NOTE (June 25, 2015)

The research described here was funded as NIH grant R01 HD040353 from 2006 through 2011. Data generated by the project is still working its way through to publication. Additional information on the project and its antecedent can be found at:
http://projectreporter.nih.gov/project_info_details.cfm?aid=7797627&icde=22844820

Key Personnel:

Primary Investigator: Donald Shankweiler

Investigator: David Braze

Investigator: James Magnuson

Investigator: Einar Mencl

Investigator: Whitney Tabor

Investigator: Julie Van Dyke

A. Specific Aims

In our initial project (hereafter Project 1) our goal was to provide an account for an overriding fact about reading: that comprehension of print is much more difficult than comprehension of speech for a large sector of the population, including many of the young adults we studied. Our earlier research had identified printed word decoding skill as the key factor responsible for the gap between comprehension by eye and by ear. Indeed, the Simple View of Reading (Gough & Tunmer, 1986) proposes that this factor and general language ability represented by listening comprehension together account for all of the variance in reading comprehension. As to the mechanism, we and others have maintained that insufficient phonologic decoding skill at the word level creates a bottleneck that impedes integrative processes in reading sentences, i.e. the Phonological Bottleneck Hypothesis (PBH; Hogaboam & Perfetti, 1975; Shankweiler & Crain, 1986). Although our findings with a diverse group of young adults confirm the importance of decoding skill and listening comprehension, a finding in Project 1 suggests that the theory needs to be extended: Spoken vocabulary knowledge captured significant unique variance in reading comprehension beyond that captured by the combination of decoding ability and listening comprehension. We have implemented a hierarchical, activation-based model which explicates this finding and which will guide us in examining an implication of the Lexical Quality Hypothesis (LQH; Perfetti & Hart, 2002) that strength of interconnections within the lexicon can partially compensate for weakness in decoding (Braze, Tabor, Shankweiler & Mencl, submitted). In the continuation, we will use this model to deepen our understanding of reading difficulties of non-college bound young people at several levels of linguistic structure. First, we plan to replicate and extend our current findings about the role of vocabulary in reading ability, using innovative methods that assess the time course of the availability of lexical information. Then we will test the prediction that limitations originating at the lexical level produce sentence level reader differences, especially when demands on the syntactic processor are high. Additionally, we will explore reader differences in the ability to integrate lexical and inter-lexical information, as well as information from visual context, into a coherent meaning representation. Finally, we extend the investigation begun in Project 1 on the contribution of sequence learning to reading difficulty by examining reader differences in learning new lexical representations, focusing on morphosyntactic compounds which link lexical and syntactic knowledge.

Aim 1 is to investigate the link between the quality of an individual's word knowledge and the ability to quickly access that knowledge during on-line speech and print comprehension tasks (Experiments A, B, C, D). Both current linguistic theory, in which syntax is lexically driven, and connectionist-based approaches to sentence processing, predict such a link, envisioning a strong influence of lexical quality on parsing.

Aim 2 is to investigate sentence processing in relation to reading-related differences in lexical processes (Experiments F, G, H). The PBH predicts that difficulties at the lexical level will accentuate difficulties at the sentence level. We will look for evidence in measures of processing speed and brain circuitry, focusing on complex grammatical constructions. We expect that poor readers will show greater success with difficult syntactic structures when the demands of orthographic processing are removed through auditory presentation. To this end, we extend our Project 1 program of contrasting comprehension in print and speech modalities.

Aim 3 is to identify the learning mechanisms, and the relevant brain patterns, that underlie individual differences, both lexical and syntactic, focusing on bits of language that are larger than words and smaller than sentences (Experiments I and J). We suspect that learning routinized sequences of familiar words is a stumbling block for many less-skilled readers. The aim of a second learning study is to test a further prediction of the PBH: Because poor readers have less well-specified phonological representations, they should have greater difficulty than good readers in adding words to dense lexical neighborhoods.

B. Background and Significance

Studies of reading have allowed great gains in understanding the processes, and the relevant brain circuitry, by which a reader gets to the word from its encipherment in print, and the sources of the enormous individual differences that characterize the course and outcome of reading acquisition. Fluent decoding skill depends on phonological awareness (Bradley & Bryant, 1983; Liberman, Shankweiler, & Liberman, 1989; Shankweiler et al., 1999). Poor phonological awareness correlates with other difficulties of phonological processing, including slow object and symbol naming and difficulty in encoding and retaining phonological patterns of new words and word sequences (Brady, 1997; Wolf, 1991; Wolf & Bowers, 1999).

Neuroimaging findings have provided additional confirmation and neurobiologic grounding of these findings. Studies by the group at Haskins Laboratories and Yale have shown that reading-disabled and non-impaired readers differ most at brain regions that are centrally implicated in decoding letter patterns and printed words into phonology (Pugh, Mencl, Jenner et al., 2000; Pugh, Mencl, Shaywitz et al., 2000; Shaywitz et al., 1998). Reading difficulty in adults and older children is marked by low levels of activity (relative to those of non-impaired readers) in two left posterior cortical regions (the inferior temporo-occipital region and the superior temporal area and adjoining portions of the parietal lobe). In addition, disabled readers show comparatively elevated levels of activity in inferior frontal region bilaterally and in the right posterior region. Greater activity in frontal regions has been widely interpreted to reflect disabled readers' over-reliance on articulatory coding in the absence of well-automated word recognition procedures. These findings have been confirmed and extended in recent studies (e.g., Sandak et al., 2004) showing changes in brain activity in response to one or another aspect of word learning. Little information previously existed about reader ability differences in brain circuits critical for sentence processing. One of the goals of Project 1 was to acquire such information (C1.4, 1.5), and this is also reflected in the continuation proposal.

Our earlier studies of reading and related abilities in high school and college students show that failure to acquire accuracy and fluency in word recognition and poor non-word reading are characteristic not only of dyslexics, but also of many individuals exhibiting less extreme reading difficulty (Cunningham, Stanovich, & Wilson, 1990; Lundquist, 2003; Shankweiler, Lundquist, Dreyer, & Dickinson, 1996). Although word decoding is a persistent source of difficulty for poor readers up to adulthood, older poor readers suffer from a wider range of verbal deficits (Fowler & Scarborough, 1993), the cumulative result of poorly established literacy skills. One consequence is a widening gap between comprehension of speech and print: adult poor readers often fail to comprehend printed material that they could readily comprehend by ear (Stanovich, 1986, 1991).

In Project 1 we have been studying reading comprehension and its supporting abilities in young adults representing a wide range of reading ability, using conventional educational and cognitive measures and obtaining online experimental behavioral measures and neuroimaging measures based on parallel studies in speech and print modalities. As in much earlier work with younger readers (Foorman, Francis, Shaywitz, Shaywitz, & Fletcher, 1997; Shankweiler et al., 1995; Shankweiler et al., 1999; Shaywitz et al., 2003), we are finding substantial correlations between skill in reading isolated words and nonwords, and ability to read connected material with comprehension, confirming the existence of a bottleneck in these readers.

A parsimonious form of the bottleneck hypothesis can be stated in terms of regression. It has been claimed that if word decoding and listening comprehension are measured appropriately, their product accounts for all of the variance in reading comprehension (i.e., the Simple Model; Gough and Tunmer, 1986). Thus, if listening comprehension is entered first into a regression analysis, decoding accounts for all of the remaining variance in reading. In this spare model, listening comprehension subsumes the many variables reflecting spoken language, and decoding subsumes the various operations required in getting from print to phonology. However, in Project 1 we found that *three* measures were required to get all the reliable variance in reading comprehension; in addition to decoding and listening comprehension, orally-assessed vocabulary knowledge is also a unique contributor. Although it is possible that the additional contribution of spoken vocabulary reflects measurement problems, we will explain later why we claim that vocabulary knowledge contributes more weight to reading comprehension than to listening comprehension. Decoding problems may therefore not be the only source of the bottleneck in reading connected material in adult readers, as Cunningham, Stanovich, and Wilson (1990) have also maintained.

The influence of separately measured word knowledge is a phenomenon deserving further study in order to understand the difficulty associated with incorporating known, accurately decoded printed words into sentence schemas. To more fully reveal the contribution of vocabulary to this equation, we have implemented a model of the Phonological Bottleneck Hypothesis (PBH), described in Section C1.3, which incorporates this empirical result and makes specific predictions that we intend to test in the current continuation. In particular, this model suggests that once accurate phonological representations are obtained, poor comprehension may still arise because of inferior lexical semantic representations. While on the surface this conclusion is trivial, what is not trivial is the specific prediction that weak semantic representations will affect *how quickly* semantic information becomes available on line. This relationship is key to understanding the etiology of poor reading comprehension because efficient comprehension requires that all components of word knowledge become available *at the right time* for integration (Perfetti, 1985, 1990).

This is the central claim of Perfetti and Hart's Lexical Quality Hypothesis. Many problems in reading comprehension are caused by incomplete word identification processes relating to the timely access of orthographic, phonologic, and semantic aspects (Perfetti, 1992, 1985; Perfetti & Hart, 2002). The current proposal seeks to investigate the role of vocabulary knowledge using methods that allow measures of the quality of lexical representations and the time course of their access. We will test the prediction that our adult poor readers have low quality lexical representations and thus slower access to words, difficult integration across words, and poor comprehension even in the auditory modality, when deficiencies related to decoding skill are removed. In addition, our model explains working memory effects in sentence processing in terms of similarity-based interference. We predict that poor readers will be less susceptible to phonologic and semantic interference due to their weaker lexical representations. Finally, our model predicts that a poor lexical representation will affect readers' ability to encode new sequences that incorporate that lexical representation. The proposed experiments will fill a gap in understanding the role of lexical quality as a variable underlying reading skill and a major determinant of comprehension differences.

Target Population. Most past studies of sentence processing have focused on university students. We have widened the base by focusing our sample on non-degree community college students and adult education students in order to make a realistic appraisal of the range of individual differences and their effects, obtaining badly needed information about literacy skills of non-university bound young people, a large and neglected sector of the population.

Methodologies. Research on comprehension has suffered from reliance on gross measures that are not diagnostic. An important aspect of our approach is to use sentence materials as well as longer texts (sentence length materials afford greater possibilities for control than longer texts), and to use on-line measures to study comprehension processes while they are in progress. Eye tracking offers a promising method for studying skill differences in reading that has been under-exploited. Monitoring eye movements during reading provides a running record well suited for tapping the temporal structure of sentence processing, and can reveal, with greater precision than other behavioral techniques, locations within the sentence where the processor becomes derailed, charting its attempts at recovery. We have new evidence that our unskilled readers respond differently than college students in eye-movement measures of reading sentences containing anomalies of grammatical form and meaning. Coordinating these measures with neuroimaging findings offers a rare opportunity to assess the associated patterns of brain activity during the same cognitive task. Our strategy has been to design dual eyetracking and fMRI experiments, based on controlled comparisons of reading and listening, to investigate differences in brain circuitry related to type of sentence structure, input modality, and reader skill (see Sec. C). We continue this approach in the current proposal. Another distinguishing feature is a focus on continuous, regression-based analyses which make use of the full range of variability along individual difference continua. Follow-up tests will employ discrete classification into good/poor reader groups in order to check linearity of these effects, and simplify interpretation of some experiments.

New to this proposal is the Speed-Accuracy Tradeoff (SAT) procedure, which is uniquely suited to assess effects of lexical quality on retrieval speed, as it measures how the confluence of speed and accuracy jointly determine comprehension. This is an advance because speed and accuracy of word (and higher level) processing are continuous variables, with large variability expected in each measure. Moreover, traditional measures of reaction time or reading speed can result from a confluence of factors, only some of which reflect true differences in processing speed (McElree, 1993; McElree & Griffith, 1995, 1998). Consequently, a record of how accuracy changes as a function of response time will provide more information about the source of individual differences than correlations of individual measures with some external measure of comprehension.

Practical Goals. The planned studies are relevant to two important practical goals. First, they will elucidate the mechanisms for promoting reading effectiveness for further education and job training within this sector of the population. Secondly, our inquiry is relevant to the task of creating more effective learning environments for the low-functioning secondary student. Are these goals best advanced by sharpening decoding and spelling skills, or should vocabulary building also be stressed? Much effort has been directed to development of basic literacy programs, but less is known about the factors that promote or retard advancement from the basic skill level to more advanced levels compatible with the adult learner's listening comprehension.

C. Progress Report and Preliminary Studies

C1. Progress Report

C1.1 Summary of progress based on the Project 1 period, which began March 1, 2002. Project 1 had three specific aims. **Aim 1** asked whether the sources of sentence comprehension difficulties in this population, either in speech or in print, reflect difficulties in word recognition mechanisms specifically or whether the difficulties reflect more general sequence learning problems. In sections C1.2 - C2.5 we discuss the pertinent results from Project 1. **Aim 2** explored the consequences that individual differences in phonological processing ability have for sentence comprehension (also addressed in sections C1.4-C1.5). **Aim 3** asked whether sequence learning problems give a better account of individual differences in sentence comprehension than the Phonological Bottleneck Hypothesis. In our work so far, we found sequence learning differences between less and more experienced readers (C2.5). Due to limitations of time and budget, a deeper investigation of this aim has been deferred to the continuation project, in which our ***new*** **Aim 3** proposes to extend our original Experiment 5, on learning sequences of words smaller than the sentence. All three specific aims of the new proposal represent extensions of research on these issues. One extension is based on the finding that phonologic decoding differences together with listening comprehension did not explain all the variance in sentence level reading. Problems with vocabulary learning and their consequences need to be better understood. Hence we propose an intensive study of lexical knowledge and lexical processing with new predictions based on a partially implemented activation-based model developed in Project 1. The model regards quality of lexical encoding as the principal source of skill differences.

The participants in Project 1 were young people, aged 16-24 years. In keeping with our intention of studying people from poor urban neighborhoods, including those whose schooling had been interrupted, we recruited from adult education centers, from community colleges, and high schools serving these neighborhoods. Many who volunteered were seeking a high school equivalency certificate or resuming regular high school classes. Few would have met the requirements for a four-year college. Our project is unique in being based on a diverse sample, reflecting a range of abilities more nearly representative of the age group than most previous reading studies of young adults, which typically have been based on university students.

Regardless of educational background, those selected to participate in Project 1 had to be capable of reading simple material with understanding in order to perform our reading tasks. We used the Fast Reading subtest of the Stanford Diagnostic Battery as a screening measure, requiring an accuracy of at least 70% correct of the items attempted. This cutoff excludes some severely reading-disabled individuals, while admitting others who are fairly accurate but slow. Screening also ensured that all participants learned English as their first language and that they had a minimum estimated WASI IQ of 80. All those selected to participate received ancillary educational and cognitive tests, organized into the following groupings: a) Print Mapping and Reading Skills; b) Oral Language Measures; c) Mental Facility and Speed; d) Print Familiarity.

To characterize our lower-skilled readers relative to norms, we examined the lower half of the group separately on our screening (reading comprehension) measure. Word-level reading skills as well as comprehension are low in these readers compared to a 12th grade benchmark; notably, their listening comprehension grade equivalent scores are substantially ahead of their reading comprehension scores. Reading speeds were slow, based both on the scores on a standardized test of oral reading/fluency, the Gray Oral Reading Test (Widerhold & Bryant, 2001), and on results from the eye-tracking studies. They showed wide variability in measures of sentence span and many had poor decoding skills and poor phoneme awareness. They also showed low levels of estimated print experience based on their responses to author recognition and magazine recognition checklists (Stanovich & Cunningham, 1992).

An eye-movement facility to study reading comprehension was created at Haskins Laboratories at the beginning of Project 1. Software developed by members of the project provides a variety of dependent measures needed to assess eye-movement patterns across sentence regions, including first pass fixation time, incidence of first pass regressive eye movements, and regression path targets. Participants who met the selection criteria and who completed the ancillary tests took part in a series of double experiments; each experiment had an eye-tracking sentence processing part, and a coordinated neuroimaging part. Data collection for Experiments 1 and 2 is complete. A paper discussing the behavioral-cognitive findings from Exp. 1, Syntactic and pragmatic anomaly detection (N = 44), has been reviewed favorably and is currently under revision (Braze, Mencl, Tabor, & Shankweiler, 2005; Braze, Shankweiler, & Tabor, 2004; Braze, Tabor, Shankweiler,& Mencl, submitted). These findings, the eye-movement findings, and a first-pass analysis of the fMRI data (each discussed below) have each been presented to professional groups and conferences (Braze et al., 2004; Van Dyke, Shankweiler, & Tabor, 2005). For Experiment 2: Resolving ambiguities: syntactic and pragmatic effects, data analysis is underway. Prior to conducting this experiment, pilot work using word-by-word self-paced reading was conducted under subcontract at the University of Connecticut. This work contributed to the development of a self-organizing parsing model (Tabor, Galantucci, & Richardson, 2004; Tabor & Hutchins, 2004a), also conducted under subcontract, which forms part of the hierarchical activation model described below. In preparation for Experiment 3: Crossing word decoding and sentence parsing complexity, we carried out a systematic analysis of the Celex database to characterize the decoding complexity of multi-syllable words. In preparation for Experiment 4, Processing idioms as exceptionally-learned sequences, we compiled a database of suitable idioms, rated by their transformability; we also devised and are piloting a questionnaire to establish which idioms are familiar to the people in our participant pool. A pilot version of the behavioral portion of Experiment 5 (Novel phrase learning and reading difficulty) was run and the results are reported below (N=56).

