

Research Strategy

(a) Significance

Reading disability (RD) in children represents a seriously limiting intellectual disorder, causing children difficulty in learning, frequently across their lifespan. While estimates of the degree and severity of RD vary, at least 15% of children have RDs (IRA, 1998; NICHD/NRP, 2000a). Reading is necessary for success in school, thus is critical for success in life. Schuele & Boudreau (2008) argue that phonemic awareness (PA), including segmentation and blending, is vital for learning to read, but does not occur naturally, without training. If current PA training were generally successful, RD would be resolved early instead of presenting a lifelong challenge. The reasons why children with RD have such poor PA have yet to be determined, but to our knowledge, the sensory abilities of children with RD, to discriminate and identify the full speech-sound repertoire of English consonants (Cs) and vowels (Vs), have yet to be examined. Some recent work on a small number of speech contrasts in French and Dutch speaking children with Specific Language Impairment—a disorder resembling RD—is suggestive (Ziegler et al., 2005; Vandewalle et al., 2012).

A major barrier to progress in the field is that normal speech perception is a complex, poorly understood process, with many potential barriers. Speech is spontaneously learned with very little direct feedback on success. Children first learn to understand speech from their parents (caretakers) when learning to talk, and only then, how to read. We propose that RD problems start early, prior to actual reading instruction, and are fundamentally related to the auditory perception of speech sounds (phones), which we will denote as *phonetic perception* (PP), but others have labeled *phonology* (i.e., receptive rather than expressive phonology; e.g., Dawes and Bishop, 2009; Mody et al., 1997; Snowling, 2001; Studdert-Kennedy and Mody, 1995; Vandewalle et al., 2012). Early auditory difficulties (e.g., from otitis media with effusion) is known to limit the speech learning process (Klein, 1984; Ptok, 2005; Paden, Novak & Beiter, 1987; Mody et al., 1999; Rosenfeld et al., 2011). When this happens, the child and parents may never be aware of any hearing loss. Although our study focuses on residual speech-sound confusions in older elementary school children with RD, the possibility that they once experienced early undetected hearing loss is supported by possible correlations in their pathologies, supported by our **Preliminary Studies** and by Co-PI Allen's published work on CV confusions in the hearing impaired (Phatak et al., 2009; Allen 2012, 2011).

The present study will improve scientific knowledge and clinical practice by exploring new methods of evaluating PP. In most children, learning to read can be efficient and automatic, yet highly error prone in the child with RD. If we find that school children with RD have phonetic perception problems, our study could point the way to why, and how, RD develops, and suggest an onset at a very early age (before kindergarten reading instruction). Our study has the potential to change clinical concepts and interventions by providing child-dependent measures of deficiencies in PP. These measures should allow us to increase the effectiveness of training, by concentrating on the underlying source of PA and reading.

Another barrier to progress is that RD studies often use average performance measures, thus masking key individual differences. For clinical application, individual performance is crucial (as in the fitting of glasses, hearing aids, or cochlear implants). In the field of speech-language pathology (SLP), it is well established practice to map the profile of speech production errors of the individual child with an articulation-phonological disorder. By mapping *perceptual* error profiles in the same way, for the individual child with RD, we may be able to discern the range of perceptual errors and error rates that can be found in the RD population. While improving scientific knowledge, this should also advance clinical practice by enhancing the diagnosis of (a) RD, (b) assessment for intervention and ultimately, (c) individual intervention outcomes.

(b) Innovation

Our study is innovative in five ways: *First*, we pair the PI's expertise in their respective fields. Cross-disciplinary research is an important advantage when investigating such complex problems. A licensed and certified SLP, the area of Co-PI Johnson's research expertise is in phonological (DeThorne et al., 2009; Ha et al., 2009; Lin & Johnson, 2010), language (Marinellie & Johnson, 2002, 2003; Lyons et al., 2010), and literacy development and disorders in children (Frame et al., 2008; Yang & Johnson, 2009; Johnson et al., 2011a). The area of Co-PI Allen's expertise is in auditory and cochlear speech processing (Allen, 2008), middle ear disease (Allen et al., 2005), speech perception in the normal hearing (Li et al., 2010; Singh & Allen, 2012), speech hearing loss (Allen & Li, 2009; Phatak et al., 2009) and RD.

