Phonological instability in young adult poor readers

James S. Magnuson (james.magnuson@uconn.edu)[1,2]
Anuenue Kukona (anuenue.kukona@uconn.edu) [1,2]
David Braze (braze@haskins.yale.edu) [1]
Clinton L. Johns (johns@haskins.yale.edu) [1]
Julie A. Van Dyke (jvandyke@haskins.yale.edu) [1]
Whitney Tabor (whitney.tabor@uconn.edu) [1,2]
W. Einar Mencl (einar@haskins.yale.edu) [1]
Kenneth R. Pugh (pugh@haskins.yale.edu) [1,2]
Donald P. Shankweiler (donald.shankweiler@uconn.edu) [1,2]
[1]Haskins Laboratories, 300 George St., New Haven, CT 06510
[2]Department of Psychology, University of Connecticut, Storrs, CT 06269-1020

Abstract

Phonology is held to play a central role in typical reading development (Shankweiler et al., 1979) and sensory or phonological deficits are often held to be a primary cause of reading disability (Snowling, 2008). However, little is known about the nature of phonology at the endpoint of atypical reading development -- that is, in adult poor readers. We examined the time course of (auditory) lexical activation, competition, and learning in a community sample with a high proportion of poor readers in two experiments. In Experiment 1, contrary to our expectations, we found that poor readers were *more* sensitive to subphonemic coarticulatory cues than better readers. In Experiment 2, we examined the time course of word learning along with the time course of phonological competition. Poor readers differed from better readers in the trajectory of learning, and also in phonological competition: typical readers exhibited strong competition between rhymes, but poor readers did not. Simulations with a computational model suggest that instability in phonological organization (simulated via reduced lateral inhibition) can explain differences in both studies in counter-intuitive ways, shedding new light on an old problem.

Keywords: phonology; reading; dyslexia; reading disability; spoken word recognition; computational modeling; visual world paradigm.

Introduction

A fundamental principle shared by nearly all theories of reading is that phonology plays a key role mediating the mapping from print to meaning (Harm & Seidenberg, 2004; Shankweiler et al., 1979; Snowling & Hulme, 2005; Ziegler & Goswami, 2005). This follows from repeated findings that impairments in reading are correlated with deficits in phonological abilities (Shankweiler et al., 1977; Snowling, 1981). While multiple hypotheses exist, linking the deficit to poor phonological quality (Joanisse, 1994) or low-level sensory impairments (e.g., Tallal, 1980), the precise nature of the phonological deficit in dyslexia and its causes remains a subject of intense debate.

Fairly little is known about the nature of phonological processing at the endpoint of atypical reading development, since studies of reading disability logically focus on developing samples. An exception is recent work by Szenkovits, Ramus, and colleagues (reviewed by Ramus & Szenkovits, 2008). They point out that deficits in phonological abilities in college-aged poor readers (self-reported "presumed dyslexics") are most readily detected in tasks with significant working memory demands (phonemic awareness tasks, or verbal short-term memory tasks) or under time pressure (as in rapid auditory naming). However, in tasks that do not impose such demands, poor readers are not strikingly different from typical readers (most notably, they report that poor readers in their sample exhibit phonological similarity effects similar to those exhibited by good readers, contra Shankweiler et al., 1977, who reported that poor readers fail to show such effects). Ramus and Sjenkovits suggest that the phonological deficit in dyslexia therefore may not be one of phonological representation, but rather one of phonological *access* -- and so manifests as difficulty in rapidly retrieving phonological forms into working memory. This new take on phonology in dyslexia has the potential to illuminate the nature and basis of the phonological deficit in new ways.

Techniques for examining the time course of on-line language processing provide the means to examine this hypothesis more closely. We report preliminary results of a project investigating the phonological abilities of adult poor readers. We use stimulus manipulations and time course measures that have been used to investigate lexical activation and competition at a fine timescale (Dahan, Magnuson, Tanenhaus, & Hogan, 2001) and lexical learning (Magnuson, Tanenhaus, Aslin, & Dahan, 2003) in typical adults.