C1.2 New Result: Vocabulary knowledge captures unique variance in reading comprehension. Reading comprehension showed high correlations with the measures of vocabulary, listening comprehension, reading isolated words, print exposure, and the sentence span measure of working memory. Some nonlinguistic measures (e.g., WASI Matrices and Corsi Blocks) also yielded significant correlations, but none contributed to the most successful regression model for predicting reading comprehension (Braze et al., 2005; Braze et al., submitted).

We adopted the Simple View of Reading (Gough & Tunmer, 1986) as a point of departure for regression analyses aimed at predicting reading comprehension. This theory separates the variables governing reading skill into two groups: One group, called D (decoding), comprises the orthographic-to-phonologic mapping skills that are needed to recognize printed words and build a sight vocabulary. The other group includes the many abilities that reading shares with spoken language: vocabulary, syntax, semantics, pragmatics. Collectively, these are called L; in principle, they can be assessed by measures of spoken language comprehension independently of reading. The Simple View states that these two factors account for all of the variance in reading comprehension. Based on work with younger readers (Hoover & Gough, 1990; Shankweiler et al., 1999), we expected to confirm the theory, and we did, with one proviso. Regression analyses based on the cognitive and skills measures for our diverse sample of young adults showed that two measures, decoding skill and listening comprehension, together accounted for 73% of the variance in a composite measure of reading comprehension.

However, a third measure, oral vocabulary, adds another 8% of unique variance, improving prediction non-trivially (Braze et al, submitted; Braze et al., 2005). To check the generality of this result, we evaluated all possible regression models that could be computed from our battery of verbal and nonverbal measures with the proviso that listening comprehension be included. No other model captured as much variance. Thus, we concluded that listening comprehension, narrowly construed as ability to understand sentences and short narratives, does not capture all the variance in reading comprehension beyond decoding. It is still possible that the additional contribution of spoken vocabulary reflects measurement problems, such as could occur if our test of listening comprehension (based on the PIAT-R) and our measures of vocabulary (PPVT and WASI) sampled different subsets of the vocabulary. It is unlikely, however, that this can explain our result. Substantially the same outcome of a regression analysis is obtained when Listening Comprehension and Reading Comprehension measures are each derived exclusively from items drawn from the same test (the PIAT-R). Further, vocabulary quality contributes more to comprehension of print than of speech. New Aim 1 seeks to confirm this result and extend it with focused investigations of selected words, guided by the Lexical Quality Hypothesis.

C1.3 An activation-based model of the Phonological Bottleneck Hypothesis (PBH). We are presently implementing the PBH as a continuous-activation connectionist model which is a version of a Hopfield Network (Hopfield, 1982) and is similar in concept to recurrent network models like that of Kawamoto (1993), Harm & Seidenberg (2001), Plaut, McClelland, Seidenberg, & Patterson (1996). Words are represented by distributed patterns of activation representing fine-grained phonological, syntactic, and semantic features. We focus initially on a subset of the model (shown in Figure 1), the mapping from phonology to semantics, which is one edge of what is called the "Triangle Model" (Seidenberg & McClelland, 1989). The weights between units in the model are tuned such that each lexeme corresponds to a stable, attractive state of the network. This means that one can activate some features of a word (e.g., its phonological features) and the network will tend to turn on other features associated with that word (e.g., its semantic features). As Hopfield noted, one can think of such a model as traveling on an irregular energy landscape as shown in Figure 2. Activating a subset of a word's features corresponds to starting on the side of a basin in the energy landscape. The model's response, which is to gradually activate the other features, corresponds to sliding down the side of the basin. Being at the bottom of the basin corresponds to fully understanding the word. The Lexical Quality Hypothesis posits that good readers have deeper and steeper basins, while poor readers' basins are shallower and have more gentle slopes. The hypothesis fits with the assumption that good readers have had more opportunity to adjust the weights between units on the basis of experience, resulting in higher quality lexical representations.

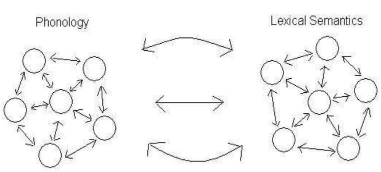

Figure 1

Explaining how vocabulary interacts with reading ability. In keeping with the Lexical Quality Hypothesis, poor readers should be slower than good readers at achieving full understanding of a word presented orally. This follows because the dimples in their energy landscapes are less steep so the system settles more slowly into the appropriate basin. Poor readers should also misidentify more words. Since they spend more time near the tops of the basins, noise is likely to knock the system into a nearby basin. Nearby basins, or "neighbors," correspond either to linguistic interpretations with similar features (e.g., reading "cap" as CAT, or "cot" as BED) or they correspond to what are called "spurious attractors" -- basins that do not correspond to any real interpretation. Falling into a spurious basin corresponds to a failure of comprehension. Spurious attractor basins correspond to partially coherent linguistic input, for example pseudowords. In Braze et al. (submitted) we combine the piece of the model above with other groups of units mediating auditory input, orthographic input, and syntactic processing. The resulting model predicts trade-off effects: e.g., shallow attractors in the orthography-

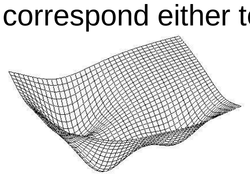
Figure 2

phonology circuit can be compensated for by deeper attractors in the lexical-semantic group. This assumption predicts our Project 1 result that vocabulary ability picks up significant additional variance beyond that accounted for by Decoding and Listening Comprehension.

Extending from words to sentences. Our work in Project 1 on the self-organizing parser model (Tabor et al., 2004; Tabor & Hutchins, 2004a) extends the energy landscape framework to the sentence level by assuming that basins correspond to grammatically valid parse trees (or subtrees). In this system, the perception of each successive word activates a syntactic subtree with open attachment sites bearing syntactic and semantic features. These act as basins, attracting all other subtrees to create potential syntactic links. Links with well-matching featural specifications dominate the competition, and if the emerging linked structure is a well-formed syntactic tree, the sentence is successfully interpreted. This model predicts the sentence-level interference effects found in Van Dyke and Lewis (2003) and Tabor et al. (2004) (see Section C2.1).

This implementation illuminates the critical link between the word and sentence levels described by the Phonological Bottleneck Hypothesis: shallow basins in the orthography-to-phonology circuit will slow the activation of syntactic and semantic nodes crucial to the formation of tree structures. Thus, decoding difficulty hampers the formation of sentence interpretations and we predict that poor readers should be slower, more prone to error, and less sensitive to syntactic and semantic anomaly than good readers. These predictions are largely borne out by our Project 1 findings (see Section C1.3) and form the motivation for our new Specific Aims 1 and 2.

C1.4 New Result: Comparing brain activity during sentence processing in speech and print. An fMRI experiment (Constable et al., 2004), comparing sentence processing by speech and reading, was completed during the Project 1 period, and provides essential background data for all our experiments using this methodology. Controlled listening/reading comparisons are essential for interpreting reading level differences. Our goal in Constable et al. was to map the cortical regions engaged by sentence processing by skilled readers to assess the influence of input modality (speech or print) and parsing difficulty (sentences containing easier subject-relative or harder object-relative clauses). Auditory presentation was associated with pronounced activity in the region surrounding the primary auditory cortex and across the superior temporal gyrus bilaterally. Printed sentences, by contrast, evoked major activity at several posterior sites in the left hemisphere, including the angular, supramarginal, and the fusiform gyri in the occipito-temporal region. In addition, modality independent regions were isolated, with greatest overlap seen in the inferior frontal gyrus (IFG; also observed by Michael, Keller, Carpenter, & Just, 2001). With respect to sentence complexity, object relative sentences evoked heightened responses in comparison to subject relative at several left-hemisphere sites, including IFG, middle/superior temporal gyrus, and the angular gyrus. These sites showed modulation of activity as a function of sentence complexity, independent of input mode, arguably form the core of a cortical system essential to sentence parsing, providing a unitary framework within which information from each modality can be incorporated into a common currency. Our subsequent studies based on other sentence structures and with readers differing in skill have incorporated parallel speech and print conditions. The results (see C1.5) support our conjecture that strongly overlapping areas of response to sentence processing in speech and print is characteristic only of the more skilled readers.

C1.5 New Result: Individual differences in sentence processing in speech and print. A central question in examining the relationship between language and the brain is whether and to what extent different aspects of language processing can be separated. The distinction between syntactic form and pragmatic interpretation is usually considered to be fundamental. Earlier research from our group has studied how this distinction is manifested in patterns of eye movements during reading and in patterns of brain activity. We regard sentence anomalies as a good tool for evoking specific processes, dependent upon the anomaly type. Accordingly, we use sentences containing syntactic and pragmatic anomalies to probe for separate markers of syntactic and pragmatic levels of processing. Within our group, several lines of research using eyetracking, and BOLD signal recording indicate that the distinction has objective manifestations that are revealed in the time course of sentence processing, regressive eye movements, and activation of specific cortical sites.

The base stimulus set consists of sentences that are either syntactically anomalous (1.a), pragmatically anomalous (1.b), or non-anomalous (1.c):

1a The daisies will slowly wilting in the hot weather this afternoon.
1b The puddles were slowly wilting in the hot weather this afternoon.
1c The daisies were slowly wilting in the hot weather this afternoon.

In a series of eye-tracking studies with college students students (Braze, 2002; Braze, Shankweiler, Ni, & Palumbo, 2002; Braze et al., 2004; Fodor, Ni, Crain, & Shankweiler, 1996; Ni, Fodor, Crain, & Shankweiler, 1998), each type of anomaly triggered a distinct pattern of eye movements, which differed depending on the type of anomaly. Sentences containing syntactic anomalies caused abrupt increases in eye regressions at the anomalous verb before a quick return to normal reading, whereas sentences with pragmatic anomalies resulted in perturbations that increased gradually and peaked at the end of the sentence. These findings suggest that distinguishably different cognitive processes are associated with syntactic and pragmatic operations, and is consistent with other results (Boland, 1997; McElree & Griffith, 1995, 1998).

Reader differences had been little studied with materials such as these. The PBH leads us to expect that reader differences will be magnified in sentences containing pragmatic anomalies because their detection requires an open-ended search through the world knowledge base, a demanding task when working memory is already under stress due to difficulties with word decoding. In contrast, syntactic anomaly requires only a domain-limited query of the language system or a portion of it. We extended this eye-tracking study in Project 1, and found that our diverse sample of young adult readers also differs in their responses to the two kinds of sentence anomalies, partially confirming our earlier findings based on more homogeneous college student samples. Moreover, there are clear individual differences correlated with reading skill: lower-skilled readers made fewer regressive eye movements from the syntactically anomalous verb than higher skilled, suggesting that they do not consistently attend to the anomaly (See Fig.3). This is confirmed in the case of pragmatic anomaly where low-skilled readers also differ from normal college students in failing to show the end-of-sentence regressions, a signature effect of pragmatic anomaly (Braze et al., 2004).

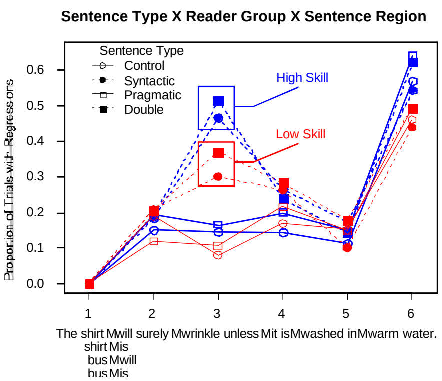

Figure 3

<u>Neurobehavioral findings.</u> We have used event-related fMRI successfully to identify cortical events in online anomaly detection (Ni et al., 2000). In a study of normal young adult readers, the anomalous sentences were presented as rare “oddballs.” Our purpose was to find out (1) whether critical cortical sites registered isolated occurrences even when no detection response was required (subjects judged whether each sentence contained a living thing), and (2) whether there were distinct cortical sites for processing syntactic and pragmatic anomalies. Unlike earlier studies using the block-subtraction method and requiring explicit detection responses, the event-related, single-trial method clearly teased apart cortical manifestations of the two anomaly types. Syntactic anomaly significantly increased activity in and around posterior inferior frontal gyrus (IFG) whereas pragmatic anomaly activated several additional sites, including the posterior, superior temporal gyrus (STG). The two activation patterns were largely non-overlapping. The dissociation between syntactic and pragmatic processing conforms to indications from eye-movement studies discussed above, as well as ERP (Hagoort, Brown, & Osterhout, 1999; Kutas & Hillyard, 1983; Osterhout, 1997; Osterhout & Holcomb, 1992) and fMRI findings from other laboratories (Friederici, Ruschemeyer, Hahne, Fiebach, & Friederici, 2003 ; Helenius, Salmelin, Service, & Connolly, 1998).

In Project 1, we investigated brain responses to syntactic and pragmatic anomaly in speech and reading in our young adult sample (N=42) using anomalous and corresponding non-anomalous sentences similar to 1.a, 1.b and 1.c. In accord with Ni et al. (2000), syntactic and pragmatic anomalies have distinct effects on the BOLD signal. For the speech condition, syntactic anomaly minus nonanomalous control sentences activates IFG strongly, as Ni et al. showed. STG is also clearly active for syntactic anomaly, which was not observed in the Ni et al study, perhaps because of reduced power (but was reported by Friederici et al., 2003). The print version of the syntactically anomalous sentences yielded robust frontal activity, diffusely in IFG and spreading to the insula. STG is also activated by the print version, more so than in speech.

Skill differences are clearly apparent in our data. Our results for higher skilled readers confirm the findings of Ni et al in showing a significant increase for pragmatic anomalies minus control sentences in the posterior temporal lobe (MTG). Less skilled readers did not show significant activation in this region, in keeping with the findings of Helenius et al (1999) with another imaging methodology (MEG) and another language. Moreover, the low level of responsiveness of poor readers in posterior cortical regions is a recurrent feature in earlier fMRI studies based on tasks taxing phonologic processing with isolated word stimuli (Pugh, Mencl, Jenner et al., 2000; S. E. Shaywitz et al., 1998).

Reader skill differences in brain activity are even more striking in response to syntactic anomaly. Figure 4 shows a representative slice showing activity generated in reading syntactically anomalous minus control sentences by higher skill (left) and lower skill readers (right). Higher skill readers show activity predominantly in the left hemisphere in IFG (area 44), insula, and STG,

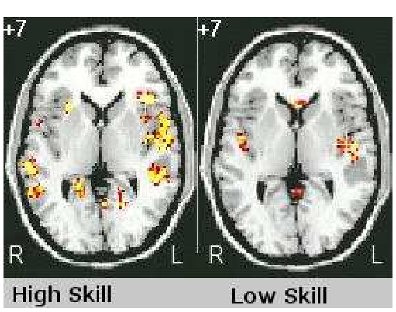

Figure 4

anteriorly and posteriorly. In the right hemisphere, smaller activations were observed in the STG and insula. Less skilled readers show sharply less activity overall in response to syntactic anomaly than more skilled readers. Direct comparison of higher and lower skilled shows a significant interaction in the most inferior aspect of the IFG and insula. The differences associated with skill level are not confined to reading; differences implicating many of the same regions are found also in response to speech. The analysis of imaging data from Project 1 is ongoing.

C2. Additional Motivation for the Current Proposal

<u>C2.1 Interference effects in poor readers and the time course of information availability.</u> As noted in Section C1.3, the Phonological Bottleneck Hypothesis suggests that poor readers will have difficulty with syntactically complex sentences not because of the complex syntax as such, but because working memory is overtaxed by decoding and lacks sufficient capacity to handle the syntax (Crain, Ni, & Shankweiler, 2001; Shankweiler & Crain, 1986). In support of this, we found that poor readers perform better on the same structures when processing load is reduced (Crain, Shankweiler, Macaruso, & Bar-Shalom, 1990; Macaruso, Bar-Shalom, Crain, & Shankweiler, 1989; Smith, Macaruso, Shankweiler, & Crain, 1989). Even when the words can be read correctly, poor readers' difficulties are expected to be exacerbated when they read sentences, compared to listening to them, a prediction under test in our current Project (See Section C1.)

Our limited-resource hypothesis is consistent with much research focusing on the storage demands required to process complex sentences, which contain multiple grammatical dependencies, and which strain the memory of even good readers (e.g., Caplan & Waters, 1999; Chomsky & Miller, 1963; Gibson, 1998, 2000). However, recent work has suggested that, at least in skilled readers, some of these complexity effects may be due to properties of retrieval. Specifically, interference effects have been demonstrated at the level of referential category (Gordon, Hendrick, & Johnson, 2001, 2004), grammatical structure (Kaan & Vasic, 2004; Van Dyke, submitted; Van Dyke & Lewis, 2003), and semantics (Van Dyke, submitted). In concert with global-matching models of memory retrieval (e.g., SAM, Gillund & Shiffrin, 1984; MINERVA2, Hintzman, 1984), Van Dyke and Lewis (2003; Van Dyke, 2002) have argued that these effects result from an associative, direct-access retrieval mechanism that uses cues to identify to-be-retrieved items. For example, Van Dyke (2002) found that sentences with a subject intervening between the dependent and its verb (i.e., sentences with syntactic interference, as 2a&b) were more difficult to process than sentences with the same number of intervening words, but without an intervening subject (2c&d). Moreover, when the intervening noun matched the semantic requirements of the verb, (i.e., semantically interfering, as in 2b&d), the effect was increased, even when that noun did not fit the grammatical properties required by the verb (2d). Van Dyke (submitted) suggested that these results reflect the ambiguous retrieval cues supplied by the verb, which cannot uniquely identify its correct subject. Notably, these interference effects are not predicted by most accounts of complexity effects, which focus on storage limitations as the source of processing breakdown, since the memory load required to process each sentence is identical.