Second, we seek to shift current research paradigms by choosing the broader term, *reading disability*, rather than the more narrow term, *dyslexia*, to accommodate the possibility that children with reading problems have co-morbid and perhaps undocumented problems with spoken language (e.g., speech perception or lexical or grammatical difficulties; cf. Catts, 1993; Catts & Kamhi, 2005; Catts et al., 2002; Flax et al., 2003), attention (King et al., 2003; Loo et al., 2004; Shaywitz & Shaywitz, 1994; Willcutt & Pennington, 2000;), and importantly, PP. A related argument is in Dawes & Bishop, 2009, pp. 443-444.

Third, child RD experts fully appreciate *phonemic* awareness (PA; Ehri et al., 2001; Stahl & Murray, 1994; Wagner & Torgesen, 1987) and understand its relevance to reading (Catts & Kamhi, 2005; IRA, 1998; NICHD/NRP, 2000b; Rvachew et al., 2003;). Yet RD experts do not either have a well defined working hypothesis of how PA operates during reading, or even an agreed upon measure of PA. Measures of PA vary widely, from simple syllable-level awareness tasks to complex, phoneme-level manipulation tasks (Schuele & Boudreau, 2008).

Theorizing about PA's relation to RD has been limited because the speech science community has not fully understood the nature of consonant perception (i.e., PP). Prior to 2005, the basic perceptual units (acoustic features) of consonants had not been identified (Allen, 1994; Allen et al., 2005; Blumstein & Stevens, 1980; Liberman, 1996; Shannon et al., 1995).

Co-PI Allen's recent investigations on phonetic features (Phatak and Allen, 2007; Phatak et al., 2008; Li et al., 2010) have

naturally led us to the hypothesis that basic perception of phones (PP) has gone astray in the child with RD, and underlies poor PA. Allen’s studies have identified the acoustic features used by normally hearing adults to identify hundreds of tokens of CV syllables and demonstrated that consonants are identified by normal hearing ears with zero error, even given large amounts of noise (Singh & Allen, 2012). This newly acquired understanding of consonant feature decoding opens a door for innovative RD research. Our *Preliminary Studies* have been designed to further test this hypothesis, on possible interesting parallels between the hearing impaired and the RD populations (Phatak et al., 2009). If our instinct regarding this parallel is correct, it could pave the way to a new level of understanding of RD ears.

Fourth, to measure PP, which we theorize occurs earlier in the auditory processing stream than PA, we break from the methodological approaches traditionally used in the PA literature, namely, we ask children only to judge and imitate whole nonsense syllables. Unlike the PA literature, our tasks never require the child to parse and recognize the *phoneme* as a linguistic unit of real words, an ability that Ziegler and Goswami’s Grain Size Theory of reading acquisition and dyslexia (2005; also Goswami, 2010) argues is learned *from* successful reading rather than serving as its foundation.

Lastly, though addressing C and V perception, our work represents a shift away from the algorithm used by Tallal (1980), Tallal et al. (1996) and Merzenich et al. (1996), which yields substantially modified tokens of speech to train children with RD to better hear consonant distinctions. We use only naturally produced tokens by multiple talkers, to retain a better match between our experimental tasks and spontaneous speech. This should improve: (a) accuracy in diagnosing a child’s perceptual difficulties in conversational speech, (b) translation of our method to typical settings in which reading specialists work, and (c) generalization of a child’s learning to natural situations. We focus on precise controls. Based on the results of Allen et al. from 2005 to 2012, we suggest that isolated consonants and vowels are the source of PP, not CV transitions.

(c) Approach

Preliminary Studies: In two earlier studies (Johnson et al., 2007a,b,c,d; 2010a,b; 2011b,c), we reasoned that children’s perceptual difficulties might be more fundamental than poor PA or poor auditory to visual phonic to phone mapping, arising from acoustic-phonetic properties of speech sounds, rather than their linguistic properties (in meaningful words). Consequently, we systematically measured confusions of a full array of English nonsense syllables, to examine specific phonetic perceptual confusions in children with a history of RD.

Participants: In collaboration with The Reading Group, a local nonprofit reading tutoring center, we compared two groups of children: 11 children with *reading disabilities* (RD) and 6 *reading controls* (RC), age 8-11 years. The RDs attended weekly reading lessons at the center. All participants were administered an assessment battery of reading, speech, language, phonological awareness, nonverbal cognition and hearing, by speech-language pathologists (SLPs).