Experiment 1

In Experiment 1, we sought a sensitive test of the fine-grained phonological processing of our sample, but in a task that minimizes cognitive demands. The study reported by Dahan, Magnuson, Tanenhaus and Hogan (2001) fits the bill. Dahan et al. investigated the impact of misleading coarticulation (subcategorical -- i.e., subphonemic -- mismatches). They achieved misleading coarticulation by cross-splicing recordings of words. For example, they took the initial consonant and vowel (CV) from "neck", cut as late as possible before the final stop consonant, and spliced it together with the final consonant of "net". This sounds like "net", but the vowel includes coarticulation consistent with /k/. They labeled this sort of item "W2W1" (word 2 spliced to word 1). They also had cases where the initial CV

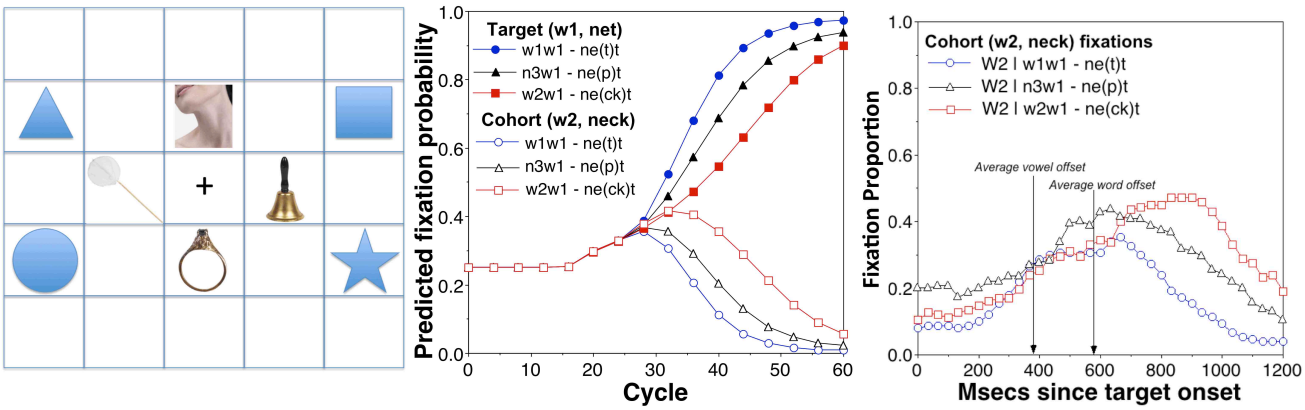

Figure 1. Left: Sample display. Center: TRACE predictions. Right: Competitor fixations over time from Dahan et al. (2001).

came from a nonword ("nep" + "net" → N3W1). Finally, they included cross-spliced items without misleading coarticulation by splicing together two recordings of a target word like "net" (W1W1).

Dahan et al. presented these items with displays like the one shown in Figure 1, using the Visual World Paradigm (VWP; Tanenhaus et al., 1995). Subjects heard instructions like "point to the net". Eye movements were recorded as subjects followed the spoken instructions.

The motivation for their study was the apparent deficiency in the TRACE model (McClelland & Elman, 1986) identified by Marslen-Wilson and Warren (1994) using these kinds of materials, in that lexical decision reaction times appeared inconsistent with the time course of activation in TRACE. However, the time course measure provided by the VWP (Figure 1, right) showed that the TRACE predictions (Figure 1, center) were remarkably accurate. Crucially, subjects fixated the competitor, "neck," most when there was misleading coarticulation consistent with that word (W2W1 condition), and least when the coarticulation was fully consistent with the target (W1W1). Fixation proportions were intermediate when misleading coarticulation did not map onto a word (N3W1). TRACE predicts the W1W1 and W2W1 patterns intuitively; the word with best bottom-up match is initially activated most strongly. The N3W1 results follow because neither net nor neck has an advantage as the nonword coarticulation is heard; thus, both reach a relatively high level of activation before the disambiguating final consonant.