2a. (high syntactic, low semantic interference) The <u>assistant</u> forgot that the client who implied that the meeting was important <u>was waiting</u> for a signature.

2b. (high syntactic, high semantic interference) The <u>assistant</u> forgot that the client who implied that the visitor was important <u>was waiting</u> for a signature.

2c. (low syntactic, low semantic interference): The <u>assistant</u> forgot that the client who had asked about the important meeting <u>was waiting</u> for a signature.

2d. (low syntactic, high semantic interference): The <u>assistant</u> forgot that the client who had asked about the important visitor <u>was waiting</u> for a signature.

Reader differences in memory retrieval and sensitivity to syntactic or semantic interference have so far not been investigated, although our model does account for these effects. As noted in Section C1.3, we expect less-skilled readers to have shallower basins, corresponding to lexical representations that are less well-defined than those of more-skilled readers. This leads to the counter-intuitive prediction that skilled readers should be more hampered by irrelevant similarity (i.e. interference) between words in a sentence than less-skilled readers. This is because high similarity between words will cause readers to fall into spurious basins, as described above. Since their basins are deeper, skilled readers will have more difficulty getting out of them, giving rise to more pronounced and sustained interference effects. This means that the model will stabilize more quickly for a good reader whose basins are deep. Therefore, if irrelevant similarity is present, the model is more likely to settle into an incorrect state.

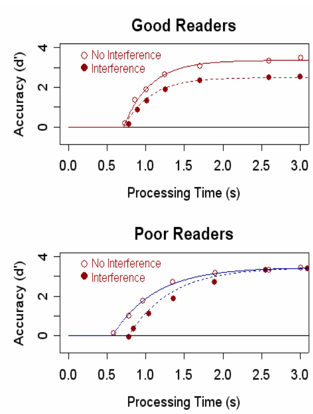

Figure 5

The current proposal aims to test these time course predictions using two complementary methodologies: the Speed-Accuracy Tradeoff (SAT) procedure and the Visual Worlds eye-tracking paradigm. The SAT procedure provides detailed time-course data of individual subject performance as accuracy increases from chance to asymptotic performance. The prediction that good readers will be hampered more quickly and to a greater extent by similarity between words corresponds to the SAT curve depicted in Figure 5, Top Panel. A single access speed (i.e. slope and intercept) results in two different asymptotes, corresponding to differences in response accuracy. Conversely, for poor readers the model will settle on a particular lexical item more slowly, even if they end up in the correct basin. In terms of the SAT response curve, this corresponds to Figure 5, Bottom Panel, which shows different access speeds resulting in the same asymptotic accuracy. The SAT experiments proposed in the current project aim at testing these predictions at both the word and sentence levels. Section D3 explains the experimental paradigm for obtaining these response curves in more detail.

Further tests of these time-course predictions will be gained through measuring eye movements as participants hear spoken instructions to move objects in a computer display. This *visual worlds paradigm* has been used to study both phonological and semantic aspects of the time course of spoken word recognition (Allopenna, Magnuson, & Tanenhaus, 1998; Tanenhaus, Magnuson, Dahan, & Chambers, 2000; Yee, Blumstein, & Sedivy, 2004). This will complement the SAT studies by providing fine-grained estimation of the time course of target activation as well as that of specific phonological and semantic neighbors. We will examine predictions about rate of lexical activation as a function of the quality of phonological and semantic representations that follow from our model of the PBH. Additional studies examine the temporal resolution of phonological representations by examining sensitivity to changes in coarticulatory cues (a more direct test of the quality of phonological representations), and the exact nature of weaknesses in lexical knowledge via familiarity ratings and lexical database analyses.

<u>C2.2 Individual differences in learning novel words and word sequences.</u> This module tests for reader differences in learning new words and new two-word sequences. Sandak et al. (2004) and Katz et al. (in press) found that training on novel phonological forms can produce differentiation in BOLD (fMRI) signal pattern and is consistent with prior work on the cortical locus of components of printed word recognition (Pugh, Mencl, Jenner et al., 2000). Experiment 5 of Project 1, "Novel phrase learning and reading difficulty," seeks parallel evidence on the learning of new sequences of familiar words. In a pilot for the fMRI study, university students read novel noun-noun sequences (e.g. DIME-YACHT) repeatedly during a training phase and then were tested for recognition of the items and their reversals (e.g., YACHT-DIME). An interaction between reading experience and the reversal manipulation was found: more experienced readers were less prone than less experienced to identify the reversals as items they had seen before, but the two groups performed similarly on recognition of non-reversed sequences. Further, additional studies find evidence for a transition from "holistic" (order-insensitive) to "incremental" (order-sensitive) encoding in children's learning of words (see Charles-Luce & Luce, 1990, for a review). Magnuson et al. (2003) produced a similar transition with adults in an artificial lexicon paradigm.

Our Hierarchical Activation model also predicts these results. If we assume, in keeping with work on sequence learning in connectionist models (e.g. Elman, 1990, 1991), that sequential information is stored by the same error-correcting mechanism by which pattern associations are stored, then we expect to see poor readers also showing difficulty in learning new sequences of familiar words. This is due in part to a bottleneck effect: because the poor readers' basins are shallow, they don't settle fully into them during the processing of the familiar words; this makes the inputs to the sequence learning less distinct, so the basins associated with the sequences end up being shallow as well. Consequently, poor readers are expected to be less sensitive to order contrasts than good readers. Moreover, noise will be more likely to bump the system into an erroneous basin, disrupting the learning of new words because they will be easily confused with existing words. This will happen most often in "dense neighborhoods" (Luce & Pisoni, 1998), where there are many competing basins nearby. These considerations provide the motivation for Specific Aim 3.

D. Research Design and Methods

D1. Overview of Research Plan

This project targets individuals in their late teens and early twenties including those whose limited reading abilities seriously restrict their educational and occupational prospects. All of those selected to participate will receive a series of tests to assess reading comprehension, listening comprehension, decoding skill and IQ. In addition to receiving a common set of background measures for use in regression and structural modeling (Exp. A), each individual will participate in one of 9 additional experimental studies, which fall into three categories. The first group (B – D) focuses on individual differences in word knowledge and its relationship to reading skill. The second group (E – H) focuses on sentence parsing and other comprehension processes involving word combinations. The final pair (I and J) focus on learning mechanisms that underlie sentence-level reader differences. The proposed research will build upon and extend the adolescent and adult literacy research initiated under HD-40353.

D2. Participants

In the renewal project, we would continue to recruit individuals, aged 16-24, from the New Haven area, who exceed reading comprehension and IQ cutoffs. We will recruit a total of 466 participants from urban high-schools, adult-education centers, and community colleges with which we have ongoing relationships. From our experience in Project 1, we expect to find large differences in reading skills in this population. The less skilled readers in the sample will have literacy limitations that are educationally significant. These individuals are often overlooked in favor of those who meet explicit criteria for learning disability.

Screening measures : Measures of reading comprehension and intelligence, and a brief questionnaire will be used for participant selection. The Fast-Reading subtest of the Stanford Diagnostic Reading Test (Karlsen & Gardner, 1985) will be used to screen potential participants in order to ensure that participants are not at the floor on experimental reading tasks. The Vocabulary and Matrix Reasoning subtests of the Wechsler Abbreviated Scale of Intelligence (WASI; The Psychological Corporation, 1999) will be used to estimate Full-Scale IQ. Participants must have an estimated Full Scale IQ of 80 or above. Extrapolating from Project 1, the expected range should be approximately 80-120. Research indicates that patterns of cognitive abilities that characterize reading differences are largely stable across those variations in IQ (Fletcher et al., 1994; Stanovich & Siegel, 1994). All will have English as their first language, as determined from responses to the questionnaire.

D3. Methods

D3.1 Skill and Experience Measures

All participants will receive the following tests, which yield independent measures of reading and related abilities. Standard regression analyses and sequential regression will be used to study the relations between these measures and the dependent measures of the experimental tasks (eye-tracking, SAT, and fMRI). Multiple indicators of most variables are included to provide stable estimates of underlying factors for use in the regression analyses and in structural equation modeling. We will calculate a composite reading comprehension score as the mean standardized score of the following three tests (listed below): (1) PIAT sentence comprehension subtest, (2) GORT, and (3) Woodcock-Johnson Passage Comprehension. This continuous score will be used as a predictor variable for primary analyses in all experiments. A categorical version of this score will also be created for use in some follow-up analyses, typically for a check against predictor nonlinearity, and to assist description and interpretation of interaction effects. In these secondary analyses, we will recode subjects in the top third of the sample for that experiment as "good readers", subjects in the bottom third as "poor readers", and subjects in the middle as "average readers." When this is applied as a between-subjects grouping factor in ANOVA, all three groups will be included, to retain the full sample size for main effects. For group effects and interactions, we will focus on planned comparisons between the good and poor groups only. In the sample assessed in the previous project period, the top and bottom thirds were separated by approximately one standard deviation in the composite score.

1. Normed Reading and Listening Comprehension: To assess the relation of reading comprehension to listening comprehension, it is necessary to have measures of each that are well-calibrated to one another. We accomplish this with 3 paired indices of comprehension: (a) Peabody Individual Achievement Test -- Revised (PIAT; Markwardt, 1998) sentence comprehension subtest(Leach, Scarborough, & Rescorla, 2003; Spring & French, 1990). Odd numbered items are administered in the standard way to assess reading comprehension, while even numbered items are presented orally to assess listening comprehension (Leach, Scarborough, & Rescorla, 2003; Spring & French, 1990); (b) parallel forms of the Gray Oral Reading Test Version 4 (GORT; Wiederholt & Bryant, 2001) will be used to assess comprehension in print and speech modes; (b) the Woodcock-Johnson Passage Comprehension and the Oral Comprehension subtests (Woodcock, McGrew, & Mather, 2001) will provide parallel measures of print and speech comprehension. These tests were selected to represent a range of materials – sentences, narrative and expository discourse – and a range of dependent measures.

2. Reading Fluency: GORT passages read aloud by participants (previous paragraph) will yield reading

fluency measures. The Woodcock-Johnson Reading Fluency subtest, a sentence reading task, will yield a second measure reading fluency (Woodcock, McGrew, & Mather, 2001).

3. Normed Word and Nonword Reading: The Word ID and Word-Attack sub-tests from the Woodcock-Johnson Psycho-educational Test Battery -- Revised (Woodcock, McGrew, & Mather, 2001) and the word and nonword reading tasks from the Test of Word Reading Efficiency (Torgeson, Wagner, & Rashotte, 1999) will be used to provide coordinated tests of isolated word and nonword reading skill. The Peabody Individual Achievement Test (PIAT) will be used as a third standardized test for word recognition (Dunn & Markwardt, 1970). A third test of nonword decoding skill, the Pseudohomophone Identification Test, is not normed but has been used extensively by our group. This test, based on Olson, Forsberg, Wise & Rack (1994) consists of 40 nonwords, presented by computer. Half of the items are pseudohomophones, like “brane”, and half, like “brone”, are similar to these, but are not homophonic with actual words. The task is to judge whether each item is a word homophone or not. Correct “yes” judgments indicate that access to the lexicon has occurred and that it was done by way of rule-based decoding.

4. Vocabulary: The tests were chosen to represent a diversity of tasks, including recognition and production, and diverse types of items. The Peabody Picture Vocabulary Test (Dunn & Dunn, 1997), WASI vocabulary sub-test (The Psychological Corporation, 1999), and the Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983) will be used to assess word knowledge. In addition, we include a Synonym Identification Task, which elicits judgments as to whether pairs of words are synonyms. This task is based on materials drawn from the “Kit of Factor Referenced Tasks” (Ekstrom, French, Harman, & Derman, 1976). Response measures for Synonym ID task include accuracies and reaction times.

5. Sentence Span Test: This measure of verbal working memory (Daneman & Carpenter, 1980) taps resources that are taxed variably in reading different kinds of sentences. We use an auditory version of the task to avoid confounding memory differences with reader skill differences.

6. Spatial span: We use a computerized version of the Corsi blocks task . Participants reproduce increasingly long visuo-spatial patterns by tapping on an irregular arrangement of nine circles displayed on a touch sensitive computer screen. The score is the longest sequence that can be reproduced successfully.

7. Print Familiarity: We assess familiarity with printed material using author and magazine checklists based on the work of Cunningham and Stanovich (1990; Stanovich & Cunningham, 1992). The participant has to distinguish actual authors (or magazine titles) from foils consisting of fictional authors (or titles). True positive and false negative responses are used to compute scores adjusted for guessing biases.

8. Phonological Sensitivity: The CTOPP battery (Wagner, Torgeson, & Rashotte, 1999) will be used to assess phonological awareness (segmentation and blending subtests), rapid naming, and phonological memory (nonword repetition and digit span subtests).

9. Articulation Rate: We will assess articulation rates for word pairs and for letter and digit sequences. In the word-pair task, participants repeat a series of 10 word pairs with increasing numbers of syllables (e.g., bear-pen; piano-umbrella) while being timed. The first 8 error-free articulations of each pair will be averaged to produce a measure of articulation rate (Scarborough & Domgaard, 1998). In letter and digit forms of the task, participants will be asked to count from 1 to 10 five times in a row and then to repeat 'ABCD' ten times in a row. For both tasks, time to complete the specified number of repetitions is used as the measure of articulation rate (Ackerman & Dykman, 1993; Ackerman, Dykman, & Gardner, 1990).

10. Anti-saccade: This task measures executive control (Guitton, Buchtel, & Douglas, 1985; Roberts, Hager, & Heron, 1994). The participant’s goal is letter identification. Trials begin with a fixation point at center screen. At random intervals, the point is replaced by an attention-grabbing cue on one side of the screen. Immediately thereafter, a target letter is displayed on the side of the screen opposite the cue for 150 ms and then masked. The subject responds with a keypress to indicate whether the letter corresponds to a pre-determined target. Dependent measures are correctness of target identification, reaction time to correct trials, proportion of incorrect initial saccades to the cue, and latency of the first saccade in the direction of the target.

11. Familiarity with Standard English (SE) is assessed by a sentence repetition task (Radloff, 1991). This is important as a significant proportion of study participants will be speakers of African-American English dialect. Learners whose dialect is more distant from SE may be at a disadvantage in acquiring print skills (Charity, Scarborough, & Griffon, 2004). Dependent measures are numbers of phonological and grammatical discrepancies from SE.

D3.2 Experimental and Analytic Techniques

Regression Analyses and Structural Equation Modeling (experiment A)

Experiment A makes use of the behavioral battery gathered from all subjects participating in this study, and the expected large sample size allows strong tests of our critical hypotheses. Standard and sequential regression analyses will be used to (a) examine relationships among listening comprehension, decoding skill, verbal memory, print experience, and vocabulary in predicting reading comprehension, and (b) to more specifically test whether the initial Gough & Tunmer (1986) formula should be modified to include vocabulary knowledge. With these univariate regressions, we expect to replicate and extend the initial findings of a unique predictive effect of vocabulary on reading comprehension. In Braze et al. (submitted), effect sizes for partial correlations of interest range from modest to small. For example, in a model predicting reading comprehension, the effect size for the partial correlation of vocabulary above decoding skill and listening comprehension is $f^2 = .08$ (Cohen, 1988). With our anticipated N of 466 subjects, power to detect a comparable effect is about .99, with significance level .05.

To better understand the mechanisms mediating the effect of vocabulary quality on listening comprehension and to provide a stronger test of the hypothesis that vocabulary quality has an independent causal role, we turn to structural equation modeling (SEM; Fletcher, Foorman, Shaywitz, & Shaywitz, 1999; Pedhazur & Schmelkin, 1991) implemented in SPSS with AMOS. With 466 subjects, we will be able to include up to thirty indicator variables and still maintain stability and replicability of the SEM results.

<u>fMRI imaging and analysis (experiments G and J)</u>

Functional magnetic resonance imaging (fMRI) is performed on a Siemens 3.0T Trio scanner located at the Yale University School of Medicine. Subjects' heads are immobilized within a circularly polarized head coil by using a neck support, foam wedges, and a restraining band drawn tightly around the forehead. Prior to functional imaging, 32 axial-oblique anatomic images (TE (echo time), 4 ms; TR (repetition time), 350 ms; FOV (field of view), 40 x 40 cm; 4 mm slice thickness, no gap; 256 x256 x 2 NEX (number of excitations)) are prescribed parallel to the intercommissural line based on sagittal localizer images (TE, 6.83; TR, 20 ms; FOV, 25.6 cm; 4 mm slice thickness, no gap; 256 x 256 x 1 NEX). Activation images are collected using single shot, gradient echo, echo planar acquisitions (flip angle, 80 degrees; TE, 30 ms; TR, 2000 ms; FOV, 20 x 20 cm; 4 mm slice thickness, no gap; 64 x 64 x 1 NEX) at the same 32 slice locations used for anatomic images. High-resolution anatomical images are then gathered for 3-D reconstruction (sagittal MPRAGE acquisition, flip angle, 15; TE, 2.83 ms; TR, 1500 ms; FOV, 25.6 x 25.6 cm; 1 mm slice thickness, no gap; 256 x 256 x 2 NEX; 160 slices total). Total time required for subject setup and all scans is approximately 80 minutes. For each subject, a nonlinear transformation to Montreal Neurological Institute (MNI) template space is obtained using the intensity-only module of the algorithm described in Papademetris et al. (2004), mapping between the subject-space anatomic and the MNI-space "Colin" brain (available at http://www.bic.mni.mcgill.ca). Prior to across-subjects analysis, this transformation is applied with trilinear interpolation to bring the single-subject activation maps into 2 mm isotropic MNI space.