Methods: We measured speech perception in two studies, with 17=11RD+6RC and 15 participants, respectively; with 10 hour-long sessions per study. In the first study, we presented a *Syllable Confusion Oddball (SCO)* task: a speech discrimination task to determine which of 24 consonants (Cs) and 15 vowels (Vs) caused confusion errors for a child. The child listened over headphones, via a computer, at a comfortable listening level, to random sequences of three nonsense CV or VC syllables *spoken by three different talkers* drawn from a set of 20 professionally recorded talkers (Linguistic Data Consortium database LDC2005S22 “Articulation Index Corpus,” University of Pennsylvania; Fousek et al., 2004). For example, the child would hear [da] (Voice 1) – [da] (Voice 2) – [fa] (Voice 3), in which two of the three stimuli were the same CV, but the 3rd stimulus differed, in this example, in /f/. The child’s task was to indicate the odd syllable. The child could request a trial be repeated, but received no performance feedback. The computer program randomly selected the syllable triads and the three talkers, thus the number of trials varied across sounds (mean $\mu = 41$).

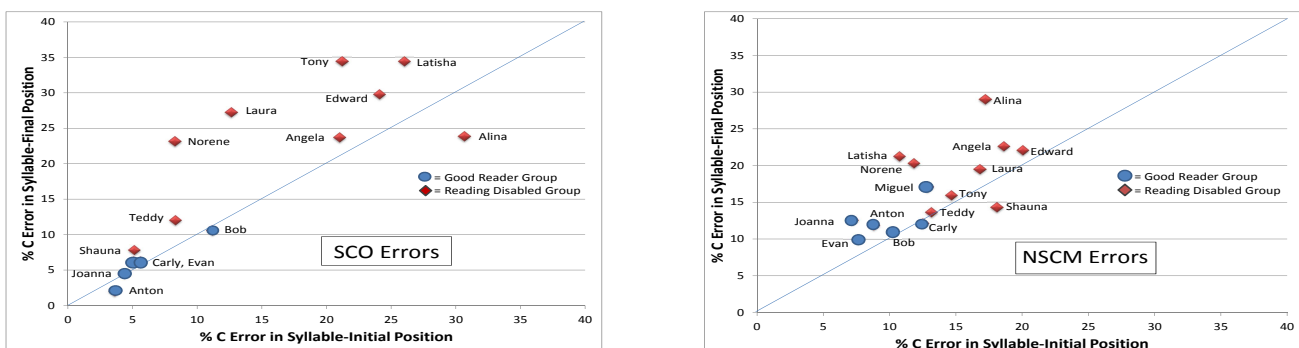


Figure 1: Percent of consonant error in syllable-initial and syllable-final position for the Reading Disabled (RD; red diamonds) and Control groups (RC; blue circles) on the SCO (left) and NSCM (right) tasks. Three participants were excluded from the left panel, due to too few trials and two subjects who did not return for the NSCM task were removed.

In the second study, we presented a *Nonsense Syllable Confusion Matrix (NSCM)* task: a C,V speech identification task, designed to discern each participant’s idiosyncratic confusions. With NSCM we presented one syllable at a time and the child repeated it back. Responses were entered into the laptop computer by one examiner and transcribed phonetically (in the International Phonetic Alphabet) by a second. Because of random presentation, the number of trials varied ($\mu = 69$).

Results: For the SCO task, RCs (mean $\mu = 93\%$ correct, SD $\sigma = 3\%$) perceived speech sounds significantly better than children with RD ($\mu = 81\%$ correct, $\sigma = 10\%$), $t(14.9) = 4.09, p = .001$. All individuals in the RD group performed well above chance (chance being 33% on the SCO task). All individuals had some sounds for which the perceptual accuracy was 80% or higher, indicating that both groups could perform the task: Five of the six RCs perceived 100% of the 24 Cs and 15 Vs correctly. The sixth RC perceived 92% of Cs and 100% of Vs correctly; three children with RD performed similarly. Four RDs perceived 58 to 88% of Cs correctly, and from 27 to 100% of Vs. Four children with RD perceived only 13 to 35% of Cs correctly, and only 0 to 40% of Vs. Thus, the RC group had low error for nearly all sounds, whereas the RD group ranged from low error, to substantial difficulty, with many sounds. With respect to Cs, 88% of the 17 RD and RC participants perceived /s, h, n, r/ accurately; only the four lowest performing children in the RD group had difficulty with these four sounds (and with /w/). This suggests that RD difficulty with these five sounds might provide a quick screen.