Predictions What might we predict for our sample? If their linguistic difficulties arise from imprecise phonological representations (e.g., the phonological quality hypothesis of Joanisse, 2004) or slow-to-activate phonological representations (e.g., the generalized slowing hypothesis; Kail, 1994), we might expect them to be less affected by misleading coarticulation, and so show weaker competition effects. On the phonological access hypothesis (Ramus & Szenkovits, 2008), if the task minimizes cognitive demands, our sample ought to look no different from a typical sample.

Methods

Participants The participants were 56 college-aged adults (mean age = 21) recruited from community colleges and GED programs in the New Haven area. Previously, we have documented linguistic and other cognitive abilities in samples from this population (Braze et al., 2007), and demonstrated that the degree to which reading is subserved by common, supramodal brain areas also subserving speech is correlated with reading ability (Shankweiler et al., 2008). We examine this sample with a battery of 25 linguistic and other cognitive assessments. In this brief report, we only have room to mention that this population tends to lag in language and other cognitive domains, but a wide range of abilities is observed. Our goal is to conduct individual differences analyses. Given space constraints for the current report, though, we will compare the top 50% of readers in our sample with the bottom 50%. The most intuitive measure for conducting this median split is the standardized score from the Peabody Picture Vocabulary Test (which correlates closely with, e.g., a composite score derived from all subtests of the Woodcock-Johnson battery). The bottom 50% had standard scores ranging from 67 to 90, with a mean of 81. The top 50% had scores ranging from 91 to 137, with a mean of 104. The results we report do not differ if we remove, e.g., participants with low approximated IQ, and so the full sample is included.

Materials The auditory materials were those used by Dahan et al. (2001), and consisted of 15 word 1-word 2-nonword 3 triples (W1, W2, N3), such as *net, neck,* and *nep* (for the full set, see the Appendix B of Dahan et al.). The visual materials were similar to those used by Dahan et al., except that their line drawings were replaced with photographs.

Procedure The procedure was identical Dahan et al.'s. There were 3 lists, with 5 items assigned to each condition (W1W1 [consistent coarticulation], W2W1 [misleading cohort coarticulation], N3W1 [misleading nonword coarticulation]) in each list. Participants were randomly assigned to lists. On each trial, a fixation cross and four simple shapes appeared on the screen. When the participant clicked the cross, the trial began, and pictures of four objects appeared. A spoken instruction was presented over speakers, such as "point to the net; now click on it and put it below the circle." We tracked eye movements using an SR-Research Eyelink II head-mounted eye tracker, sampling at 250 hz. We tracked the probability of fixating each item