<u>Statistical Treatments:</u> Data analysis is performed using software written in MATLAB (Mathworks, 2001). Images are first corrected for motion with SPM-99 (Friston et al., 1995) and sinc-interpolated to correct for slice acquisition time. For single-subject event-related analysis, a regression-based method is used, which allows direct estimation of the hemodynamic response for each trial type, at each voxel separately, without prior specification of a reference function (Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000; Ollinger, Shulman, & Corbetta, 2001; for a recent application, see Sandak et al., 2004). Linear contrasts for effects of interest, including the evoked response of each trial type, simple subtractions among trial types, main effects, and interactions, are applied to these regression estimates to obtain contrast images for each subject. Across subjects, each voxel in these contrast images is tested versus zero with an F-test, implementing a mixed-model or repeated measures ANOVA (Holmes & Friston, 1998; Kirk, 1982; Woods, 1996). Primary analyses for reader differences will employ continuously-valued behavioral scores, and composite measures derived from clusters (see *D3: Skill and Experience Measures*, above) as regressors on activation values for each voxel separately, generating maps of correlations and p-values (also termed brain/behavior analysis; see Shaywitz et al., 2002, for a recent application).

In addition to standard univariate regression and ANOVA analyses, Partial Least Squares, or PLS, (McIntosh, Bookstein, Haxby, & Grady, 1996; Mencl et al., 2000; Shaywitz et al., 2003) will be used to further study the distributed systems underlying word reading and sentence comprehension. Since PLS is a multivariate analysis, it can supplement and corroborate findings from univariate analyses, as well as uncover effects not accessible through univariate tests (McIntosh et al., 1996). Univariate tests identify relationships between activation levels and a predictor variable (a contrast of experimental tasks, or a behavioral covariate) on a region-by-region basis; in contrast, PLS identifies brain-wide distributed patterns of activity, or components, which relate to one or more predictor variables. For example, in Experiment G, our initial hypothesis-driven ANOVA will test whether the activation levels in IFG in response to word consistency and sentence complexity are modulated by reading ability. Using PLS, we will then bring to bear two continuous behavioral measures, the Woodcock-Johnson Word Attack and Peabody Listening Comprehension, simultaneously as predictors of the brain activation values in the set of experimental conditions. Since the analysis assesses both common and unique predictability from the behavioral scores, we expect that (1) at least one of the extracted multivariate components will identify a brain-wide pattern of activity that is predicted by common variance from the two variables, related to general differences in effort and performance levels; and (2) other components will isolate activation patterns that specifically relate one behavioral component to activations in specific tasks, for example a relationship between the Word Attack scores and differences in brain activity for consistent versus inconsistent words.

To assess the statistical power of our proposed fMRI experiments, we make use of data collected in Project 1. In Experiment 1 of that project, we collected fMRI images from good and poor readers (N=21 and N=18, respectively), while they heard and read short sentences. We observed a strong reader group difference during the print reading trials in the ventral occipitotemporal cortex, extending earlier observations using single-word identification tasks (Salmelin, Service, Kiesilae, & Uutela, 1996; Shaywitz et al., 2002). We extracted the data from a single voxel at this location (Talairach coordinates x:-38; y:40; z:-8) that showed this effect (mean activation z-score for good readers, 0.6037; mean for poor readers, 0.3077; S.D., 0.3400), and used these parameters to compute statistical power (Cohen, 1988) for across-subjects effects in our proposed fMRI studies. With a two-tailed alpha level of 0.05, 22 subjects per group are needed to attain the standard power level of 0.8. The currently proposed sample size of N=50 (translating to 25 per dichotomized group) results in an estimated power level of 0.854, which should be sufficient to detect similarly-sized effects. We expect similar, or better, power levels for our proposed analyses, which utilize the full variation of the continuous predictor variables.

<u>Eye-movement monitoring during reading (experiment G)</u>

Eye movement recording provides critical on-line measures of the processing costs that arise as a reader encounters words and syntactic structures, while allowing reading to proceed normally. Patterns of eye fixations and regressions (refixations of earlier material) reflect, in part, the cognitive processes involved in reading because eye-movement characteristics co-vary with changes in processing demands over the course of a sentence (Murray & Kennedy, 1988; Rayner, 1998; Rayner, Reichle, & Pollatsek, 1998).

As part of Project 1, Haskins Laboratories acquired an SR Research Eyelink II head-mounted eye-tracking system (http://www.eyelinkinfo.com). The system consists of three miniature cameras mounted on a headband. Two eye cameras allow binocular eye tracking. The EyeLink II has high resolution (noise-limited at <0.01 deg.) and a fast sample rate (500hz max, 250hz in corneal reflection mode). EyeLink II control software integrates calibration, gaze position collection, head position compensation, and saccade and fixation analysis. The EyeLink system provides raw eye position relative to the head, and gaze position in pixel coordinates relative to the computer display. This is done by measuring head position via an infrared camera mounted on the head band which monitors four infrared LEDs mounted at the corners of the stimulus display CRT; software determines head distance and orientation relative to the display based on the relative positions of the LEDs in the infrared image. This arrangement allows the EyeLink to compensate for modest head movements.

<u>Statistical treatments:</u> *Dependent measures* of eye movement patterns: a) First-pass fixation duration, which is the summed duration of eye fixations within a defined region, provided that the fixations are not the result of a regressive eye movement. First-pass fixation times are statistically adjusted to compensate for inequalities in the length (number of letters and spaces) of each region across test conditions. b) Incidence of first-pass regressive eye-movements: A region is counted as having a first-pass regression if its final first-pass fixation ends in a backward glance to an earlier portion (left of the present region) of the sentence. c) Regression path target: further regressions that follow the first-pass regressions, from which the frequencies of revisits to a particular region are calculated. Each measure is tabulated for predefined regions of the test sentences. <u>Power Analysis:</u> We estimated the sample sizes required for within-subject effects based on our previously acquired eye movement data on ambiguity resolution (Ni et al, 1996a). A standard ANOVA was performed on the 22 subjects' first pass fixation times at the region of sentences where an ambiguity may arise. The result demonstrated that fixation times are significantly longer for ambiguous sentences than for unambiguous sentences ($F(1,21) = 9.91$; $p< 0.01$). Effect size for the difference between the two sentence

types was computed at 15%. This is the estimated proportion of population variance accounted for by the experimental manipulation, and is used to estimate necessary sample size for desired power levels. For a power level of 0.8 and a significance level of 0.05, a total of 32 participants is required.

Eye-movement monitoring over visual worlds (experiments C, E and F)

This eye-tracking paradigm can illuminate the fine-grained time-course of lexical activation during speech perception (Allopenna et al., 1998). A participant's eye-movements are monitored as they are asked to click on an image on a computer screen. From the participant's perspective, the task demands are quite low and are similar to those of conventional picture vocabulary tests like the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997). The technique has important advantages over other methods of assessing lexical activation or word knowledge. First, the linguistic stimuli occur in the context of a naturalistic task, resembling situations that occur outside the laboratory. Second, eye movements are typically not under conscious control. Third, they are fast and incur a low cognitive load. Depending on the task at hand, we typically make 2-3 saccades per second. Because they can be made so quickly, there is little cost associated with making eye movements to an "incorrect" location (as opposed to reaching for the wrong object). Thus, the decisional criteria for launching saccades are apparently lower than those for manual motor movements, and we often see eye movements based on partial information (see Allopenna et al., 1998 for discussion on interpreting eye movements).

The EyeLink II eye tracker, noted in the previous section, will be used for data collection. The 250 Hz sampling rate of this device is more than sufficient for the grain of question being asked; most studies on which ours are based used 60 Hz trackers. Operation in corneal reflection mode assures calibration stability. Gaze position computed by the EyeLink software will provide the primary source of data, although raw eye and head position will be recorded to allow periodic comparisons between the raw and processed data.

Statistical treatments: The time course data provided by eye tracking will be analyzed using standard ANOVA techniques over critical time windows (e.g., from 200 ms after word onset to the point at which target fixations asymptote). Various strategies have been used in the past, from repeated tests on series of smaller time windows (e.g., Allopenna et al., 1998), to mean proportion to each item in each window. Most strategies violate various assumptions of ANOVA. For example, using time sample as a factor is problematic because a value at time t is not independent of the preceding value, and tests on multiple windows increase the probability of false positives. Analyses on mean proportion within a single time window violate the fewest assumptions. When tests on smaller windows are deemed necessary, very conservative corrections will be made for multiple comparisons. In addition, the time course of comparable items (e.g., targets in different conditions) may be compared by examining differences in slope (derived from logistic regressions).

Speed-Accuracy Trade-offs (experiments B,D,H)

The SAT procedure provides conjoint measures of processing speed and accuracy by tracking changes in the accuracy of a response as a function of processing time. For each trial, experimental stimuli are displayed in Rapid Serial Visual Presentation (RSVP) format with 250 ms SOA. In auditory versions, stimuli will be played while a series of changing fixation points (i.e., *,%,@) are displayed on the screen to maintain attention. Following the final item, a response cue occurs both visually (i.e. a screen flash) and auditorily (i.e., a beep) and participants are asked to make a judgment appropriate to the particular experiment. For phonological similarity the judgment is "Do these words rhyme?" For semantic similarity, "Are these words from the same semantic category?" For sentences, the judgment is "Is this sentence acceptable?" Readers are trained in a separate 1-hour training session prior to the experiment to respond within 300 ms of the tone, even if they have to guess, so as to minimize elaborate decision processes in which subjects select a criterion for responding that balances speed and accuracy (Dosher, 1979; Ratcliff, 1978; Wickelgren, 1977). The onset of the response cue is varied across an appropriate range of times (e.g. 100–3000 ms. after the onset of the final stimulus) so that the full time-course of processing is sampled, from times when performance is at chance to times when performance has reached an asymptotic level. It is critical to sample processing *before* subjects have reached an asymptotic level, since it is in this period that differences in the rate of information accrual are observed. To accomplish this, six-eight appropriately spaced response-signals are usually sufficient.

Statistical treatments: The data at each response point is transformed to d-prime (d') scores by scaling the z-scores of the hit rate for an acceptable sentence against the z-score of the false alarm rate for the corresponding unacceptable sentence. Accuracy at each response delay will be plotted as a function of the delay of the response tone and the average reaction time to respond to the tone, ensuring that condition-specific differences in latency are factored into the estimates of processing accuracy at each lag (McElree &

Dosher, 1989, 1993). As illustrated in Figure 6, these full time-course functions typically display three distinct phases, namely a period of chance performance, followed by a period of increasing accuracy, followed by an asymptotic period beyond which further processing does not yield increases in accuracy. The three phases of SAT functions can be quantitatively summarized by the exponential approach to a limit shown in Equation (1) (Dosher, 1979; McElree, 1993; McElree & Dosher, 1989; Reed, 1973; Wickelgren, Corbett, & Dosher, 1980):

$$d'(t) = \lambda(1-e^{-\beta(t-\delta)}), \text{ for } t > \delta \text{ else } 0. \quad \text{(EQ. 1)}$$

Accuracy (here in d' units) is a function of three parameters, corresponding to the three phases of the SAT curve: λ serves to estimate the asymptotic level of performance, δ estimates the intercept or discrete point in time where performance begins to rise from chance, and β indexes the rate at which accuracy grows from chance to asymptote. The λ parameter describes the overall accuracy of the process, usually providing a measure of information availability. Differences in these parameters are illustrated in Panel A of Figure 6, where the functions differ in asymptote, but are associated with the same intercept and proportional rate of information accrual. The intercept (δ) and rate (β) parameters jointly describe the dynamics (speed) of processing. SAT dynamics reflect either the underlying accrual of information if processing is continuous or the underlying distribution of finishing times if processing is discrete or quantal (Dosher, 1976; McElree & Dosher, 1989; Meyer et al., 1998; Ratcliff, 1988). Panel B of Figure 6 illustrates disproportional dynamics, reaching a proportion of the asymptote at different times.

Figure 6 also illustrates how the SAT procedure provides additional information for interpreting reading times, and how it can discriminate between factors associated with memory strength and those associated with retrieval speed. Notice that panels A and B display quite different retrieval dynamics for the two conditions represented, despite identical response times (filled data points). In Panel A, the two conditions have similar intercepts and rates, while the asymptotes differ. In Panel B, the conditions have the same asymptote, though retrieval speeds differ. Such a retrieval dynamic may arise from an increasing tendency to retrieve incorrect items which are then revised (McElree, 1993), or in terms of our model, slower convergence in the presence of competing attractors (Tabor & Hutchins, 2004b).

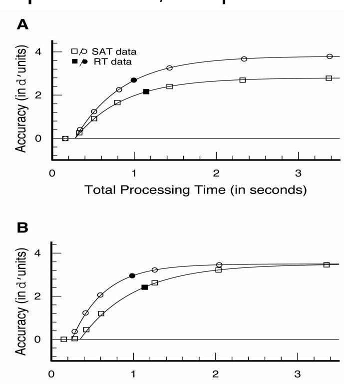

Figure 6

The proposed SAT studies involve contrasting the time-course of several conditions to determine whether they vary in response curve parameters. To do so, we use a form of hierarchical model testing in which various nested models are evaluated for their goodness-of-fit. We will explore the full parameter space ranging from a simple "null" model, in which conditions are fit with a single δ, β, and λ, to a fully saturated model, in which each condition is fit with a unique set of parameters. Goodness-of-fit is evaluated by a standard adjusted-R2 statistic, which assesses the proportion of variance accounted for by the model, adjusted by the number of free parameters (Judd & McClelland, 1989; McElree & Dosher, 1993; see McElree & Griffith, 1995).

Power analysis. To obtain sufficient data for model fitting, which is based on individual subject data, 30-40 trials per data point of interest are required (i.e., 30-40 trials in each condition at each response delay). This amounts to a minimum of 720 experimental trials for a standard 2x2 experiment with 6 response delays (not counting filler trials). In order not to overtax subjects, SAT experiments are typically divided into several 1-hour sessions. Each proposed SAT experiment states the specific number of sessions required.

Lexicalization of new word combinations (Experiment I)

Experiment I requires standard ANOVA and regression techniques. Retrospective power analyses indicate our sample size provides power of .80 with alpha at .05.

D4. Experimental Projects

D4.1 Vocabulary knowledge, Lexical access, and Reader Ability Differences

It is well established that syntactic and semantic processing occurs incrementally whether the linguistic signal is in print (e,g., Altmann & Kamide, 1999; Crain & Steedman, 1985; Ni, Crain, & Shankweiler, 1996) or speech (e.g., Delphine Dahan & Brent, 1999; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1996). A little explored question is what role individual differences in the quality and accessibility of lexical representations play in reading and listening comprehension. We have maintained that the print modality accentuates individual differences in language comprehension due to the increased cognitive load of the phonologic decoding step (Shankweiler & Crain, 1986). Poor and good readers differ in their ability to access phonetic and phonologic representations (Mark, Shankweiller, Liberman, & Fowler, 1977; Mody, Studdert-Kennedy, & Brady, 1997), and semantic representations (Byrne & Shea, 1979) in the speech context, as well. They also differ in their encoding abilities (Brady, 1997; Carroll & Snowling, 2004), and in the degree of lexical competition they experience in spoken word recognition (Metsala, 1997). However, it is difficult to differentiate between processing explanations (e.g., differences in speed of lexical activation) and structural explanations (e.g., vocabulary gaps or weaknesses). We will use eye tracking measures from which we can estimate the time course of lexical activation to ask whether poor reading ability is correlated with:1. Slowed lexical activation? (Experiment B); 2. Decreased temporal phonetic resolution? (Experiment C); 3. phonological or semantic interference? (Experiment D); 4. Increased or decreased semantic competition (Experiment E)?

Experiment A: Testing a regression-based model of reading comprehension. In this experiment we ask: **Does the Simple View of Reading accurately predict reading comprehension in a diverse sample of readers? Does orally assessed vocabulary knowledge contribute to reading comprehension over and above the components of the simple theory (listening comprehension and decoding skill), as predicted by the Lexical Quality Hypothesis?**

The primary goal is to replicate and extend our findings from Project 1 which found that vocabulary knowledge and listening comprehension make separate contributions to reading comprehension. This finding runs counter to a strict interpretation of Gough and Tunmer's Simple View of Reading (1986; Hoover & Gough, 1990), but is predicted by the Lexical Quality Hypothesis (Perfetti and Hart, 2002). A further goal is to seek a deeper understanding of the precursors of vocabulary knowledge. Two measures have predicted significant variance in word knowledge in other studies: experience with print (Cunningham & Stanovich, 1991; Cunningham et al., 1990; Stanovich & Cunningham, 1992; Stanovich, West, & Harrison, 1995) and verbal working memory (Aguiar & Brady, 1991; Atkins & Baddeley, 1998; Avons, Wragg, Cupples, & Lovegrove, 1998; Baddeley, Gathercole, & Papagno, 1998; Gathercole & Baddeley, 1990, 1997). Project 1 suggests that the contributions of these two factors are partially non-overlapping. Their relative contribution to word knowledge is not well understood.