For the NSCM task, the RC group ($\mu = 87\%$ correct, $\sigma = 1\%$) perceived speech sounds significantly better than the RD group ($\mu = 80\%$ correct, $\sigma = 7\%$), $t(9.68) = 3.55, p = .006$. Effect size was determined via arcsine transformed values on the average group scores. Cohen's d was 1.96, indicating a large effect. As with the SCO task, all participants had some sounds for which accuracy was $\geq 90\%$, demonstrating that they could do the task. Ranges for the number of accurate sounds went from RD: 6-14 to RC: 12-18 for consonants and RD: 2-8 to RC: 8-10 for vowels. The best consonant perception in RD and RC children was for /d, k, w, j/ ($\leq 6\%$ error): RCs additionally perceived /t, s, f, h/ with a similar low error. Low error was seen for the vowels /i, u/ for both RD and RC children. RCs additionally perceived /e, o, ai, au, oi, ɜ:/ with low error. These results are summarized in Fig. 1 which compares percent error for Cs in syllable initial (CV) and final (VC) position for the two tasks. RCs show substantially fewer errors than RDs on both tasks, with only a small overlap. RCs show little influence for syllable position, whereas a number of children with RD make more perceptual errors in syllable-final (VC) position.

From the assessment battery, three reading measures from the WRMT-R significantly correlated with performance on the SCO task: Average percentile rank (PR) for reading fluency correlated highly and significantly with SCO performance ($r = .73, p < .001$), as did average PR for word attack ($r = .67, p < .01$). Average PR for reading comprehension on the GORT-4 also correlated significantly, but more weakly, with SCO performance ($r = .55, p < .05$). One oral language measure, PR for Recalling Sentences on the CELF-4, also correlated significantly, but only weakly, with SCO performance ($r = .53, p < .05$).

Figure 2 compares the RD vs. RC log-error on the NSCM VC task. Here the RD group has substantially greater error for most final Cs. Next, consonant confusion matrices were generated for each child. As shown in Fig. 3, RDs showed nearly twice as many confusions as RCs. Only severe errors ($\geq 20\%$) are shown. Most confusions were for fricatives and affricates. The degree of confusions may be ordered as place, voicing, and finally manner being the least. Five confusion were shared by the groups, however the RDs had twice the final consonant error (15 confusions) as the RCs (7), including some confusions of nasals and plosives. This indicates more quantitative than qualitative group differences. Similar errors were evident in syllable-initial position, except that only fricatives and affricates were affected.

Conclusions: Children with RD have poor PP that relates to all aspects of reading, including fluency, decoding, and even comprehension. All children with RD perceived some sounds well, indicating that they could do the SCO task, and that poor performance on certain sounds was not due to a general auditory memory problem, but rather to selective difficulties with PP for certain Cs and Vs. Particular C confusions on the NSCM task were primarily for fricatives and affricates, with children in the RD group showing many more idiosyncratic confusion than the controls. Thus, our preliminary studies suggest that although children with RD do not experience extreme PP difficulty (i.e., they are well above chance on most sounds), they are significantly and measurably worse than the controls. *Cumulatively*, this increased *level* of confusions could well result in considerable difficulty when learning to reading. Patterns of confusion are child-specific (idiosyncratic), and thus would require identification on an individual basis, for effective intervention.

Feasibility of the Proposed Project: In our preliminary collaboration with The Reading Group center over more than a three-year span, we were successful in recruiting and running 17 participants. That collaboration continues, allowing the recruiting of participants for the proposed study. Children regularly attended our experimental sessions (generally twice a week), completing 23 sessions for both studies in 3 or 4 months. Having regularly-scheduled lessons at The Reading Group center (held immediately following experimental sessions) boosted attendance in our preliminary studies and allowed us to document the children's success in their concurrent reading intervention program. Consequently we have included funding for each participant's reading lessons at The Reading Group center in our proposed budget. Furthermore, the methods we have developed have been shown to be feasible. Children enjoy the experimental sessions, including snacks, short play breaks and a modest remuneration per session (all having IRB approval). The SCO and NSCM tasks appear to be a feasible way to define and measure PP, because overall performance was well above chance for all participants. Nor did performance