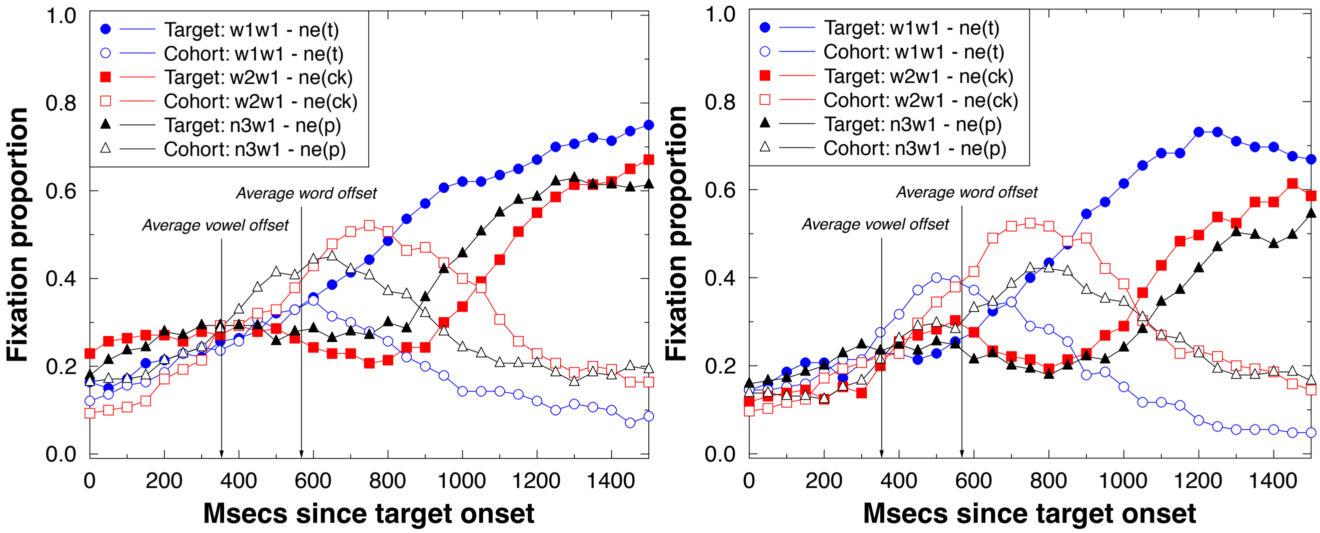

Figure 2: Subcategorical mismatch data for the top 50% (left) and bottom 50% (right) of readers from our community sample.

over time from the onset of the target word (e.g., *net*).

Results and discussion

Eye movements were parsed into saccades and fixations. Saccade time was attributed to the following fixation, since saccades are essentially ballistic; the initiation of a saccade is the earliest indicator of the choice to fixate the next gaze position. Eye tracking results are presented Figure 2. Qualitatively, the observed patterns for both halves of our sample resemble the (competitor) pattern in Figure 1. Notably, there is no apparent delay in the response to the bottom up signal in either half, when compared with the university sample in Figure 1. There are some differences between the two subsets in Figure 2 in the relative magnitude and timing of competitor proportion curves, but the most salient difference between the groups is in the target fixations in the mismatch conditions. The top 50% show the same ordering observed by Dahan et al.: W1W1 > N3W1 > W2W1. However, the pattern for the bottom 50% is W1W1 > W2W1 > N3W1. We explored this using a 2 (subset) x 2 (W2W1, N3W1) ANOVA on mean target fixation proportion in the window from 200 msecs after word onset (the expected average latency for a signal-driven saccade) to 1200 msecs (approximate target peak latency).

There was a main effect of Subset (top=0.40, bottom=0.32; $F(1,54)=6.8$, $p=.01$), but not Condition ($F < 1$), and a significant interaction ($F(1,54)=4.2$, $p<0.05$). This was due to a reliable effect of condition for the top subset (W2W1=.37, N3W1=.44; $F(1,54)=5.4$, p=.03), but not for the bottom (W2W1=.34, N3W1=.30; $F<1$).

Thus, there are several interesting patterns. There is no apparent delay in bottom up response. However, the later time course is different in both subsets compared to the sample of Dahan et al. (2001), and the subsets differ from each other. Most notably, it appears that lexical competition differs in the bottom subset. Target proportions for the mismatch conditions are depressed throughout the analysis window in comparison to the top subset, and the two mismatch conditions do not differ reliably in the amount of target interference they cause for the bottom subset.