All participants in the renewal project (N= 466) will receive the cognitive assessments listed below, in addition to participating in the various experimental tasks, assuring a large enough group to lend confidence to the results of a regression analysis based on these measures. The measures are as follows (referenced by task number from section D3.1): reading and listening comprehension (1); reading fluency (2); decoding skill (3); vocabulary knowledge (4); familiarity with print (7); articulation rate (9); and familiarity with standard English (11). The analysis will first evaluate the contributions of print decoding skill and listening comprehension to reading comprehension. Then, measures of vocabulary and Familiarity with Print will be entered to see if either captures additional unique variance. A second set of models will explicate the relationship between verbal working memory, print experience and vocabulary, controlling for Familiarity with Standard English.

We will follow these univariate regressions with specific model tests, implemented with Structural Equation Modeling (SEM). In our initial analyses, a majority of the variables (indicators) from the behavioral battery (see D3: Skill and Experience Measures) will be included, grouped into eight hypothesized latent constructs based on some of our earlier work with younger readers (Fletcher et al., 1996; Katz et al., in press): (1) reading comprehension (PIAT reading comprehension, GORT comprehension, WJ-III passage comprehension); (2) word identification (WJ-III Word ID; TOWRE word reading); (3) nonword decoding (WJ-III Word Attack; TOWRE non-word reading; pseudohomophone decoding); (4) listening comprehension (PIAT speech comprehension; GORT speech comprehension; WJ-III oral comprehension; sentence span); (5) fluency & automaticity (GORT total reading time; WJ-III reading fluency); (6) phonological awareness (CTOPP phonological awareness subtests); (7) memory (CTOPP memory for digits; CTOPP nonword repetition, sentence span) and (8) vocabulary (print exposure; PPVT; WASI vocabulary; Boston Naming; Synonym Identification).

We will employ the two-stage approach of first testing and modifying a measurement model, wherein each indicator is initially assumed to influence only a single construct and inter-construct relations are unconstrained. This analysis will indicate which indicators should be allowed to load on multiple constructs, and whether some indicators should be excluded. Second, we will test for goodness-of-fit differences among specific a priori structural models, where linkages among certain constructs are constrained or eliminated. We begin with a model with five basic ability constructs (4-8 above), that directly affect two mediating constructs of word and nonword decoding (2-3), that, in turn, determine the outcome measure of reading compehension (1). With this general approach, we expect to confirm a major result of Project 1 and to test questions raised by the newly identified influence of vocabulary within the reading system.

Predictions: As explained, the Lexical Quality Hypothesis and Project 1 findings lead us to expect that measures of vocabulary knowledge will continue to capture significant unique variance after decoding skill and listening comprehension are accounted for, allowing us to extend the model based on the Gough and Tunmer's Simple View. We predict further that vocabulary produces gains in reading comprehension both indirectly, by way of facilitating word identification, as well as directly by facilitating supralexical processes. This direct influence can be implemented in the model by allowing a single additional linkage from the vocabulary construct to the outcome comprehension measure, and tested by assessing the change in the model fit. Further, Project 1, and studies cited in C, indicate that verbal working memory and print experience are both related to vocabulary knowledge. We expect that, even after other measures of (visual) memory and experience (age, education) are taken into account, verbal memory and print experience each make unique contributions to vocabulary knowledge.

Experiment B: Phonological and semantic interference in recall. This experiment exploits the SAT procedure to investigate whether reading ability covaries with retrieval speed. It will ask: **1) Do poor readers retrieve phonological and semantic information more slowly than good readers? 2) Does retrieval of this information depend on the lexical quality of items being retrieved?**

The hypothesis that poor readers are slow in retrieval of phonological information is suggested by their poor performance on serial naming tasks (Bowers & Wolf, 1993; Cobbold, Passenger, & Terrell, 2003; Scarborough & Domgaard, 1998). The supporting data has been correlational, however. We propose to evaluate this hypothesis directly by requiring participants to judge phonological and semantic category membership over a range of response times that will reveal the time course over which this information becomes available. In the phonological similarity task, we manipulate lexical quality via orthographic consistency (i.e., whether a spelling has only one phonological pronounciation or many). This factor which has been shown to produce slower naming speeds and increased errors (e.g., Jared, McRae, & Seidenberg, 1990; Ziegler, Montant, & Jacobs, 1997) due to weak backward feedback from non-unique phonological representations. We include rhyming and non-rhyming words as baseline comparisons.

In the semantic task, we manipulate the quality of semantic representations via word frequency, following the assumption that high quality representations are the result of experience (Perfetti & Hart, 2002). Participants will be oriented to the concept of category membership prior to the task. In order to obviate any difficulty accessing the semantics of words caused by poor decoding skills, the semantic similarity task will be done in the auditory modality only. The phonological similarity task will be done in the print modality only.

Participants and Materials: We will include this experiment as 4 extra sessions for participants from each of the two larger SAT experiments described below, yielding N = 40. Reading ability as a continuous measure will be defined from the composite reading comprehension score. For follow-up categorical analyses, a categorical recoding will be applied to this variable (see D3: Skill and Experience Measures). We will use 40 sets of items, with 4 or 5 conditions in each set, as exemplified below.

Task A: Phonological similarity task (print modality): Do these words rhyme?

A1. (Consistent): DUCK-LUCK; A2. (Inconsistent): PINT-MINT; A3. (Nonwords): BLOAT-PLOTE; A4. (Word-control): TAKE-FAIR; A5. (Non-word control): CLOAT-PRAKE

Task B: Semantic similarity task (auditory modality): Are these words from the same category?

B1. (Similar High Frequency): CAR-TRUCK; B2: (Similar High-Low Frequency): CAR-VEHICLE; B3. (Dissimilar High-Frequency Control): CAR-DRINK; B4: (Dissimilar Low Frequency Control): CAR-DAGGER.

High and Low frequency synonyms will be identified with the help of WordNET; (Fellbaum, 1998), which generates sets of synonyms ordered by frequency, and LSA, (Landauer & Dumais, 1997). Phonologically consistent words will be identified using the database in Ziegler, Stone & Jacobs (1997).

Procedure: Tasks A and B will be presented using the SAT procedure described in D3. However, the comparison items will be presented immediately after a fixation screen, with variable processing time following their appearance. There will be 1200 trials in Task A and 960 trials in Task B (40 sets of 4 or 5 conditions presented at 6 response lags), presented over 2 experimental sessions for each subject (4 sessions for both tasks). Each set of 600 or 480 trials will be presented in 4 blocks of 150 or 120 each, with breaks in between.

Predictions: For both tasks we predict a reading ability x lexical quality interaction in speed parameters, with good readers less affected by the consistency/frequency variable because they have more highly tuned representations. In the phonological task, poor readers will be especially slowed by inconsistent spellings, as their stored phonological representations for these words are less well tuned. Similarly, poor readers will be slower to judge category membership when the semantics of a low frequency word must be accessed, as the meaning of these words is less elaborated (i.e., more shallow wells in our energy landscape model.) Poor readers will be more susceptible to errors, producing lower asymptotes on the SAT curve, especially for the inconsistent words and non-words in the phonological task and the similar and dissimilar high-low frequency conditions in the semantic task. If poor readers have a generally slower access mechanism than good readers, they will be slower for even the high quality (phonologically consistent or higher frequency) lexical items.

Experiment C: Temporal resolution in phonetic processing.

This study investigates a possible basis for word-level reader differences, asking: **Do poor readers show delayed access to the phonetic properties of words relative to good readers?** We will test this by examining the sensitivity of good and poor readers to conflicting coarticulatory information within words.

Participants: 60 readers will participate; the composite reading comprehension score will again be used as a continuous measure of reading ability, and a categorical recoding will be applied for follow-up analyses.

Procedure. Participants will see four objects on a computer display, and hear an instruction to click on one of them using the computer mouse (e.g., "click on the net"). The primary dependent measure will be the proportion of fixations over time to targets, competitors and unrelated distractors.

Materials: We will use 27 triplets like those used by Dahan, Magnuson, Tanenhaus & Hogan (2001). These are word/word/nonword sets like "net", "neck", and "nep". Three cross-spliced versions of each word are created by cutting the final consonant off versions of all three, and splicing each initial CV onto the final Cs of the two words. So for "net", we have three versions: ne(t)t (W1W1, the initial CV of word 1 spliced onto the final C of another recording of word 1, to ensure differences in other conditions do not result from cross-splicing), ne(k)t (W2W1, word 2 spliced onto word 1), and ne(p)t (N3W1, nonword spliced onto word 1). The effect of these cross-splicings is that coarticulatory information in the vowels of W2W1 and N3W1 mismatches the final target, potentially activating a competitor. In each case, listeners perceive "net" and do not report that the items sound unusual. However, eye tracking reveals differences in the time course of lexical activation (different from those indicated by, e.g., response times in lexical decision reported by Marslen-Wilson and Warren (1994) see Dahan et al. for details). As shown in Figure 7, from trials in which cross-spliced targets were displayed among unrelated distractors, typical adults recognize ne(t)t fastest, then ne(p)t, and are slowest on ne(k)t (W1W1 < N3W1 < W2W1). The explanation is that the coarticulatory information in ne(k)t strongly activates "neck", whereas the coarticulatory information in ne(p)t does not favor either "net" or "neck". We will create three cross splicings for every word (e.g., ne(t)t, ne(k)t, and ne(p)t and also (ne(k)ck), ne(t)ck and ne(p)ck), giving us 54 items in each of the W1W1, W2W1 and N3W1 categories. There will be three types of trials in this study.

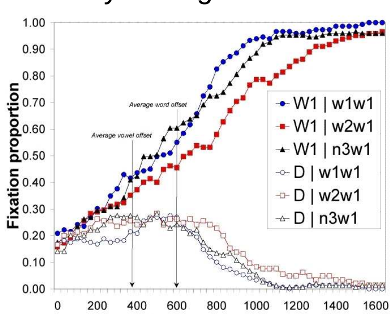

Figure 7

(1) Subcategorical mismatch, target only (18 trials). The two words from 9 triplets will each be randomly assigned to the W1W1, W2W1, or N3W1 condition and presented among three unrelated distractors. We will have 6 trials in each condition (compared to 5 per condition used by Dahan et al.).

(2) Subcategorical mismatch, competitor present (18 trials). Dahan et al. found larger differences between target types when competitors were simultaneously displayed (Figure 8). These items will come from the other 18 triplets. The target from each will randomly be assigned to the W1W1, W2W1, or N3W1 condition, so that we will have 6 trials in each condition.

(3) Fillers (64 trials). In half of these trials, no items will be related to one another. In the other half, a pair of onset competitors (e.g., “hammer” and “hammock”) will be displayed, but neither will be the target, to avoid any contingency between the presence of competitors and the identity of the target.

Predictions. Again, we will compare the time course for poor and normal readers. The goal will be to look for differences in the temporal resolution of phonological representations, which would be revealed by differences either in the rank ordering from fastest to slowest of the three cross-splice types, or from differences in the relative magnitudes of differences. If poor readers' phonological representations are less fully-specified or permit less temporal

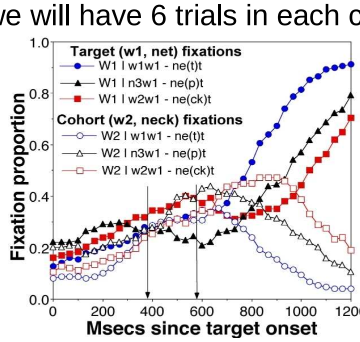

Figure 8

resolution (consistent with the PBH), we may see little difference between the three cross-splice conditions; if the bottom-up activation to lexical representations is relatively weak already, the subtle differences introduced in this experiment should not have much influence. On the other hand, on Metsala's (1997) weak inhibition explanation for differences observed in spoken word recognition competition effects for poor readers, we might expect larger effects of cross-splicing, especially in the W2W1 condition. In that condition, conflicting coarticulatory cues give W2 a "head-start." Once the final C is heard, W1 must quickly inhibit W2 in order to resolve the competition. If inhibition is weak, the time needed to resolve the competition should increase.

Experiment D: Susceptibility to phonologic and semantic interference in memory. This two part experiment utilizes the SAT procedure to investigate whether good and poor readers differ in their susceptibility to phonological and semantic interference in a memory task. It asks: **1) Are there (modality specific) reader skill differences in susceptibility to interference? 2) Does susceptibility to interference depend on the type or the amount of interference?**

Our model predicts that poor readers will be unaffected by interference if lexical representations are weakly tuned. This experiment will simultaneously test the speed and accuracy of item recognition in print and auditory modalities at high and low levels of interference when the lexical quality of the word to be recalled is high or low. There are six conditions, each containing 5 memory words and three probes (positive, negative related, and negative unrelated) associated with each. Items 1-3 show these conditions, with the lexical quality manipulation (consistent or inconsistent orthography for Part 1; high or low frequency for Part 2) presented in brackets [CONS/INCONS] or [HIGH/LOW]. All words will be displayed in lower case during the print experiment.

Part 1: Phonological interference:

1. Consistent Target Baseline: yellow, [CAKE/GREAT], library, surprise, sausage, vehicle. Positive probe: [CAKE/GREAT]; Negative related probe: [RAKE/LATE]; Negative unrelated probe: hammer.
2. Consistent Target 2 similar items: yellow, [CAKE/GREAT], [FLAKE/CRATE], surprise, sausage, vehicle. Positive probe: [CAKE/GREAT]; Negative related probe: [RAKE/LATE]; Negative unrelated: hammer.
3. Consistent Target 5 similar items: yellow, [CAKE/GREAT], [FLAKE/CRATE], [SNAKE/GATE], [MAKE/MATE], [SHAKE/FATE]. Positive probe: [CAKE/GREAT]; Negative related probe: [RAKE/LATE]; Negative unrelated probe: hammer.

Participants and Materials: 10 good and 10 poor readers will participate. The composite reading comprehension score will be used as a continuous measure of reading ability, and a categorical recoding will be applied for follow-up analyses. Consistent and Inconsistently spelled items will be identified using the published database (Ziegler, Stone, & Jacobs, 1997).

Part 2: Semantic interference:

1. High Frequency Target Baseline: yellow, [CAR/COUPE], library, surprise, sausage, vehicle. Positive Probe: [CAR/COUPE]; Negative related probe: van; Negative unrelated probe: hammer.
2. High Frequency Target 2 Similar: yellow, [CAR/COUPE], auto, surprise, sausage, vehicle. Positive probe: [CAR/COUPE]; Negative related probe: van; Negative unrelated probe: hammer.
3. High Frequency Target 5 Similar: yellow, [CAR/COUPE], auto, jeep, motorcar, vehicle. Positive probe: [CAR/COUPE]; Negative related probe: van; Negative unrelated probe: hammer.

Participants and Materials: 20 readers representing the range of ability will be chosen; the reading ability predictor is defined by the composite reading comprehension score. To easily visualize the predicted three-way interaction, a follow-up descriptive analysis will employ a categorical recoding to form groups. This experiment is an extension to the domain of individual differences of an experiment proposed by Van Dyke in NRSA proposal #1 F32 HD049215-01 and the design of materials will be informed by those results. In particular, that project will include collecting norming data on sets of similar items to verify the potential for similarity-based interference. In order to control for similarity among the different items in each memory list, they will be chosen from different semantic categories (i.e., colors, locations, abstract nouns, and foods).

Procedure: Ten sets of 324 memory lists (each set consisting of 18 instances of the 18 conditions) will be constructed from these conditions. The same 10 sets will be recorded for auditory presentation. Subjects will participate in 10 1-hour testing sessions, each consisting of both a print and auditory presentation of one of these sets, counterbalanced across sessions. The sets will be divided into blocks of 108 trials, with breaks in between. Prior to the experimental sessions, subjects will participate in a 1-hr practice session for familiarization with the SAT procedure and training to respond within 100-300 ms of the response tone. Within each session, the order of list presentation and the assignment of response lag for each memory list will be randomized. The test procedure follows that described in D3, with subjects seeing each word for 250 ms followed by a high contrast mask and then the probe word presented in uppercase. In the auditory condition, participants will sit at a computer while words are played through headphones. In the print condition, the probe word will remain on the screen until the tone signaling the subject to respond occurs (i.e. during the entire response lag). In the auditory condition, a visual and auditory response cue will occur after the onset of the probe. In each modality, subjects respond with yes-no keys indicating whether the probe word was in the study set. Feedback on response time will be displayed following each response.