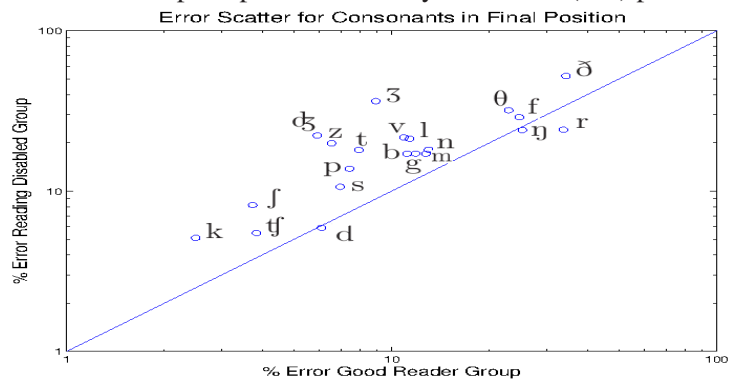


Figure 2: NSCM task final C error [log-%], RD vs. RC groups.

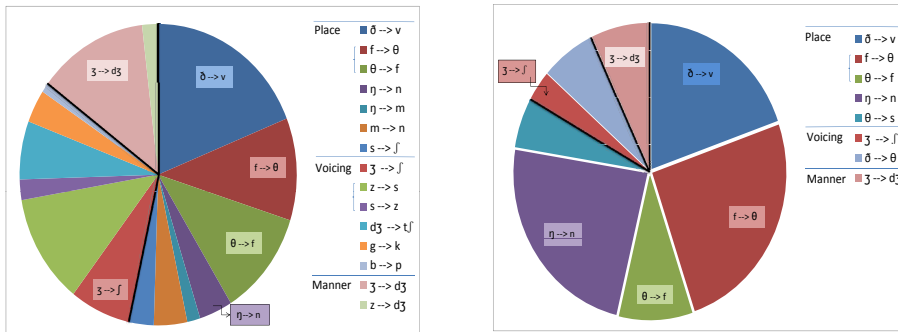


Figure 3: Distribution of severe C confusions in VCs, occurring $\geq 20\%$ of the time for any child, for the RD group (left) and RC group (right). The RD group had 74% more $\geq 20\%$ confusions in VCs ($N = 222$ judgments by 2 examiners for 9 children) than the RC group ($N = 127$ for 6 children).

appear to decline due to boredom or fatigue, as all RDs had target sounds with $< 20\%$ error in both tasks.

Overall Strategy for the Proposed Project

Specifically designed tasks corresponding to each of the four hypotheses (H1-H4) are to be tested via a specific experiment, hierarchically organized to determine if poor PP contributes to memory and print difficulties. Each experiment includes RD and RC groups. Data sets for the two groups will be compared using repeated measures ANOVA and more general statistical tests, such as Fisher’s exact test (Singh and Allen, 2012).

Methodology

Participants: Participants will be 39 children having documented RD and 30 RC children (good readers). All participants will be 8 to 12 years old, to ensure that they have had a number of years of reading instruction and have adult-like articulation, and to optimize the chances that they can do our experimental tasks.

Participants with documented histories of reading difficulties (RDs) will be recruited from The Reading Group center. Parents will be asked to fill out an extensive questionnaire about the child’s physical, speech-language, and reading development; vision, hearing, and health; and educational history. Control children will be recruited from local schools and other community facilities, lab web site postings, and local newspaper announcements. Their parents should report no history of reading difficulties or any remedial services for reading. Control children will be paid \$15 per session. For both the RD and control groups, two to three sessions will be devoted to a battery of standardized tests or protocols for hearing screening, nonverbal cognition (Matrices subtest, Kaufman Brief Intelligence Test-2), speech (Goldman-Fristoe Test of Articulation-2), language (Peabody Picture Vocabulary Test-4; Clinical Evaluation of Language Fundamentals-4; and a spontaneous language sample), phonological awareness (Comprehensive Test of Phonological Processing; and the Nonword Repetition Test, Dollaghan & Campbell, 1998), and reading (Word Identification and Word Attack subtests, Woodcock Reading Mastery Tests-3; and the Gray Oral Reading Tests-5), administered by a graduate RA with an MA in SLP, under the supervision of Co-PI Johnson, a licensed and certified SLP. Most of these procedures are identical to our preliminary studies.