Computational modeling To make sense of these patterns, we turned to the jTRACE re-implementation (Strauss, Harris, & Magnuson, 2007) of the TRACE model (McClelland & Elman, 1986) that includes several additional features (graphical user interface, plotting and scripting utilities). Starting with the default parameters used by Dahan et al. (2001) to obtain the simulations shown in the middle panel of Figure 1, we explored a wide range of changes to several parameters, one at a time. The goal was to determine whether any parameter could be changed to produce the observed changes in the bottom subset: increased competition effects without slowing initial lexical access. We tested a variety of parameters in TRACE (feedforward and feedback gain at various points, addition of input and "sensory" [model-internal] noise). *Lexical decay* was of particular interest, as the parameter McMurray et al. (2010) claim best fits individual differences in a lexical competition in a group of adolescents with a range of language and cognitive abilities; however, its influence is too weak and late. Two parameters could simulate the general trends: reducing *phonemic* or *lexical lateral inhibition* by approximately 50% from default levels. Reducing inhibition does not affect initial activation rates, but it allows larger competition effects because it delays the impact of late-arriving bottom-up disambiguation. In particular, it predicts larger cohort competition effects (note slight trends in this direction in the bottom subset) as well as less differentiation in target trajectories for the mismatch conditions.

Experiment 2

In Experiment 2, we continued our exploration of our sample's phonological abilities by examining lexical competition in the context of an artificial lexicon learning task (based on Magnuson, Tanenhaus, Aslin, & Dahan, 2003). This allowed us to simultaneously study

Figure 3: pictures of unusual animals used in the artificial lexicon study (Experiment 2).

phonological competition effects in word recognition (how strongly do "cohorts", like /pibo/ and /pibu/, compete? How strongly do rhymes, like /pibo/ and /dibo/, compete?) and word learning ability. Magnuson et al. (2003) were motivated in part by the goal of precisely controlling lexical characteristics such as phonological similarity, frequency, and neighborhood density. This approach has an added advantage for our sample. To the degree that our sample diverges from the performance of typical participants using real words, it is very difficult to determine the locus of the difference. There may be deep reasons, such as differential organization of processing mechanisms, or shallow ones, like simple differences in vocabulary size. An artificial lexicon paradigm allows us to put participants on maximally similar footing. While participants differ in linguistic and cognitive abilities, the items are equally unfamiliar to all.

Predictions Virtually any variant of the phonological deficit hypothesis might predict poor readers would perform worse in learning the artifical lexicon. With respect to the time course of cohort and rhyme competition, two precedents using familiar words in the visual world eye tracking paradigm suggest possible outcomes. Desroches, Joanisse, and Robertson (2006) examined cohort and rhyme competition in children with dyslexia. Unlike typically developing peers, they did not exhibit rhyme competition effects. In contrast, McMurray et al. (2010) reported that adolescents meeting criteria for SLI showed stronger cohort *and* rhyme effects, though only in the late time course. Thus, we might expect to see typical cohort effects but weak or absent rhyme effects (consistent with Desroches et al.) or late-enhanced cohort and rhyme effects (consistent with McMurray et al.).

Methods

Participants A subset of participants from Experiment 1 participated in Experiment 2: 14 individuals from the top 50% and 20 from the bottom 50%.

Materials 8 artificial words were constructed with one "cohort" (onset) competitor in the artificial lexicon and one rhyme. The words were /pibo, pibu, dibo, dibu, tupa, tupi, bupa, bupi/. The visual materials were pictures of 8 unusual animals (see Figure 3). Names were mapped randomly to pictures for each subject.

Procedure Each trial had identical structure. A fixation cross appeared in the center of the screen. When the participant clicked the cross, the trial began. Two pictures appeared, to the left and right of the cross. 500 ms later, an instruction was played, such as "find the pibo." At first, participants could only guess. If they clicked on the incorrect object, they heard "try again." When they clicked the correct object, they heard feedback, such as "that's right, that's the pibo!" The experiment consisted of 8 blocks of 24 trials. Each item appeared as the target 3 times per block, once each with its cohort, its rhyme, and an unrelated item. Thus, each block had 8 cohort, rhyme, and unrelated trials. There was no formal test; we measured behavior continuously over learning.

