiment 1b). The mean error curves as a function of ISI were, therefore, flat at, or very close to, zero—strikingly different from the falling functions of Experiment 1a (see Figure 1 of Mody et al., 1997, and, for /ba/-/da/ discrimination, see Figure 1, below). If we are skeptical, we can test the significance of the difference between stimulus conditions at each ISI for both tasks by a series of simple sign tests. The results are the same for both tasks: no significant difference between conditions at 100 ms ISI, and significantly more errors on /ba/-/da/ at both 50 ms ($p < .05$) and 10 ms ($p < .01$).

Far, then, from there being "no meaningful interpretation of the lack of findings in Experiment 1b," as Denenberg states (p. 381), the lack of an effect of ISI for poor readers in Experiment 1b contrasted with the presence of a strong effect of ISI for the same participants in Experiment 1a, and served to establish our conclusions as

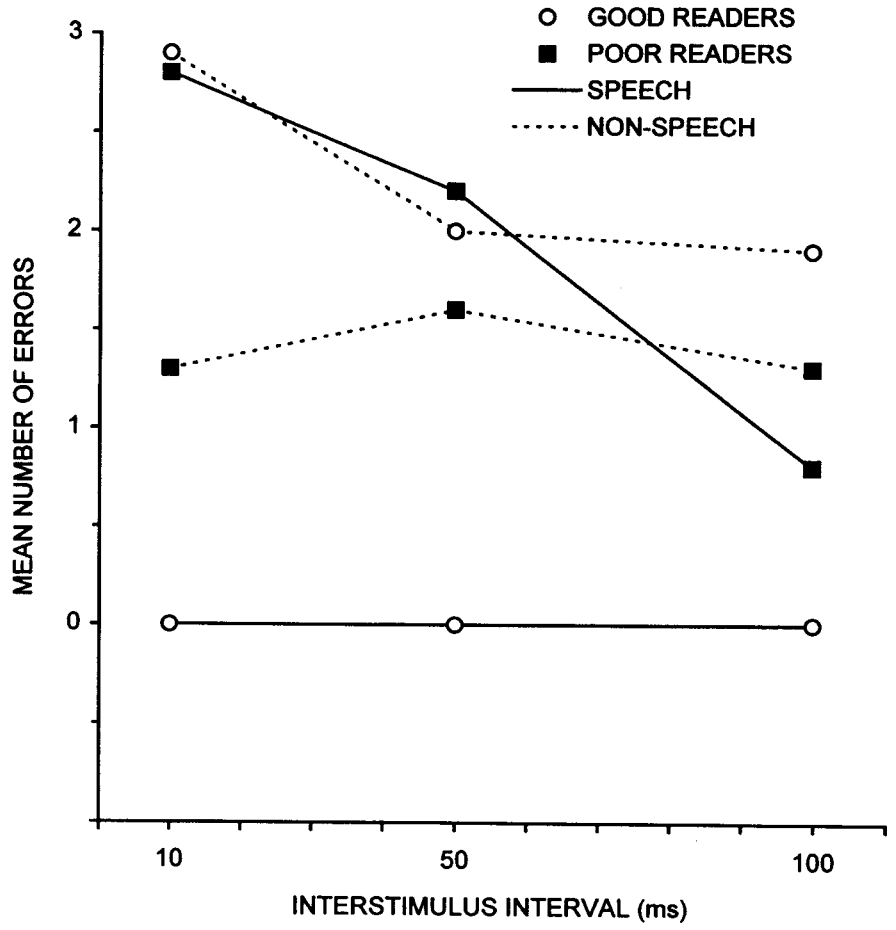

FIGURE 1. Mean number of errors by good and poor readers as a function of ISI on speech (/ba/-/da/) and nonspeech discrimination. (Reprinted from Mody et al., 1997.)

originally stated: "These findings demonstrate that the poor readers' difficulties with /ba/–/da/ TOJ do not reflect a general problem with temporal order analysis: Poor readers judge temporal order accurately, even at rapid rates of presentation, if they can identify the items to be ordered" (Mody et al., 1997, pp. 215–216).