Data Collection: By providing funding for two lesson packs at The Reading Group center, we hope to secure regular, twice weekly attendance of children with RD at 1-hr experimental sessions (held in conjunction with the child’s 1-hr reading lesson) and to thus document any progress in intervention. (We have been successful in this in the past at The Reading Group center.) Intervention plans and performance reports are provided by the child’s reading teacher, at the end of each 10-session lesson pack. In our experience the nature and success of past and concurrent intervention is seldom documented in studies of children with RD. Thus we will collect current and past plans/reports for all participants with RD, to assist with data interpretation. All participants will be run at The Reading Group, or in our labs, in the Dept. of Speech and Hearing Science or the Beckman Institute, when appropriate.

In all tasks, the nonsense syllables will be drawn from the LDC database of 20 professionally recorded talkers as our preliminary studies (Fousek et al., 2004), and played from a laptop computer. The child will listen over headphones and wear a close talking microphone mounted on the head, to record responses. Participants will be assigned to one of three cohorts, on a rolling admission basis. Each cohort will consist of 13 participants with RD and 10 control children. Each cohort will participate in one experiment composed of two of the following tasks.

Combined SCO/NSCM tasks (H1: PP). The first task is a combined one and examines the nature of PP by measuring confusion between randomly presented CV or VC syllables. Data for each portion of the task will be collected on each trial: the SCO portion is similar to the one we used in our preliminary studies, only in the proposed project we have changed the number of comparison units in a trial from three to four, to set chance at 25%. In this task, the participant will listen to four syllables spoken by four different talkers. Three syllables will be the same and the fourth will differ by one C or V, in either syllable initial or final position. The participant will use a touch-sensitive computer screen to indicate which of the CVs (1 to 4) is the “oddball.” Sufficient trials will be presented to test all English Cs and Vs in both initial- and final-syllable position, with sufficient trials to be statistically significant at the $\alpha = 0.05$ level, based on either ANOVA or Fisher’s exact test.

To complete each trial, the participant will then repeat back the similar and dissimilar sounds. This portion is similar to the NSCM task used in our preliminary studies, but integrated into each SCO trial. The NSCM data will be used to form individual C and V confusion matrices. Immediately following the participant’s SCO response, the computer software

will turn on the head-mounted microphone and record the child's spoken responses. This is intended to provide accurate response time data. During the NSCM portion of a trial, the examiner will electronically record IPA symbols for Cs and Vs the participant spoke into the microphone. A second transcriber will also electronically enter transcription judgments. The SCO and NSCM tasks will help us to accomplish Aim 1, to measure the ability of children with RD to aurally perceive speech sounds, and Aim 2, to map out individual confusions for the two groups.

Memory task (H2: Short-term Auditory Memory). In this task, the duration of each of the SCO stimulus strings will be increased, to push the subject's phonetic memory limits. Thus, the SCO task will be replicated with CVCV and VCVC bisyllabic stimuli. Bisyllables will be created by combining individual syllables from our prerecorded database, with the constraint that the second syllable must contain both a different C and V than the first syllable. Statistically sufficient numbers of trials will be presented, to test all English Cs and Vs in syllable initial and final position, of both the first and second syllable of the bisyllables. (We have the option of testing triphones, i.e., CCVs, VCCs, and CVCs, as an alternative.)

Integration task (H3: Integration of Auditory and Visual Streams). In this task, participants will view a random printed string of four different nonsense syllables (all CVs or all VCs) on the computer screen while listening to an auditory version that differs in only one C or V. Here the task is to detect the point of mismatch between the auditory and visual signals and point to the corresponding letter or digraph in the visual display. Our hypothesis is that *the addition of the visual stream* will cause participants with RD to experience an *increased processing burden*, consequently their performance should decrease, whereas RC individuals should do well with the added visual input. Immediately after pointing, the participant will read the printed sequence aloud for recording. The examiner will have a matching copy of the print display on a second computer linked to the first and enter (in IPA symbols) or tag any errors made by the child, for later review.