Results and discussion

Accuracy and response time Accuracy and response time (for accurate trials) are shown in Figure 4 for the two groups. We conducted ANOVAs with factors Type (Cohort, Rhyme, Unrelated) and Block for accuracy and RT. In the interest of space, we will only briefly summarize the results. The two subsets were both reliably more accurate for Unrelated than Rhyme trials, and more accurate in Rhyme than Cohort trials. In RT, the main effect of Type was not reliable for the top subset; in planned comparisons, none of the Types differed another. But for the bottom subset, Cohort trials were significantly slower than both Rhyme and Unrelated trials, which did not differ from each other. Thus, the bottom subset seemed to show less rhyme interference.

Fixation proportions over time are presented in Figure 5 by just showing target fixations (competitor fixations are essentially complementary) averaged over all correct trials (as the patterns did not change substantially with training). For qualitative comparison, results from a sample of 14 U. of CT (UConn) undergraduates are presented. Qualitatively, there is a very striking result. There are clear effects of both Cohort and Rhyme for the UConn sample. The Cohort effect is stronger and earlier, as with real words (Allopenna, Magnuson, & Tanenhaus, 1998; Desroches et al., 2006), while the Rhyme effect emerges later. Growth curve analysis (Mirman, Dixon, & Magnuson, 2008) revealed reliable intercept differences for the TD group (Unrelated > Rhyme > Cohort), analogous to differences in mean proportion over the analysis window. In contrast, the two community sample groups shows strong Cohort effects, but delayed Rhyme effects. The Rhyme condition differs reliably from the Unrelated condition for the top 50%, but not for the bottom 50%.

Our results are consistent with those of Desroches et al. (2006), who reported an absence of rhyme effects in children with dyslexia using a similar eye tracking paradigm with familiar, real words. They are partially consistent with the recent report of McMurray et al. (2010) that adolescents with SLI show larger but *later* competition effects than typically developing peers. We again turned to the model in order to explore possible bases for such a pattern.

Computational modeling As with Experiment 1, we used the jTRACE re-implementation (Strauss et al., 2007) of TRACE. Because TRACE is not a learning model (though see the Hebbian version of TRACE version developed by

Mirman, McClelland & Holt, 2006), we treated TRACE as a model of the stabilized system at the end of learning. Again, we changed one parameter at a time, looking for a change that would leave the magnitude and timing of the cohort effect intact while ideally wiping out the rhyme effect. We again tried several parameters. Lexical decay does not selectively affect rhyme effects. Reduced lexical lateral inhibition actually boosts rhyme effects. Only one parameter could generate the correct trends: a *reduction in lateral inhibition at the phoneme layer*. As it is reduced, rhyme effects are weakened and delayed, while leaving the cohort time course largely intact (though cohort effects are somewhat amplified). This counter-intuitive outcome follows from what happens to phonemes other than the initial phoneme of the target word. With inhibition reduced, similar phonemes get much more activated. Even though the phoneme inhibition parameter is lower, there is actually greater inhibitory flow at the phoneme level, putting rhymes that differ from the target in initial phoneme by more than a single feature at a disadvantage. Interestingly, lateral inhibition at the phoneme level was one of two parameters that could achieve the correct pattern to fit the bottom 50% subset behavior in Experiment 1.

Summary In Experiment 2, good and poor readers achieved similar accuracy in artificial lexicon learning. However, the time course of learning was substantially different, with poor readers exhibiting slower learning in early trials. Poor readers showed similar on-line onset (cohort) competition effects as better readers, but failed to exhibit a reliable effect of rhyme competition (instead showing a weak, delayed effect). This converges with a report that children with dyslexia did not exhibit rhyme effects in a similar study using real words (Desroches et al., 2006). In TRACE simulations, the only way to substantially reduce rhyme effects without inappropriately perturbing cohort (onset) effects was to reduce lateral inhibition at the phoneme level -- a parameter change that can also capture the poor reader differences in Experiment 1.