Predictions: Predictions for both semantic and phonological tasks are the same. In keeping with the predictions of our model (see Section C1.3), we expect that retrieval speed will decrease as the amount of interference increases. This will only be true if readers perceive the similarity between items, however. Hence we predict a reading ability x interference interaction in which poor readers are unaffected by interference. This is consistent with our previous work showing that younger poor readers are deficient in perceiving phonological similarity in both print and auditory modalities (Mark, Shankweiler, Liberman, & Fowler, 1977; Shankweiler & Liberman, 1976), in contrast to good readers, who show marked phonological interference effects. As to the effect of modality, the PBH suggests that any deficit related to poor phonological encoding of memory items will disappear in the auditory modality. Thus, we predict a modality x reading ability x frequency interaction in which poor readers resemble good readers on the low frequency items in the auditory modality only.

Experiment E: Reading skill variations in the structure of item-specific neighborhoods. Even if the SAT studies reveal reader differences in lexical activation and competition from similar items, and Experiment C finds such differences in the specificity of phonological representations, these differences may not fully account for effects in spoken word recognition. The current experiment asks: **1) Do vocabulary differences associated with reading ability affect the time course of lexical competition? 2) Do processing differences result specifically from weaker lexical knowledge?**
The correlation discovered in Project 1 between spoken vocabulary and reading ability suggests that weakness in lexical knowledge may play an important role in the time course of word recognition. Characteristics like neighborhood density (Luce & Pisoni, 1998; Metsala, 1997) are calculated from corpus-based estimates of normal readers' language experience. These norms are particularly problematic for poor readers: absence of neighborhood effects (e.g., Metsala, 1997) may result from vocabulary differences rather than (or in addition to) differences in processing or phonological representation. We will examine this possibility with an eye tracking study followed by an assessment of item-specific neighborhoods. We will describe the eye tracking portion and predictions first, and then turn to the follow-up assessments.

Participants: 120 readers will participate; reading ability will be gauged by the composite reading comprehension score. In follow-up analyses a categorical recoding will be applied to assist interpretation.

Eye tracking Procedure. Participants will see four objects on a computer display, and hear an instruction to click on one of them using the computer mouse (e.g., "click on the net"). The primary dependent measure will be the proportion of fixations over time to targets, competitors and unrelated distractors.

Eye tracking materials. We will use a set of 128 items like those from Experiment 3 in Magnuson (2001), which manipulated word frequency, neighborhood density (summed frequencies of items differing from a target by no more than one phoneme; Luce & Pisoni, 1998), and cohort density (summed frequencies of items overlapping with a target through the vowel of the first syllable). The difference will be that we will control the proportion of neighbors that are also onset competitors. Magnuson presented the targets among unrelated distractors, and compared the time course of fixation proportions to targets. The comparisons revealed faster activation of high frequency words, words in sparse neighborhoods, and words in sparse onset cohorts.

Eye tracking predictions. First, the relative time course for each manipulation will shed light on the results of Metsala (1997), who found that reading impaired children do not show the standard neighborhood effect (faster performance on low-density items). We will examine whether an absence of such an effect is apparent throughout on-line processing. Metsala's explanation (weaker inhibition) would suggest that we should see a depressed time course for low-density items in poor readers, i.e., that even weak competition slows processing relative to typical readers. Alternatively, we may see equally fast performance for high- and low-density items in poor readers because functionally, the "high-density" neighborhoods and cohorts are not high density for them, because of vocabulary gaps or weaknesses. We can further refine our analysis of this issue with subanalyses on high-frequency targets that have primarily high-frequency competitors.

Lexical knowledge assessment procedure and materials. Following the eye tracking study, participants will hear a series of items related to a selection of the words used in the eye tracking portion and rate them on a scale from 1 (do not know) to 7 (highly familiar, I hear this word often). Familiarity ratings are highly correlated with lexical characteristics like frequency and neighborhood density (Nusbaum, Pisoni, & Davis, 1984). We will present representative (high- and low-frequency) items from the neighborhoods and onset cohorts of 25% of the words (4 from each of the frequency X neighborhood X onset cells of the design). We will use the ratings to estimate individual subjects' neighborhoods for each target word. We will re-estimate frequency, neighborhood and cohort densities accordingly, and examine whether this (a) provides better predictions of the time course of lexical activation for individuals and (b) identifies particular lexical parameters that tend to differ in poor readers (e.g., lower weight for low- to medium-frequency words).

D4.2 Testing phrase and sentence level predictions.

Experiment F: Integration of lexical and contextual information. This experiment uses a visual worlds eye-tracking paradigm to examine relations between differences in vocabulary quality, as indexed by our measures of vocabulary (D3), and differences in ability to integrate semantic information across words and with visual context in response to speech stimuli. It asks: **1) Will vocabulary weaknesses show delayed integration of spoken word meaning and visual context? 2) Are vocabulary weaknesses related to greater difficulty in integrating meaning across spoken words in adjective-noun pairs? 3) Is an individual's quality of word knowledge related to satisfaction of presuppositional constraints for restrictive modification?**

This experiment investigates the ability to integrate speech information (1) across words within adjective-noun pairs and, (2) with visual context (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). Modification of a noun by an adjective is a simple, ubiquitous grammatical function. Thus, it is an excellent test case for exploring vocabulary related differences in the ability to make syntactic and referential integrations. Additionally, our experimental materials will be built using simple vocabulary items; adjectives will all be common color names and nouns will all be names of common objects. Modification of a noun (e.g. *pen*) by an adjective (e.g. *blue*) carries the presupposition that there is an additional pen in the context that contrasts with the intended one in color. Utterances that fail to satisfy this presupposition incur an additional cognitive load (Crain & Steedman, 1985).

Participants: 60 individuals will participate in this experiment. A composite vocabulary score will be used as a continuous measure of vocabulary quality, and groups will be formed (see D3) for follow-up analyses.

Procedure and Materials: Each trial consists of a 4 picture array. Participants will follow the verbal instruction to (e.g.) "click the mouse on the blue pen", where "blue pen" is one of the images in the array.
Experimental conditions: 4 experimental conditions instantiate the interaction of 2 factors (resolution: early versus late; and contrast: present versus absent).

	Condition	**Items in array**	**Notes**
A	Early resolution No contrast	Target: blue pen; Foils: yellow bowl, red book, pink comb; green cup	Target can be identified at the word "blue". No contrasting target; presuppositions of adjectival modification are not satisfied.
B	Early resolution With contrast	Target: blue pen; Foils: red book, yellow pen, pink comb; green cup	Target can be identified at "blue". Contrasting target present; presuppositions are satisfied.
C	Late resolution No contrast	Target: blue pen; Foils: blue bowl, red book, pink comb; green cup	Target cannot be identified until "pen"; adjective is superfluous. No contrasting target; adjectival modification presuppositions are not satisfied.
D	Late resolution With contrast	Target: blue pen; Foils: blue bowl, yellow pen, pink comb; green cup	Target cannot be identified until the word "pen"; information from both "pen" and "blue" is needed for correct response. Presuppositions for restrictive modification are satisfied.

Predictions: (1) All individuals will show high accuracy in relating spoken word/phrase meaning to visual context; individual differences will emerge in the rate of convergence (2) Individuals with poorer quality vocabulary (less steep basins) will be slower, than those with better vocabulary, to converge on the appropriate item in the visual array even in conditions A and B, where no integration of meaning across words is necessary for success (3) All participants will be slower in the late resolution conditions than in the early resolution conditions (conditions C and D versus A and B), but those with poorer vocabulary will be especially delayed where integration of meaning across words (condition D) is required for success. (4) Because shallower basins result in slower cue integration, we predict an interaction between vocabulary quality and presence of presupposition: poorer readers will be slow to link the adjective in Condition D with the contrast set.

Experiment G: Crossing complexity of word decoding and sentence parsing complexity. This study continues our Project 1 program of using coordinated eyetracking and fMRI to investigate reading comprehension, with a view to understanding tradeoffs of syntactic and orthographic complexity. We ask: **1) How do eye movement correlates of comprehension skill reflect differences in processing load? 2) What are the brain-behavior correlates of comprehension skill under different processing loads? 3) What is the tradeoff between decoding ability and memory load in brain areas implicated in both types of processing?**

The PBH predicts that difficulties in processing complex sentences may stem more from decoding problems than deficiencies in higher-level syntactic processes. This may be because the ability to phonologically encode items to be remembered plays a crucial role in maintenance of material in working memory (A. Baddeley, S. E. Gathercole, & C. Papagno, 1998), or, as our attractor basin model suggests, because low quality phonology-meaning associations are less stable. Notably, both decoding demands and sentence complexity activate the inferior frontal gyrus (Caplan, 2001; Pugh et al., 1996; Shaywitz et al., 1998; Stromswold, Caplan, Alpert, & Rauch, 1996). fMRI provides an opportunity to observe the implications of the PBH directly.

Participants: 50 subjects will participate. Reading ability will be defined by the composite reading score.

Materials: Long distance dependencies provide a means for investigating memory load during sentence processing because they require retrieval of a noun that must be held in memory until its verb occurs later in the sentence. To test the effect of decoding skill we manipulate the spellingà phonology consistency of the noun being stored, under the assumption that poor phonological encoding will affect participants' ability to actively maintain this word. We test the effect of memory load in two ways: 1) manipulating the amount of material interpolated between the stored word and its verb and 2) manipulating the complexity of the material interpolated between the stored word and its verb. Examples 1-4 illustrate a 2 x 2 cross of consistency of the word to be retrieved (*clerk*/*aunt*) and complexity of the material between this word and its verb (*smiled*). The addition of conditions 5-6 provide baseline comparisons to those in 1-4 because there is no material between *clerk/aunt* and the verb *smiled.* Since we desire sentences of equal length for comparability in fMRI, we add a conjoined clause to the beginning of the sentence that will maintain semantic comparability across conditions without adding to the memory load. Phonologically consistent words will be identified using the database published in (Ziegler, Stone et al., 1997). An example of the complete paradigm is shown below:

1. (easy embedding, consistent noun)
   The boss noticed that the **clerk** <u>who had arrived in the fancy hat</u> smiled proudly.
2. (easy embedding, inconsistent noun)
   The boss noticed that the **aunt** w<u>ho had arrived in the fancy hat</u> smiled proudly.
3. (difficult embedding, consistent noun)
   The boss noticed that the **clerk** <u>who implied that the hat was fancy</u> smiled proudly.
4. (difficult embedding, inconsistent noun)
   The boss noticed that the **aunt** <u>who implied that the hat was fancy</u> smiled proudly.
5. (short, consistent noun)
   <u>The hat had been very fancy and</u> the boss noticed that the **clerk** smiled proudly.
6. (short, inconsistent noun)
   <u>The hat had been very fancy and</u> the boss noticed that the **aunt** smiled proudly.

Procedure*:* In Part A we will monitor eye movements as participants read each sentence on the screen. Following the sentence they will answer a YES/NO comprehension question testing whether they were able to complete the long distance dependency (i.e. Was it the clerk that smiled?) In Part B, they will participate in an event-related fMRI experiment (see Methods). Each sentence will be presented in a 500 ms per word RSVP format. Participants will read each sentence and press a key to indicate whether there was an automobile mentioned in the sentence. We will collect a total of 1980 full-brain images across 8 scan runs; a total of 360 trials will be presented at jittered intertrial intervals (9, 10, 11, and 12 seconds). This results in approximately 60 images at each time point for each condition.

Predictions: Regarding eye-movement patterns, we predict that the difficult embeddings will show longer reading times and more frequent regressions for all participants, particularly in the region of the final verb (i.e., *smiled*). Based on our model, we expect consistency and complexity to interact non-additively because of feedback in the attractor landscape model: one source of instability can be overcome by the feedback dynamics, but two sources working together are much more likely to devastate the parse. Similarly, the quality of lexical representations contributes a third source of instability, and is predicted to interact independently with both consistency and complexity.

As to the brain patterns, as noted above, the key region of interest is the inferior frontal region (IFG; Broca's area), where we expect to observe a tradeoff between head consistency and sentence complexity. Since poor readers overactivate IFG in apparent compensation for their poor decoding skills, we expect that fewer resources will be available for syntactic processes that also rely on this region. Consequently, we expect that poor readers will show the previously reported overactivation of the IFG in general, and this increase will be exacerbated for inconsistent words. Critically, we expect that syntactic complexity will evoke a *lesser* increase for poor readers, since resources in IFG are already allocated for phonologic processes. Finally, we will examine the three-way interaction (reader group x consistency x complexity) to see if this group difference is enhanced when resources are taxed on both dimensions. To implement this test with the continuously-varying composite reading comprehension score, we will first extract a single image for each subject that reflects the within-subject statistical contrast of consistency x complexity. Regressing the reading score onto these maps across subjects implements the desired brain/behavior correlation analysis. While not part of the primary hypotheses, we also expect to replicate the findings of reduced activity in posterior cortex for poor readers (Paulesu et al., 2001; Shaywitz et al., 1998). Here, we test if this effect holds in sentential contexts, in contrast to the single-word identification paradigms used previously.

<u>Experiment H: Interference effects in sentence processing.</u> This experiment will use the SAT procedure to assess individual differences in susceptibility to interference during the processing of complex sentences. It will extend the findings in Experiments B and C regarding susceptibility to interference in memory into the realm of sentence processing. It asks: **1) Does the presence of interference affect retrieval dynamics during sentence processing? 2) Are there reader differences in the ability to resolve long distance dependencies in the context of syntactic and semantic interference? 3) Are effects of interference on sentence processing specific to the print modality?**

Van Dyke (submitted; 2003) demonstrated semantic interference effects during sentence processing in normal readers and a further SAT study is planned in NRSA proposal #1 F32 HD049215-01 to Van Dyke. These results are consistent with a central prediction of the attractor basins model: distracting elements give rise to spurious attractor basins which tend to sidetrack the parser (Tabor & Hutchins, 2004; Tabor, Galantucci, & Richardson, 2004). This experiment extends that work into the domain of individual differences, testing the relation between retrieval dynamics and reading ability.The experimental paradigm consists of 5 conditions, a 2 x 2 cross of syntactic interference with semantic interference, plus an additional short sentence to serve as a no-distance baseline for comparing effects of distance between the subject to be retrieved and its verb. The high syntactic interference conditions contain an additional clause embedded between the target noun and the final verb; the high semantic interference conditions contain nouns that make sense as subjects of the final verb in this region. The conditions are given below in 1-5, with acceptable and unacceptable versions created by changing the final verb [ACCEPTABLE/UNACCEPTABLE]. All words will be presented in lower case during the visual modality experiment, and each condition will be presented in both auditory and visual modalities.

1. The client forgot that the man [SMILED/RAINED]. (short)
2. The client forgot that the man who had arrived for the important MEETING at the OFFICE [SMILED/RAINED]. (low syntactic interference, low semantic interference)
3. The client forgot that the man who had arrived with the important VISITOR for the BOSS [SMILED/RAINED]. (low syntactic interference, high semantic interference)
4. The client forgot that the man who had implied that the MEETING at the OFFICE was important [SMILED/RAINED]. (high syntactic interference, low semantic interference)
5. The client forgot that the man who had implied that the VISITOR for the BOSS was important [SMILED/RAINED]. (high syntactic interference, high semantic interference)

<u>Materials and Participants:</u> Twenty readers representing the range of ability will participate. Reading ability is defined by the composite reading comprehension score. Ten lists of 640 sentences will be created (32 instances of 20 conditions). These materials will be a subset of those used for Van Dyke's NRSA experiment, all of which are currently being normed to verify the potential for interference.

<u>Procedure:</u> Subjects participate in 10 1-hour sessions following the procedure described in D3. Lists of sentences are divided into blocks of 108 trials, with breaks in between. Prior to the sessions, subjects participate in a 1-hr practice session for familiarization with the SAT procedure and training to respond

between 100-300 ms of the response tone. Within each session, the assignment of response lag for each memory list is randomized. At the occurrence of the response probe, subjects respond YES/NO to the question “Is this sentence acceptable?” and response latency feedback is given.

Predictions: (2003)Differences in retrieval dynamics (i.e. slope and intercept) are predicted in response to interference. This is based on the fact that retrieval cues from the grammatical head become ambiguous when multiple similar items occur in the sentence, increasing the probability of retrieving an incorrect item, which must then be revised. Our modeling work suggests that this will interact with reading ability, such that poor readers have faster retrieval dynamics than skilled readers because they are less susceptible to ambiguous retrieval cues due to their less developed spurious attractor basins.

In addition to the interference effect, the distance manipulation may produce an interaction with reading ability if poor readers have an inherent limitation on their ability to access items in memory (a hypothesis tested directly in Experiment B). The results of Experiment B may also foreshadow a three-way interaction of distance, modality, and reading ability if it is discovered that poor readers have more difficulty accessing information presented in the print modality vs. in the auditory modality.(Van Dyke et al., 2005)

D4.3. Learning new words and combinations of words

Experiment I: Learning new words in high and low density phonological neighborhoods. **Do poor readers' weak lexical representations of words they know make it difficult to learn new words?**

We noted in section C1.3 that in activation based models, pseudowords correspond to spurious attractor basins (Plaut et al, 1996). Learning a new word involves deepening a spurious basin to make it more like the basins of existing words. Since poor readers' spurious attractor basins are shallow, noise is more likely to bump the system into a real-word attractor basin. Such disruption hampers the learning of new words—they tend to be confused with existing words. This confusion should be greater in high-density neighborhoods than low density neighborhoods. The experiment thus tests for an interaction between reader group and neighborhood density. The experiment uses a variant of the method introduced by Gaskell & Dumay (2003). In their method, subjects hear a set of nonwords 12 times per day for five days. The nonwords are similar to real words (e.g., “cathedruke” from “cathedral”). After one session, “lexicalization” effects begin to emerge in the form of increased recognition latencies for the real word counterparts, and effects are reliable after four days. Rate of nonword learning and degree of lexicalization (competition created for real words similar to the nonwords) are measured.