Experiments 1a and 2: Poor Readers

Figure 1 (reprinted from Figure 2 of Mody et al., 1997) plots discrimination errors of good and poor readers as a function of ISI for /ba/–/da/ (Experiment 1a) and for the nonspeech sine wave control stimuli (Experiment 2). Of interest here are the falling curve of the poor readers as a function of ISI for speech and their more or less flat curve for nonspeech. We reported the appropriate two-way ANOVA across speech and nonspeech conditions for poor readers alone (Mody et al., 1997, p. 218), but the test seems to have escaped Denenberg's notice. The results showed no effect of condition on the overall error rate, but a significant main effect of ISI, due to its strong effect on speech, as shown by a significant Condition × ISI interaction. From this and the previously reported three-way analysis across groups, we concluded that "the poor reader's difficulties with /ba/-/da/ were specific to speech, and cannot be attributed to a general auditory deficit in the perception of brief patterns of rapidly changing acoustic information" (Mody et al., 1997, pp. 218–219). We stand by this conclusion.

We were, of course, well aware that our data might not meet the standard assumptions for use of the *F* distribution, a normal population distribution, and homogeneous population variances. We therefore relied on the well-known sampling experiments of D. W. Norton, reported at some length by Lindquist (1953). The general conclusions, in Lindquist's words, were that (a) "the *F*-distribution seems so insensitive to the form of the [parent] distribution . . . that it hardly seems worthwhile to apply any statistical test to the data to detect non-normality"; (b) "unless the heterogeneity of form or variance is so extreme as to be readily apparent upon inspection of the data, the effect upon the *F*-distribution will probably be negligible" (p. 86).

Denenberg is apparently less liberal in these matters. We have therefore computed Kendall's Coefficient of Concordance (*W*; Ferguson, 1981) for the speech and nonspeech ISI functions of the poor readers. To do this, we ranked the errors across ISIs for each of the 20 participants on speech and nonspeech and then estimated the effect of ISI under each of the two conditions by computing *W*. (Readers interested in checking the raw data for these computations should refer to Tables D and G in the Appendix of Mody [1993].) *W* is a nonparametric equivalent of the correlation ratio (eta squared), an estimate of the proportion of the total variance of ranks that can be attributed to variations in ISI; its values range from 0 to 1 and can be tested against the *F* distribution in a procedure analogous to that for the squared Pearson product-moment correlation coefficient. The results were as follows: For speech, $W = .33$, $F(2, 36) = 9.36$, $p < .01$; for nonspeech, $W = .02$, $F(2, 36) = 0.47$, $p > .50$. Thus, for the poor readers, ISI had a highly significant effect on speech, but no effect on the nonspeech control. These results confirm our original conclusions, quoted earlier.

Experiment 2: Good Readers

Denenberg is concerned with the Group × ISI interaction for nonspeech discrimination, evident in Figure 1. An ANOVA found no significant difference between groups but a significant Group × ISI interaction, obviously due to the sharp increase in errors at the shortest ISI for good readers. A post hoc Scheffé test of the increase proved not to be significant, however, and we concluded that there was no evidence for an effect of ISI on nonspeech discrimination for either group. Denenberg chides us for this, accusing us of using "a contrived analysis" (p. 382) and "a statistical stratagem to justify ignoring . . . [the] interaction" (p. 383). In fact, our post hoc test was merely a standard procedure in statistical analysis for locating and estimating the significance of an interaction.

Denenberg's concern with this issue arises from his mistaken belief that the interaction "runs contrary to . . . [our] theoretical expectations" (p. 382). Once again, he reveals his complete failure to understand the background and rationale of the experiments. Our theoretical expectations were based on the hypothesis that the speech perception difficulties of some poor readers arise from language-related deficits in phonological representation, not from the general auditory deficit posited by Tallal. We therefore expected that poor readers who had difficulty in discriminating /ba/–/da/ at short ISIs would have no more difficulty than normal controls in discriminating matched nonspeech acoustic control patterns at the same short ISIs. This is exactly what we found. Similarly, because we do not believe that auditory discriminative capacity in normal-hearing participants has any bearing on reading capacity, we had no expectations regarding good readers' discrimination of nonspeech control patterns. We expected auditory and phonetic discriminative capacity to vary independently.