Plasticity task (H4: Auditory Plasticity). In this task, participants will receive feedback when they make a perceptual or spoken error, to allow the participant to learn what she/he is doing wrong. Then we will compare the error rate pre- and post-feedback. The idea is to determine whether a simple training approach (i.e., merely providing feedback) will help a child improve her or his perceptual performance (indicating learning, i.e., demonstrating plasticity). It is important to demonstrate that training is possible, especially if we concentrate on the sounds for which the participant has the largest confusions.

In each of the four tasks, each child will be tested in 10-min blocks of approximately 20 to 30 trials, with four blocks per session and 5-min play breaks between blocks. Pacer items will be used to reward the child for each response (e.g., one bead on an abacus, one M&M, one penny, etc.). Pacer items allow the child to judge the remaining number of trials.

Experiments 1-3: The four tasks will be combined into three experiments as follows: **Expt. 1 (Cohort 1):** A combined (a) SCO/NSCM task for eight sessions will be followed by a (b) Plasticity task for eight more sessions. The only change between tasks will be the addition of feedback to the child about his or her response accuracy during the second task. Thus, this experiment tests whether training will improve a child's PP. **Expt. 2 (Cohort 2):** The combined (a) SCO/NSCM task will be followed by a (b) Memory task (eight sessions). The change between tasks will be in the total duration of the syllable string, i.e., in the Memory task, the four comparison slots in a string (trial) will be filled by bisyllables (i.e., CVCV or VCVC, with differing syllables in each bisyllabic nonsense word). Thus, this experiment tests whether poor PP in children with RD contributes to poor short-term auditory memory for speech. If children perceive some SCO sounds highly accurately, we will know that diminished perception of other SCO sounds is not simply due to a *general* short-term auditory-memory deficit. Thus, the apparent direction of causality will be that poor PP contributes to poor auditory memory, rather than vice versa. **Expt. 3 (Cohort 3):** The combined (a) SCO/NSCM task will be followed by a (b) Integration task (eight sessions). The change between tasks will be in the addition of print to the auditory signal and in the required response. In the Integration task, the child will be asked to detect the point of mismatch between the auditory stream and the print (rather than to detect an oddball syllable or imitate syllables, as for the SCO/NSCM task).

Data Analysis and Predictions

In all our tasks, even though participants are only asked to judge or imitate whole syllables, we will examine error rates for target Cs and Vs in our analyses. (During analysis, it is possible to separate out particular target Cs or Vs from the whole-syllable stimulus presentations, because the remaining parts of the comparison syllables are held constant, as in /fa fa ta fa/ or /ipəs ibəs ipəs/).

From the SCO portion of Expts. 1 through 3, we will determine each child's overall accuracy rate per target sound, as well as the slope of the child's accuracy per target across the eight sessions, including any plateaus or points of asymptote. We will prepare repertoires of speech sounds misperceived by each child ($\geq 10\%$ error). These PP repertoires will be compared for the RD and RC groups. Eventually we aim to identify acoustic features in particular Cs and Vs that are misperceived by individual children with RD. From the NSCM portion of Expts. 1 through 3, we will prepare confusion matrices (CMs for target sound vs. responses). (If any child has consistent speech production errors for certain sounds on our assessment battery, trials for those sounds will be omitted from the NSCM analysis.). From these CMs, we will create logs of the proportion of total error represented by particular confusions (e.g., p \rightarrow t), for each participant and group. We believe individual profiles are crucial for mapping the particular confusions experienced by a child with RD and for innovative planning of PP intervention.

A One-Between, Three-Within repeated measures ANOVA will be used to analyze PP accuracy (percent correct, with an arcsine transformation; a reaction-time measure is also possible), for the three cohorts combined. The Between variable will be Group (RD vs. RC) and the Within variables, task (SCO, NSCM), Speech Sound Type (C, V target) and Syllable

Position (initial, final). This will allow us to test the PP hypothesis (H1). We predict that the RD group will be less accurate than the RC group, and that consonants will be less accurate in syllable-final position. Similar ANOVAs will be used to compare accuracy across tasks in Expts. 1 and 2. In Expt. 1, Training (SCO or NSCM without or with training) will be a Within variable, to test the Plasticity hypothesis (H4). We predict that children with RD will improve their PP accuracy with training. In Expt. 2, Duration (short VC/VC syllabic strings in the SCO task or longer CVCV/VCVC bisyllabic strings) will be a Within variable, to test the Memory hypothesis (H2). We predict that PP will diminish with longer syllable strings for the RD group, but not for the RC group.