Participants. 80 individuals representing the range of ability will participate. Reading ability will be gauged by the composite reading comprehension measure.

Materials and Procedure. Gaskell & Dumay's items were controlled for onset competitor characteristics, but not directly for neighborhood factors. Rather than using the polysyallabic materials they used, we will use monosyllabic words (for which neighborhood density is better understood; Luce & Pisoni, 1998) differing from real words by a single phoneme. We will control the frequency and cohort density of the real words on which these critical nonwords are based, but will manipulate neighborhood density: half (15) the nonwords would fall into high-density English neighborhoods (e.g., “vait”, based on “vase”) and half (15) would fall into low-density neighborhoods (“fahv”, based on “fox”). We previously pointed out that neighborhood densities depend on individual subjects' vocabularies. We will restrict our target items to high-frequency, high-familiarity words, based on Nusbaum et al. (1984). that even poor readers should know. To ensure that the neighborhood density manipulation is valid even for poor readers, we will choose targets with a sufficient proportion of high-frequency, high-familiarity neighbors that the density manipulation will hold even if some subjects do not know every word in every neighborhood. This allows us to examine the lexical density hypothesis mentioned above: poor reading ability may be correlated with increased difficulty in learning new items with relatively many existing lexical competitors. Thus, the prediction is that poor readers will have greater trouble with high-density items than will good readers. Gaskell & Dumay made extensive use of phoneme monitoring tasks (e.g., for the nonword familiarization). This task would obviously be directly confounded with phonemic awareness and therefore is inappropriate for examining differences correlated with reading ability. Instead, we will use an old-new exposure paradigm for the familiarization task. Each of the 30 critical nonwords will be presented 12 times in each day's familiarization session. These will be randomly interspersed with a set of 360 nonwords randomly selected each day from a set of 10,000 pronounceable nonwords. The task will be to press keys indicating whether the word has been heard before (“yes”) or not (“no”). We will use accuracy and latency of “yes” responses as our index of nonword learning. We will test lexicalization with a lexical decision task including the 30 base words, 60 other words (half each from high- and low-density neighborhoods, and balanced on frequency and cohort density), and 60 randomly selected nonwords. The same set will be presented each day. On each of five consecutive days, participants will participate in the lexicalization task followed by the familiarization task. This order allows the lexicalization test to index the results of the previous day's learning.

Predictions. General deficits in the ability to acquire new sound forms would be revealed by a correlation between reading ability and learning rate in the familiarization task and/or degree of lexicalization (as indexed by increasing latencies for base word compared to control words in the lexical decision task). We expect an interaction between reader group and neighborhood density with a greater contrast between groups in high-density neighborhoods than in low.

Experiment J: Learning Novel Sequences of Familiar Words. This experiment focuses on the formation of new morpho-syntactic patterns, asking **1) Is learning novel sequences of familiar words a locus of difference between good and poor readers? 2) Does the violation of recently learned sequential expectations induce activity in the same neural circuitry as our prior studies of sentence anomaly?**

As noted in C2.2, Experiment 5 of Project 1 found evidence that adult subjects can be trained to recognize new sequences of familiar words in a short (20 min.) laboratory experiment and that inexperienced readers show less sensitivity to sequential order. Other research indicates that the process of acquiring word knowledge involves a transition from a holistic stage, when the representations are relatively order-insensitive, to an incremental stage, when the representations accurately encode sequential order (Charles-Luce & Luce, 1990; Magnuson et al., 2003). Here, we test the hypothesis that good readers reach the incremental stage more quickly than poor readers when trained on novel sequences of familiar words.

Participants: 50 individuals representing the range of ability will participate; reading skill is defined by the composite reading comprehension measure.

Materials and Procedure: Compounds of one-syllable words and nonwords (e.g., ROPE-SKY, ROPE-MIB, VAKE-SKY, VAKE-MIB) are presented over headphones with English compound intonation. 160 such stimuli are repeated during 5 training blocks. The task is to detect the presence of a nonword. During a single test block, the 160 training stimuli are interspersed among 160 novel stimuli. The participant must identify items seen before. During the test, half of the training items have one of the two elements replaced. In half of these replacement cases, the second element is replaced (e.g., ROPE-SKYà ROPE-DOT). In the remaining cases, the first element that is replaced (e.g., ROPE-SKYà BOLT-SKY). The test block, but not the training blocks are presented in fMRI. Judgments and response times are collected throughout training and test.

Predictions: *Behavioral Predictions*: We expect all readers to show a decrease in speed and accuracy on the distorted elements of the test block. Given our hypothesis that better readers have more sequential representations while poor readers have more holistic representations, we expect better readers to slow down greatly on the Second-Word-Replaced condition because it violates the sequential sequence expectation they have established. On the other hand, they should quickly reject the First-Word-Replaced condition because of its novel initial constituent. Poorer readers should slow-down equally on the First-Word-Replaced condition and the Second-Word-Replaced condition because they have learned a relatively unordered association.

fMRI Predictions: Because the word-word stimuli involve word-sequence learning as well as phoneme-sequence learning, we expect to observe activated regions in addition to those reported by Sandak et al. (2004) and Katz et al. (in press). In the Second-Word-Replaced condition, we expect good readers to show engagement of circuitry associated with violation of sequential expectation, particularly IFG and adjacent insula, where syntactic violation produces high blood flow (Ni et al., 2000). We expect the poor readers to show less activity in this region in the Second-Word-Replaced condition. Finally, we expect the First-Word-Replaced condition to produce a similar activation pattern to the Second-Word-Replaced condition in poor readers.

D5 Timeline

All experiments require participants to make at least two visits to the lab. Experiments involving eye-tracking and fMRI require a total time commitment of about 6 hours per participant (Experiments: C, E, F, J, H, I). Experiments involving SAT methodology (B, D and J) are particularly demanding, involving a commitment of about 20 hours for each participant, divided across 15 visits. We anticipate a total of about 2800 hours of participant contact for the entire project. Behavioral studies are evenly distributed across years 1 - 5. fMRI studies will be conducted in years 2 and 4, reflected in the increased budget for participant fees in those years.

E. Human Subjects Research

This Human Subjects Research meets the definition of 'Clinical Research.'

E1. Protection of Human Subjects

E1.1 Risks to the Subjects

E1.1.a. Human Subjects Involvement and Characteristics

In the proposed study there will be approximately 450-470 participants. Because the studies in this project are investigating questions concerning reading-related abilities in teenage children and young adults, the participants will range in age from 16 to 24 years of age. They will have no known physical or neurological abnormalities. Participants will be limited to those who are native speakers of American English in order to avoid confounding interpretation of the relations between language and cognitive abilities. No one will be excluded from participation based on gender, ethnicity or socio-cultural background. The ages of the individuals studied have been selected to reflect particular time periods in reading development that are being investigated (e.g., reading fluency will be investigated in adolescents and young adults). Ethnic and gender make-up of our sample is anticipated to reflect the general population of New Haven, Connecticut. See Section E2, below.

We will exclude individuals with IQ scores below 80, those whose native language is not English, and those with vision, hearing, attentional, or emotional problems such that they would not be able to assent to or participate in our tasks and assessments.

All behavioral data will be collected at Haskins Laboratories. Neuro-imaging (fMRI) data will be collected at Yale University School of Medicine.

E1.1.b. Sources of Materials

The studies will be conducted at Haskins Laboratories and Yale University School of Medicine. At each location appropriate space is available for data collection (e.g., quiet rooms). All planned studies entail individual administration of tasks.

All participants will partake of the battery of linguistic and non-linguistic cognitive tasks detailed in section D3 “Skill and Experience Measures,” which provide data for regression and structural modeling (Experiment A). Each participant will also produce data relevant to one experimental study, B through J. Data to be collected include measures of response accuracy and response times (experiments B, D, H, I), eye-movements over print (Experiment G), eye-movements over pictorial arrays (Experiments C, E, F), and recordings of brain activity (fMRI) in response to linguistic stimuli.

At entry into the study, each participant will be given an identifying code number (ID) and only these code numbers will be used in any materials submitted for data analysis. The case records will be stored in locked file cabinets and their use will be restricted to members of the research team. Participant names and ID numbers will be stored separately and only senior project personnel will have access to both.

E1.1.c. Potential Risks

There are no risks associated with participation in this study. None of the cognitive assessments or experimental tasks is invasive. In all tasks involving acoustic stimuli, these are presented at sound levels within the range of normal conversational levels. Participant fatigue is guarded against by dividing data collection into multiple sessions of no more than three hours each. Experiments involving eye-tracking and fMRI require a total time commitment of about 6 hours per participant (Experiments: C, E, F, J, H, I). Experiments involving SAT methodology (B, D and J) are particularly demanding, involving a commitment of about 20 hours for each participant, divided across 15 visits. MRI equipment used in the study will be in compliance with FDA regulations. There is no known harm associated with functional magnetic resonance imaging in children or adults. See E1.2.b. Risks to confidentiality are guarded against as described in E1.1.b and E1.2.b.

E1.2. Adequacy of Protection Against Risks

E1.2.a Recruitment and Informed Consent

Both written and oral descriptions of the experiment, including purposes, methods and procedures, will be offered to potential participants (and parent/guardian, if appropriate). Potential participants will have ample opportunity to ask questions about the purposes of and procedures used in the study, and they will be encouraged to do so. Federal guidelines will be followed regarding consent to participate and, in addition, we will ask for minors’ assent to take part in the study. Each adult participant will read and sign one or more consent form as the conditions of the experiment dictate. Minor participants will read and sign assent forms, and consent forms will be read and signed by their parents/guardians.

All participants will complete a checklist before being admitted into the fMRI scanner detailing if any metal particles could be present in or on the participant's body (information for minors will be provided by a parent/guardian). All metal objects will be removed before testing. Moreover, special care will be taken to ensure that the participants are fully informed about any discomforts that may be involved.

Participants will be told that they can discontinue participation at any time, without penalty.

E1.2.b Protection Against Risk

The safety of all participants will be insured by the following precautions.

1. Eye injuries: The risks inherent in the eye movement recording procedures are no greater than those encountered during the course of an individual's normal reading or recreational activities.
2. Ear injuries: The sound pressure levels to be used in listening experiments will be set a comfortable levels that pose no threat of injury to the ear. Special protective earphone will be used to protect the subjects from the noise generated by the scanning machine.
3. Exposure to Magnetic Fields: The MRI equipment used in this study will be in compliance with all FDA regulations. Standard precautions are taken in the MRI environment, including a metal inventory to determine the presence of such items as pacemakers and steel implants that would make the high magnetic field unsafe.
4. Confidentiality: The personal privacy of all participants will be protected by deleting all participant identification from data prior to their distribution or publication. Case records will be stored in secure (locked) locations accessible only to senior project personnel, as described in E1.1.b.
5. Fatique: All experiments involve participants making at least 2 visits to the labs (on separate days). No session will last more than 3 hours and there will be ample opportunity for breaks within sessions. Participants will be made aware that they can discontinue a session at any time.

E1.3. Potential Benefits of the Proposed Research to the Subjects and Others

The individual participants who take part in these experiments will be compensated with a small monetary amount. Other than that, they receive no direct benefits to health or well-being and will be made fully aware of that fact before participating. The principal benefits that may develop from the research are likely to be realized by society at large and may include better strategies for correcting or compensating for reading and language difficulties due to anatomical or neurocognitive dysfunction, or to inadequate educational opportunity.

E1.4. Importance of the Knowledge to Be Gained

Reading is an activity of immense social, psychological and economic importance. A better understanding of the processes of reading and spoken language and their neuronal instantiations can lead to the development of teaching materials to prevent and remediate reading handicaps. This is particularly important in the relatively understudied target population of adolescent and young adult poor readers.

E2 Inclusion of Women and Minorities

Efforts will be made to recruit minority participants from the high-school, adult-school and community college populations targeted in this study. We will collect data on gender and ethnic background on anonymous forms to evaluate whether we are sampling participants in proportion to their representation in the population. The numbers anticipated in the table below for gender and minority inclusion are based on demographic information (see below) for Connecticut and for more localized information for New Haven, Connecticut. Based on our expectations of where the individual studies will be conducted, the anticipated inclusion of individuals in different subpopulations were calculated for each location and then were combined in the table below. All experiments will include equal, or nearly equal, numbers of males and female.

Demographic information for New Haven, Connecticut is as follows: 0.3% Native American, 2.3% Asian or Pacific Islander, 34.3% African American, 12.4% Hispanic and 50.7% white. Connecticut, statewide: less than 1% Native American, 2.2% Asian, 13.7% African-American, 12.1% Hispanic, 71.5% white, 48.5% female. Every effort will be made to recruit a body of participants representative of the ethnic and gender make-up of the community.

E3 Inclusion of Children

The inclusion of both adolescents and young adults in our target population is dictated by our interest in studying the reading abilities characteristic of this age group. Building on our long history of research with children and young people, the project members are sensitive to how to design and engage individuals in this age group in age-appropriate activities, and in how to make the research experience pleasant, non-threatening and comfortable. Research assistants are carefully selected to have skills for working with young people and are trained in how to ensure that the rights of participants are respected at all times.

F. Vertebrate Animals

None

G. Literature Cited

Ackerman, P. T., & Dykman, R. A. (1993). Phonological Processes, Confrontational Naming, and Immediate Memory in Dyslexia. *Journal of Learning Disabilities, 26*(9), 597-609.

Ackerman, P. T., Dykman, R. A., & Gardner, M. Y. (1990). Counting Rate, Naming Rate, Phonological Sensitivity, and Memory Span - Major Factors in Dyslexia. *Journal of Learning Disabilities, 23*(5), 325-&.

Aguiar, L., & Brady, S. (1991). Vocabulary acquisition and reading ability. *Reading and Writing, 3*(3-4), 413-425.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory & Language, 38*(4), 419-439.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition, 73*(3), 247-264.

Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19*(4), 537.

Avons, S. E., Wragg, C. A., Cupples, L., & Lovegrove, W. J. (1998). Measures of phonological short term memory and their relationship to vocabulary development. *Applied Psycholinguistics, 19*(4), 583-601.

Baddeley, A., Gathercole, S. E., & Papagno, C. (1998). The phonological loopas a language learning device. *Psychological Review, 105*, 158-173.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158-173.

Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes, 12*(4), 423-484.

Bowers, P. G., & Wolf, M. (1993). Theoretical Links among Naming Speed, Precise Timing Mechanisms and Orthographic Skill in Dyslexia. *Reading and Writing, 5*(1), 69-85.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301*(5899), 419-421.

Brady, S. A. (1997). Ability to encode phonological representations: An underlying difficulty of poor readers. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia : implications for early intervention* (pp. 21-47). Mahwah, N.J.: L. Erlbaum Associates.

Braze, D. (2002). *Grammaticality, acceptability and sentence processing: A psycholinguistic study.* Unpublished Ph.D. dissertation, University of Connecticut, Storrs, CT.

Braze, D., Mencl, W. E., Tabor, W., & Shankweiler, D. (2005). Speaking up for Vocabulary: Reading Skill Differences in Young Adults. Toronto, Ontario: Twelfth Annual Meeting of the Society for the Scientific Study of Reading.

Braze, D., Shankweiler, D. P., Ni, W., & Palumbo, L. C. (2002). Readers' Eye Movements Distinguish Anomalies of Form and Content. *Journal of Psycholinguistic Research, 31*(1), 25-44.

Braze, D., Shankweiler, D. P., & Tabor, W. (2004). Individual Differences in Processing Anomalies of Form and Content. College Park, MD: Poster presented at the 17th CUNY Conference on Human Sentence Processing.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (submitted). *Reading Outside University Walls*.Unpublished manuscript, New Haven, CT.

Byrne, B., & Shea, P. (1979). Semantic and phonetic memory codes in beginning readers. *Memory and Cognition, 7*, 333-338.

Caplan, D. (2001). Functional neuroimaging studies of syntactic processing. *Journal of Psycholinguistic Research, 30*(3), 297-320.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77-126.

Carroll, J. M., & Snowling, M. J. (2004). Language and phonological skills in children at high risk of reading difficulties. . *Journal of Child Psychology and Psychiatry, 45* 631-640.

Charity, A. H., Scarborough, H. S., & Griffon, D. M. (2004). Familiarity with "School English" in African American Children and its Relation to Early Reading Achievement. *Child Development*, 75(5), 1340-1356.

Charles-Luce, J., & Luce, P. A. (1990). Similarity neighborhoods of words in young children's lexicons. *Journal of Child Language, 17*, 205-215.

Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush & E. Galanter (Eds.), *Handbook of Mathematical Psychology, vol. 2* (pp. 269-321). New York: Wiley.

Cobbold, S., Passenger, T., & Terrell, C. (2003). Serial naming speed and the component elements of speech time and pause time: relationships with the development of word-level reading in children aged four to five years. *Journal of Research in Reading, 26*(2), 165-176.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.

Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W., et al. (2004). Sentence complexity and input modality effects in sentence comprehension: An fMRI study. *Neuroimage, 22*(1), 11-21.