General Discussion

Adult poor readers continue to differ from good readers in phonological processing. Our poor readers showed greater interference effects from misleading coarticulation than better-reading peers in Experiment 1. Poor readers learned new words with a different trajectory than better readers in Experiment 2, and exhibited late, weak rhyme competition effects. The two primary patterns of differences -- enhanced competition due to misleading coarticulation and absence of rhyme effects -- can both be modeled in TRACE via reduced lateral inhibition at the phoneme level. The convergence on phoneme inhibition in the simulations of Experiments 1 and 2 increases our confidence that this parameter manipulation is capturing something important about phonological differences in poor readers. One next step will be to use the re-parameterized model to generate predictions for poor readers in new tasks.

We do not wish to imply that we believe that there are

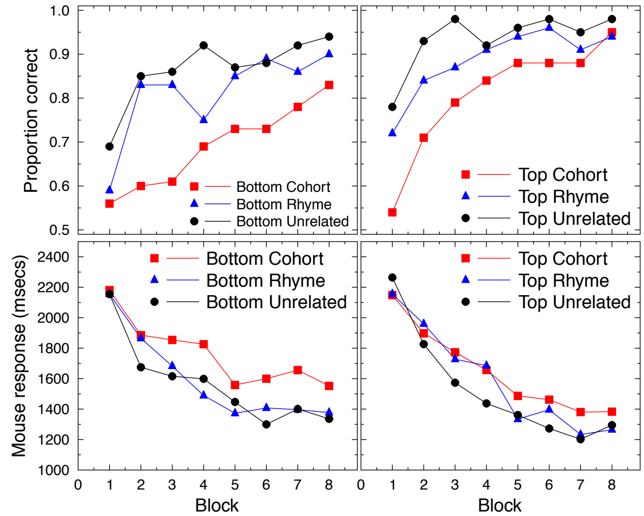

Figure 4: Accuracy (top) and RT for the bottom 50% of readers in our sample (left) and the top 50% (right) by training block.

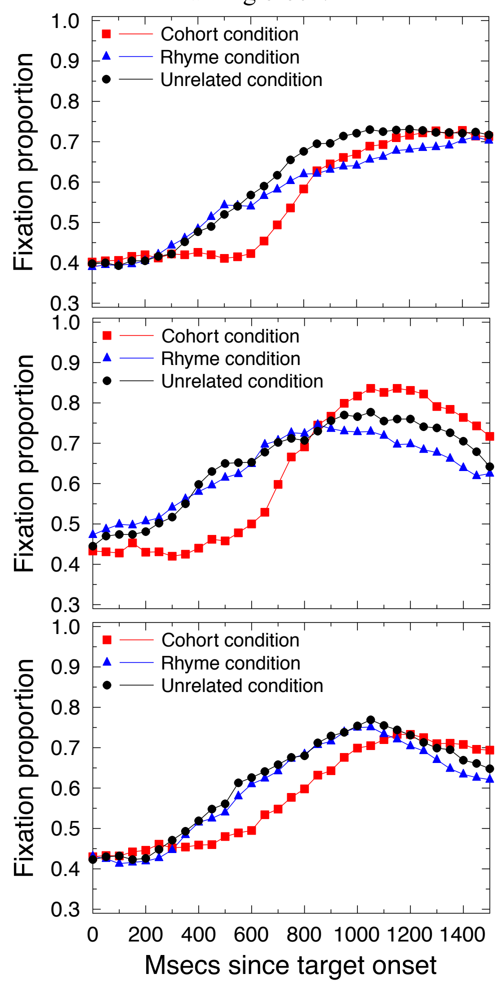

Figure 5: Target fixation proportions over time in Experiment 2, collapsed across block and only including correct trials, averaged over all 8 blocks. **Top**: typical university sample. **Middle**: to 50% of community sample readers. **Bottom**: bottom 50% of community sample readers. Patterns varied only slightly by block.

discrete representations of phonemes in the brain, let alone a discrete parameter controlling lateral inhibition. The ability of TRACE to simulate differences based on reduced phoneme inhibition instead points to the level of phonological organization in the dynamical system it is meant to simulate, i.e., the mechanisms underlying human word recognition. Thus, our simulations may identify the level of the system -- phonological organization -- that appears to be crucially different in poor readers.