In Expt. 3, SCO performance will be used to determine which syllable initial and final C and V targets (24 and 15 targets, respectively) were Confused ($\geq 10\%$ error) or Correctly Perceived by each participant, forming four sets of Cs and four sets of Vs. The dependent variable for the Integration task is the percent (proportion) of mismatches correctly detected. Following arcsine transformation, a repeated measures ANOVA will be used to examine PP. Here SCO Accuracy will be a Within variable (confused vs. correctly perceived sounds), allowing us to test Integration (H3). We predict that the RD group will fail to detect sounds confused on the SCO task more often than those correctly perceived, and that the RD group will detect fewer mismatches than the RC group. Thus, this experiment will test whether poor PP contributes directly to reading difficulty. This analysis will be repeated for the read-aloud portion of the task. Again, we predict that children with RD will make errors that mirror their confusions on the SCO task. MANOVA will be used to compare error measures on the SCO and NSCM tasks to the clinical measures collected in our reading, speech, language, and cognitive assessment battery, for the RD and RC groups.

For all children with RD, we will collect past and current intervention plans and final performance reports for lessons provided by The Reading Group center. We will review these for the top and bottom quartiles of each cohort, to assist in interpretation of performance on the experimental measures. We will determine number of years of reading and reading-related intervention, and examine teachers' estimates of the child's reading level at the beginning of each session block during the proposed study; the portion of intervention planned for word identification, word attack (reading decoding), reading fluency, reading comprehension, spelling or writing, phonemic awareness, and listening comprehension; and any progress made. These qualitative descriptions may lead to quantitative measures that can then be analyzed statistically.

Management Plan We estimate it will take 3 to 4 months, with twice weekly sessions, to collect data for each participant (total ≈ 19 sessions, depending on the child's ability to complete each session as scheduled). We anticipate running Expt. 1 (Cohort 1) and beginning Expt. 2 during the first year (due to yearly funding limits), and running Expts. 2 and 3 (Cohorts 2 and 3) simultaneously during the second year. Data processing will overlap with the sessions.

Statistical analysis will be done as the data are collected, with the final analysis for the three cohorts done during the last 3-4 months of the second year. Regarding **Resource Sharing**, we will make our data available by posting them on the Internet within 2 years of publication. Table 1 provides **Benchmarks** and a **Timetable**.

Potential Problems and Alternative Strategies (see also **Feasibility**). Should we not recruit enough participants with RD from The Reading Group center, alternate possibilities are the University of Illinois Speech-Language Clinic, or our Office of School-University Research Relations (which assists in identifying schools and other community agencies willing to participate), and lab website postings or newspaper announcements. We will keep track of the balance of race/ethnicity, dialect, and family socioeconomic status through our intake questionnaire as we recruit each cohort, and do our best to match the RD and RC groups, and our three cohorts. Minorities tend to be over represented in RD children. We also will monitor children during longer tasks, to prevent children from becoming discouraged with the task. One alternative for design of the Memory task is to use a string of four triphones (CCV, VCC, CVC) rather than four bisyllables. Another potential problem is seasonal colds and allergies that might affect hearing. Our strategy is to reschedule sessions if a child shows acute symptoms, and re-screen hearing before sessions resume. Last is the possibility that participants may choose not to complete the entire experiment. In our design, all three experiments begin with the SCO/NSCM task. Therefore, as long as the child completes the first eight sessions, we will have usable data for some of our PP analyses.

Benchmarks for Success (C1 = Cohort/Experiment 1, etc.)	Year 1 (in quarters)				Year 2 (in quarters)			
	Qt 1	Qt 2	Qt 3	Qt 4	Qt 1	Qt 2	Qt 3	Qt 4
Software development	C1	C2,3	C2,3	C2,3				
Recruitment, Assessment Battery	C1	C1	C1	C2	C2,3	C2,3	C2,3	
Data Collection	C1	C1	C1	C1,2	C2,3	C2,3	C2,3	
Data Processing			C1	C1	C2,3	C2,3	C2,3	C2,3
Analysis				C1	C1	C2,3	C2,3	
SCO/NSCM All Cohort Analysis								C1-3
Dissemination								C1-3

Table 1: Benchmarks for Success–Timetable.