Tallal, by contrast, believes that good and poor readers are distinguished by their ability to process rapidly changing acoustic stimuli: Good readers can handle them, poor readers cannot. Poor readers should therefore make more errors than good readers and should be more affected by brief ISIs. Yet, what we observed was precisely the reverse: Good readers made more errors across the board on the nonspeech task than poor readers (although not significantly) and, if we take the significant interaction seriously, were more affected by ISI. Ironically, computation of Kendall's Coefficient of Concordance to estimate the effect of ISI on discrimination of the

nonspeech control patterns by good readers gives a significant value of $W = .15$, $F(2, 36) = 3.45$, $.01 < p < .05$. Perhaps, then, we were wrong to accept Scheffé's conservative verdict on the interaction. If so, good readers were indeed more affected by ISI than poor readers, and we have here (contrary to Tallal's hypothesis and to Denenberg's understanding) further evidence that lack of skill in processing rapid frequency sweeps at short ISIs does not predict reading capacity.

Conclusions

We have tried not to bore the reader by taking up every point of misunderstanding in Denenberg's critique. We hope, however, that we have taken up enough to show that his attempt to undermine the article is without merit. We are encouraged by the fact that, since we wrote, at least five other published papers have tested aspects of Tallal's hypothesis and found it wanting (Best & Avery, 1999; Bishop, Carlyon, Deeks, & Bishop, 1999; Bradlow et al., 1999; McAnally, Hansen, Cornelissen, & Stein, 1997; Nittrouer, 1999). Thus, not one of six independent studies has found evidence in children with either specific language impairment or specific reading impairment for the auditory deficit that Tallal proposes as the basis of their disorder.

ABOUT THE AUTHORS

Michael Studdert-Kennedy, *PhD, is former president of Haskins Laboratories, New Haven, and professor emeritus of psychology at the University of Connecticut.* ***Maria Mody,*** *PhD, is assistant professor of otolaryngology at the Albert Einstein College of Medicine, New York.* ***Susan Brady,*** *PhD, is professor of psychology at the University of Rhode Island and a senior research associate at Haskins Laboratories, New Haven. Address: Michael Studdert-Kennedy, Haskins Laboratories, 270 Crown St., New Haven, CT 06511-6695 (e-mail: msk@haskins.yale.edu).*

AUTHORS' NOTES

1. *We thank Len Katz for commenting on this article and approving our statistical argument. We also thank Victor Denenberg for affording us the occasion to correct misunderstandings of the study (particularly of its experimental hypothesis and its target population) that we have heard from a number of clinicians.*
2. *Preparation of this article was supported in part by National Institute of Child Health and Human Development Grant No. HD-01994 to Haskins Laboratories.*

REFERENCES

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.

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.

Denenberg, V. H. (1999). A critique of Mody, Studdert-Kennedy, and Brady's "Speech Perception Deficits in Poor Readers: Auditory Processing or Phonological Coding?" *Journal of Learning Disabilities, 32,* 379–383.

Ferguson, G. A. (1981). *Statistical analysis in psychology and education* (5th ed.). New York: McGraw Hill.

Lindquist, E. F. (1953). *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin.

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.

Mody, M. (1993). *Bases of reading impairment in speech perception: A deficit in rate of auditory processing or in phonological coding?* Unpublished doctoral dissertation, City University of New York.

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.

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.

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of the normal distribution of reading ability. *New England Journal of Medicine, 326,* 145–150.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86,* 24–53.

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

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

Tallal, P. (1999). Children with language impairment can be accurately identified using temporal processing measures: A response to Zhang and Tomblin. *Brain and Language, 69,* 222–229.

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., Stark, R. L., & Mellits, E. (1985). Identification of language-impaired children on the basis of rapid perception and production skills. *Brain and Language, 25,* 314–322.