Crain, S., Ni, W., & Shankweiler, D. P. (2001). Grammatism. *Brain and Language, 77*(3), 294-304.

Crain, S., Shankweiler, D. P., Macaruso, P., & Bar-Shalom, E. (1990). Working memory and comprehension of spoken sentences: Investigations of children with reading disorder. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short term memory* (pp. 477-508). New York, NY: Cambridge University Press.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: the use of context by the psychological parser. In D. R. Dowty, L. Karttunen & A. M. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives* (pp. 320-358). Cambridge: Cambridge University Press.

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: a quick measure of reading experience. *Journal of Educational Psychology, 82*(4), 733-740.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83*(2), 264-274.

Cunningham, A. E., Stanovich, K. E., & Wilson, M. R. (1990). Cognitive variation in adult college students differing in reading ability. In T. H. Carr & B. A. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 129-159). San Diego, CA: Academic Press, Inc.

Dahan, D., & Brent, M. R. (1999). On the discovery of novel wordlike units from utterances: An artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology: General, 128*(2), 165-185.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Tracking the time course of subcategorical mismatches: Evidence for lexical competition. *Language and Cognitive Processes, 16*(5/6), 507-534.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466.

Dosher, B. A. (1979). Empirical approaches to information processing: Speed-accuracy tradeoff or reaction time. *Acta Psychologica, 43*, 347-359.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service, Inc.

Dunn, L. M., & Markwardt, F. C. J. (1970). *Examiner's manual: Peabody individual achievement test.* Circle Pines, MN: American Guidance Service.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research, 24*(6), 409-436.

Ekstrom, R., French, Harman, & Derman. (1976). *Kit of Factor-referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.

Elman, J. L. (1990). Representation and structure in connectionist models. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 345-382). Cambridge, MA: Mit Press.

Elman, J. L. (1991). Distributed Representations, Simple Recurrent Networks, And Grammatical Structure. *Machine Learning, 7*(2-3), 195-225.

Fellbaum, C. (Ed.). (1998). *Wordnet: An electronic lexical database*. Cambridge, MA: MIT Press.

Fletcher, J. M., Foorman, B. R., Shaywitz, S. E., & Shaywitz, B. A. (1999). Conceptual and methodological issues in dyslexia research: A lesson for developmental disorders. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders. Developmental cognitive neuroscience* (pp. 271-305). Cambridge: The MIT Press.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., et al. (1994). Cognitive Profiles of Reading Disability: Comparisons of Discrepancy and Low Achievement Definitions. *Journal of Educational Psychology, 86*(1), 6-23.

Fodor, J. D., Ni, W., Crain, S., & Shankweiler, D. P. (1996). Tasks and timing in the perception of linguistic anomaly. *Journal of Psycholinguistic Research, 25*(1), 25-57.
Foorman, B., Francis, D. J., Shaywitz, S. E., Shaywitz, B. A., & Fletcher, J. M. (1997). The case for early reading intervention. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia : implications for early intervention* (pp. 243-264). Mahwah, N.J.: L. Erlbaum Associates.
Fowler, A. E., & Scarborough, H. S. (1993). *Should reading disabled adults be distinguished from other adults seeking literacy instruction? A review of theory and research* (Technicial Report No. 93-6). Philadelphia: University of Pennsylvania, National Center on Adult Literacy.
Friederici, A. D., Ruschemeyer, S.-A., Hahne, A., Fiebach, C. J., & Friederici, A. D. (2003). The Role of Left Inferior Frontal and Superior Temporal Cortex in Sentence Comprehension: Localizing Syntactic and Semantic Processes. *Cerebral Cortex, 13*(2), 170-177.
Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping, 3*(3), 165-189.
Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*, 105-113.
Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: a study of young children learning new names. *British Journal of Psychology, 81*(4), 439-454.
Gathercole, S. E., & Baddeley, A. D. (1997). Sense and sensitivity in phonological memory and vocabulary development: a reply to Bowey (1996). *Journal of Experimental Child Psychology, 67*(2), 290-294.
Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition, 68*, 1-76.
Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz (Ed.), *Image, language, brain: Papers from the first mind articulation project symposium* (pp. 94-126). Cambridge, MA: MIT Press.
Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1-65.
Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1411-1423.
Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language, 51*, 97-114.
Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *RASE: Remedial & Special Education, 7*(1), 6-10.
Guitton, D., Buchtel, H. A., & Douglas, R. M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research, 58*, 455-472.
Hagoort, P., Brown, C. M., & Osterhout, L. E. (1999). The Neurocognition of Syntactic Processing. In C. M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 273-316). New York: Oxford University Press.
Harm, M. W., & Seidenberg, M. S. (2001). Are There Orthographic Impairments in Phonological Dyslexia? *Cognitive Neuropsychology, 18*(1), 71-92.
Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1998). Distinct time courses of word and context comprehension in the left temporal cortex. *Brain, 121*(6), 1133-1142.
Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1999). Semantic cortical activation in dyslexic readers. *Journal of Cognitive Neuroscience, 11*(5), 535-550.
Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers, 16*, 96-101.
Hogaboam, T. W., & Perfetti, C. A. (1975). Lexical ambiguity and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 14*(3), 265-274.
Holmes, A. P., & Friston, K. J. (1998). Generalizability, random effects, and population inference. *NeuroImage, 7*(Supplement), S34.
Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading & Writing, 2*(2), 127-160.
Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. In *Proceedings of the National Academy of Sciences* (Vol. 79, pp. 2554-2558).
Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effeects in word naming. *Journal of Memory and Language, 29*(687-715).
Judd, C. M., & McClelland, G. H. (1989). *Data analysis: A model-comparison approach.* San Diego: Harcourt Brace Jovanovich.

Kaan, E., & Vasic, N. (2004). Cross-serial dependencies in Dutch: Testing the influence of NP type on processing load. *Memory and Cognition, 32*(2), 175-184.
Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test* (2nd ed.). Philadelphia: Lea & Febiger.
Karlsen, B., & Gardner, E. F. (1985). *Stanford Diagnostic Reading Test* (3rd ed.). San Antonio, Texas: Harcourt Brace Educational Measurement.
Katz, L., Lee, C. H., Tabor, W., Frost, S. J., Mencl, W. E., Sandak, R., et al. (in press). Behavioral and neurobiological effects of printed word repetition in lexical decision. *Neuropsychologia*.
Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: a parallel distributed processing account. *Journal of Memory and Language, 32,* 474-516.
Kirk, R. E. (1982). *Experimental design: Procedures for the social sciences*. Belmont, CA: Wadsworth.
Kutas, M., & Hillyard, S. A. (1983). Event related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition, 11*(5), 539-550.
Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*(2), 211-240.
Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-Emerging Reading Disabilities. *Journal of Educational Psychology, 95*(2), 211-224.
Liberman, I. Y., Shankweiler, D. P., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1-33). Ann Arbor, MI: The University of Michigan Press.
Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model *Ear and Hearing, 19*, 1-36.
Lundquist, E. N. (2003). *Phonological complexity, decoding and text comprehension.* Unpublished Ph.D. dissertation, University of Connecticut, Storrs.
Macaruso, P., Bar-Shalom, E., Crain, S., & Shankweiler, D. P. (1989). Comprehension of temporal terms by good and poor readers. *Language and Speech, 32*(1), 45-67.
Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General, 132*(2), 202-227.
Mark, L. S., Shankweiler, D. P., Liberman, I. Y., & Fowler, C. A. (1977). Phonetic recoding and reading difficulty in beginning readers. *Memory and Cognition, 5*(6), 623-629.
Mark, L. S., Shankweiller, D. P., Liberman, I. Y., & Fowler, C. A. (1977). Phonetic recoding and reading difficulty in beginning readers. *Memory and Cognition, 5*, 623-629.
Markwardt, F. C., Jr. (1998). *Peabody Individual Achievement Test--Revised*. Circle Pines, MN: American Guidance Service, Inc.
Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101*(4), 653-675.
Mathworks. (2001). Matlab. Natick, MA: The MathWorks Inc.
McElree, B. (1993). The locus of lexical preference effects in sentence comprehension: A time course analysis. *Journal of Memory & Language, 32*, 536-571.
McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory: Time course of recognition. *Journal of Experimental Psychology: General, 118*, 346-373.
McElree, B., & Dosher, B. A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General, 122*, 291-315.
McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory & Language, 48*(1), 67-91.
McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 134-157.
McElree, B., & Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence comprehension: A time-course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(2), 432-460.
McIntosh, A. R., Bookstein, F. L., Haxby, J. V., & Grady, C. L. (1996). Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage, 3*(3 Pt 1), 143-157.
Mencl, W. E., Pugh, K. R., Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Constable, R. T., et al. (2000). Network analysis of brain activations in working memory: behavior and age relationships. *Microscopy Research & Technique, 51*(1), 64-74.

Metsala, J. L. (1997). Spoken word recognition in reading disabled children. *Journal of Educational Psychology, 89*, 159-169.
Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the Hemodynamic Response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *NeuroImage, 11*, 735-759.
Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: Auditory processing of phonological encoding? Journal of Experimental Child *Psychology, 64*, 199-231
Murray, W. S., & Kennedy, A. (1988). Spatial coding in the processing of anaphor by good and poor readers: evidence from eye movement analyses. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 40*(4-a), 693-718.
Ni, W., Constable, T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., et al. (2000). An event-related neuroimaging study: distinguishing form and content in sentence processing. *Journal of Cognitive Neuroscience, 12*(1), 120-133.
Ni, W., Crain, S., & Shankweiler, D. P. (1996). Sidestepping garden paths: assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes, 11*(3), 283-334.
Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. P. (1998). Anomalous strings: eye movement patterns. *Journal of Psycholinguistic Research, 27*(5), 515-539.
Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). *Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words*. Bloomington, IN: Speech Research Laboratory, Department of Psychology.
Ollinger, J. M., Shulman, G. L., & Corbetta, M. (2001). Separating processes within a trial in event-related functional MRI: I. The Method. *Neuroimage, 13*(1), 210-217.
Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243-277). Baltimore, MD: Paul H.
Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language, 59*(3), 494-522.
Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785-806.
Papademetris, X., Jackowski, A. P., Schultz, R. T., Staib, L. H., & Duncan, J. S. (2004). *Inegrated intensity and point-featured nonrigid registration.* Paper presented at the MICCAI.
Paulesu, E., Demonet, J. F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001). Dyslexia: Cultural diversity and biological unity. *Science, 291*(5511), 2165-2167.
Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, Design, and Analysis: An Integrated Approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
Perfetti, C. A. (1990). The cooperative language processors: Semantic influences in an autonomous syntax. In D. A. Balota, G. B. F. d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 205-230). Hillsdale, NJ: Erlbaum.
Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri & R. Trieman (Eds.), *Reading Acquition* (pp. 145-174). Hillsdale, NJ: Erlbaum.
Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Precursors of Functional Literacy*. Amsterdam: John Benjamins Publising Company.
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. A. (1996). Understanding Normal and Impaired Reading: Computational Principles in Quasi-Regular Domains. *Psychological Review, 103,* 56-115.
Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation & Developmental Disabilities Research Reviews, 6*(3), 207-213.
Pugh, K. R., Mencl, W. E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Constable, R. T., et al. (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity within posterior cortex. *Psychological Science, 11*(1), 51-56.
Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1996). Cerebral organization of component processes in reading. *Brain, 119*, 1221-1238.
Radloff, C. F. (1991). *Sentence Repetition Testing for Studies of Community Bilingualism* (Vol. 104). Arlington, TX: Summer Institute of Linguistics.
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59-108.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Rayner, K., Reichle, E. D., & Pollatsek, A. (1998). Eye Movement Control in Reading: An Overview and Model. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (pp. 243-268). Amsterdam: Elsevier.

Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science, 181*, 574-576.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General, 123*(4), 374-393.

Salmelin, R., Service, E., Kiesilae, P., & Uutela, K. (1996). Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Annals of Neurology, 40*(2), 157-162.

Sandak, R., Mencl, W., Frost, S. J., Rueckl, J. G., Katz, L., Moore, D. L., et al. (2004). The neurobiology of adaptive learning in reading: A contrast of different training conditions. *Cognitive, Affective & Behavioral Neuroscience, 4*(1), 67-88.

Scarborough, H. S., & Domgaard, R. M. (1998, April 19, 1998). *An exploration of the relationship between reading and rapid serial naming speed.* Paper presented at the Society for the Scientific Study of Reading, San Diego.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*(2), 109-147.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Shankweiler, D. P., & Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition, 24*(1-2), 139-168.

Shankweiler, D. P., Crain, S., Katz, L., Fowler, A. E., Liberman, A. M., Brady, S. A., et al. (1995). Cognitive profiles of reading-disabled children: comparison of language skills in phonology, morphology and syntax. *Psychological Science, 6*(3), 149-156.

Shankweiler, D. P., & Liberman, I. Y. (1976). Exploring the relations between reading and speech. In R. M. Knights & D. K. Bakker (Eds.), *Neuropsychology of learning disorders: Theoretical approaches*. Baltimore: University Park Press.

Shankweiler, D. P., Lundquist, E., Dreyer, L. G., & Dickinson, C. C. (1996). Reading and spelling difficulties in high school students: Causes and consequences. *Reading & Writing, 8*(3), 267-294.

Shankweiler, D. P., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., et al. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading, 3*(1), 69-94.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarksi, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry, 52(2) Jul 2002*, US http://www.

Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., et al. (2003). Neural systems for compensation and persistence: Young adult outcome of childhood reading disability. *Biological Psychiatry, 54(1) Jul 2003*.

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., et al. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences USA, 95*(5), 2636-2641.

Smith, S. T., Macaruso, P., Shankweiler, D. P., & Crain, S. (1989). Syntactic comprehension in young poor readers. *Applied Psycholinguistics, 10*(4), 429-454.

Spring, C., & French, L. (1990). Identifying children with specific reading disabilities from listening and reading discrepancy scores. *Journal of Learning Disabilities, 23*(1), 53-58.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-406.

Stanovich, K. E. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly, 26*(1), 7-29.

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: the cognitive correlates of print exposure. *Memory and Cognition, 20*(1), 51-68.

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*(1), 24-53.

Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance across the life span: the role of print exposure. *Developmental Psychology, 31*(5), 811-826.

Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by Positron Emission Tomography. *Brain and Language, 52*, 452-473.

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of Merely Local Syntactic Coherence on Sentence Processing. *Journal of Memory and Language, 50*(4), 355-370.

Tabor, W., & Hutchins, S. (2004a). Evidence for Self-Organized Sentence Processing: Digging In Effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 431-450.

Tabor, W., & Hutchins, S. (2004b). Evidence for self-organized sentence processing: Digging in effects. *Journal of Experimental Psychology: Language, Memory, and Cognition, 30*(2), 431-450.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eve movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*(6), 557-580.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1996). Using eye movements to study spoken language comprehension: evidence for visually mediated incremental interpretation. In T. Inui & J. L. McClelland (Eds.), *Attention and performance 16: Information integration in perception and communication* (pp. 457-478). Cambridge, MA: MIT Press.

The Psychological Corporation. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio: Harcourt Brace & Co.

Torgeson, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Tests of Word Reading Efficiency (TOWRE)*. Austin, TX: Pro-Ed.

Van Dyke, J. A. (2002). *Retrieval Effects in Sentence Parsing and Interpretation.* Unpublished Ph.D. Dissertation, University of Pittsburgh, Pittsburgh.

Van Dyke, J. A. (submitted). Effects of retrieval interference in sentence parsing.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. *Journal of Memory and Language, 49*(3), 285-316.

Van Dyke, J. A., Shankweiler, D., & Tabor, W. (2005, June 24-26). *Individual differences in the time-course of sensitivity to syntactic and semantic interference during comprehension of complex sentences.* Paper presented at the Society for the Scientific Study of Reading, Toronto.

Wagner, R. K., Torgeson, J. K., & Rashotte, C. A. (1999). *The Comprehensive Test of Phonological Processing (CTOPP)*. Austin, TX: Pro-Ed.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41*, 67-85.

Wickelgren, W. A., Corbett, A. T., & Dosher, B. A. (1980). Priming and retrieval from short-term memory: A speed accuracy trade-off analysis. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 387-404.

Widerhold, J. L., & Bryant, B. R. (2001). *Gray Oral Reading Test (GORT) (4th ed.)*. Austin, TX: Pre-Ed.

Wolf, A. (1991). Assessing core skills: Wisdom or wild goose chase?, *Cambridge Journal of Education* (Vol. 21, pp. 189): Carfax Publishing Company.

Wolf, M., & Bowers, P. G. (1999). The Double-Deficit Hypothesis for the Developmental Dyslexias. *Journal of Educational Psychology September, 91*(3), 415-438.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

Woods, R. P. (1996). Modeling for intergroup comparisons of imaging data. *NeuroImage, 4*(Supplement), S84-S94.

Yee, E., Blumstein, S. E., & Sedivy, J. C. (2004). The time course of lexical activation in Broca's and Wernicke's aphasia: Evidence from eye-movements. *Brain and Language, 91*, 62-63.

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language, 37*, 533-554.

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What's the pronunciation for _OUGH and the spelling for /u/? A database for computing feedforward and feedback inconsistency in English. *Behavior Research Methods, Instruments, and Computers, 29*(4), 600-618.