Our results are potentially consistent with any form of the phonological deficit hypothesis, although they somewhat favor accounts that assume a typical level of phonetic resolution (given that poor and better readers showed similar timing in early lexical activation), and differences in the stability of phonological representations. In particular, our results may be compatible with the phonological access hypothesis (Ramus & Sjenkovits, 2008). However, our results also suggest differences in phonological access may be more subtle than suggested by Ramus and Sjenkovits, who emphasize working memory demands in conventional tasks that most clearly identify phonological deficits. That we observed differences in the time course of lexical activation, competition and learning in poor adult readers in minimally demanding, naturalistic tasks suggests that the locus of the phonological deficit may be a more low-level property of the system, even though this deficit may require difficult tasks or sensitive measures to be detected. We hope that our continuing exploration of individual differences in adult poor readers will illuminate this possibility further.

Acknowledgments

Supported by NIH grant HD-40353 to Haskins Laboratories.

References

Allopenna, P. D., Magnuson, J. S., and Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419-439.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for Vocabulary: Reading Skill Differences in Young Adults. *Journal of Learning Disabilities, 40.3*, 226-243.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: evidence for lexical competition. *Language and Cognitive Processes, 16*, 507-534.

Desroches, A.S, Joanisse, M.F. & Robertson, E.K. (2006). Specific phonological impairments in dyslexia revealed by eyetracking. *Cognition, 100,* B32-B42.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the Meanings of Words in Reading: Cooperative Division of Labor Between Visual and Phonological Processes. *Psychological Review, 111*, 662-720.

Joanisse, M.F. (2004). Specific Language Impairments in Children: Phonology, Semantics and the English Past Tense. *Current Directions in Pysch. Science, 13,* 156-160.

Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *J. Speech & Hearing Research, 37,* 418-421.

Magnuson, J.S., Tanenhaus, M.K., & Aslin, R.N. (2008). Immediate effects of form-class constraints on spoken word recognition. *Cognition, 108(3),* 866-873.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., and Dahan, D. (2003). The time course of spoken word recognition and learning: Studies with artificial lexicons. *J. Experimental Psychology:General, 132(2)*, 202-227.

Marslen-Wilson, W.D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psy. Rev., 101*, 653-675.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psych., 18*, 1-86.

McMurray, B, Samelson, V.M., Lee, S.H., & Tomblin J.B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cog. Psych., 60,* 1-39.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory & Language, 59(4),* 475-494.

Mirman, D., McClelland, J.L., & Holt, L.L. (2006). An interactive Hebbian account of lexically guided tuning of speech perception. *Psychonomic Bulletin & Review, 13*, 958-965.

Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *Quarterly J. Experimental Psych., 61*, 129-141.

Shankweiler, D. P., Mencl, W. E., Braze, D., Tabor, W., Pugh, K. R., & Fulbright, R. K. (2008). Reading Differences and Brain: Cortical Integration of Speech and Print in Sentence Processing Varies with Reader Skill. *Developmental Neuropsychology, 33.6* 745-776.

Shankweiler, D., Liberman, I.Y., Mark, L.S., Fowler, C.A. & Fischer, F.W. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 531-545.

Snowling, M.J., 1981. Phonemic deficits in developmental dyslexia. *Psychological Research, 43*, 219-234.

Snowling, M. J. & Hulme, C. (2005). *The Science of Reading: A Handbook*. Blackwell.

Strauss, T. J., Harris, H. D., & Magnuson, J. S. (2007). jTRACE : A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. *Behavior Research Methods, 39,* 19-30.

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

Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. C. (1995). Integration of visual and linguistic information is spoken-language comprehension. *Science, 268*, 1632–1634.

Ziegler, J., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin, 131*, 3-29.