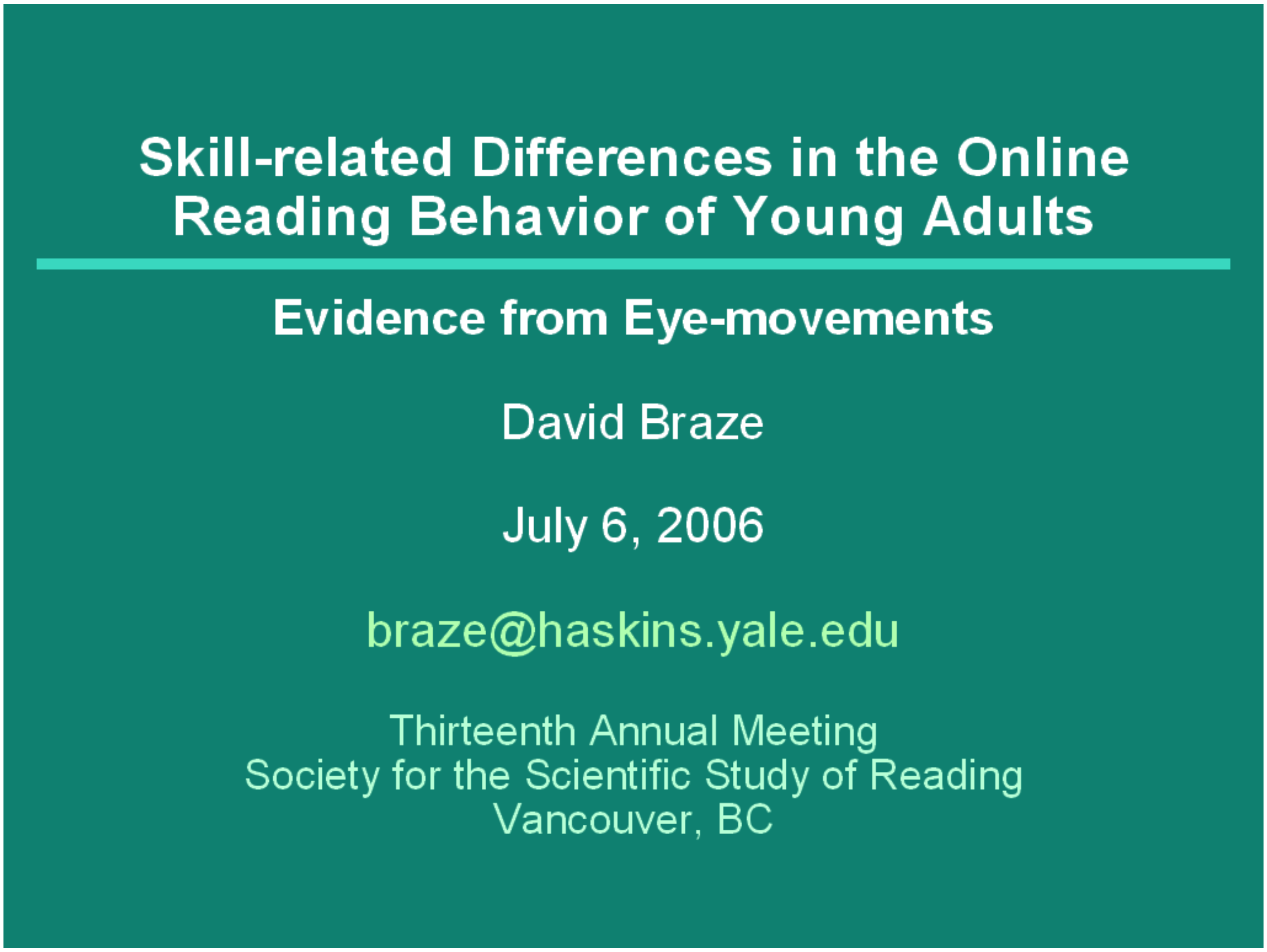
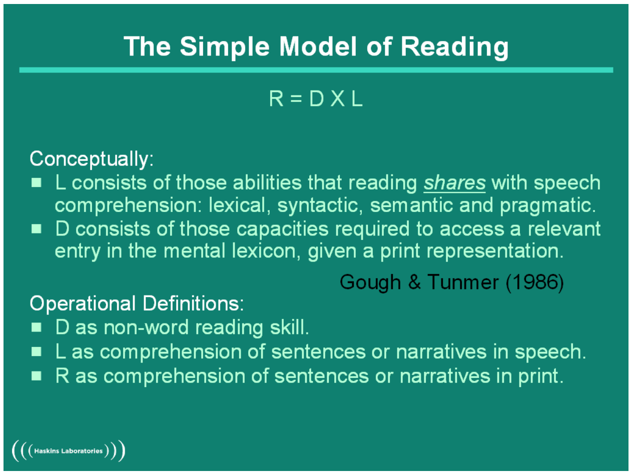


Slide **1**

[slide 1] The project I'm going to describe today has two primary goals. First, we aim to profile the reading-related aptitudes of adolescents and young adults whose limited literacy skills pose serious impediments to their occupational and educational prospects. To be clear, our focus is on garden-variety poor readers, a group distinguishable from the targets of much psycholinguistic and reading research by neither fitting usual criteria for reading disability, nor in being just another convenience sample of university undergraduates.

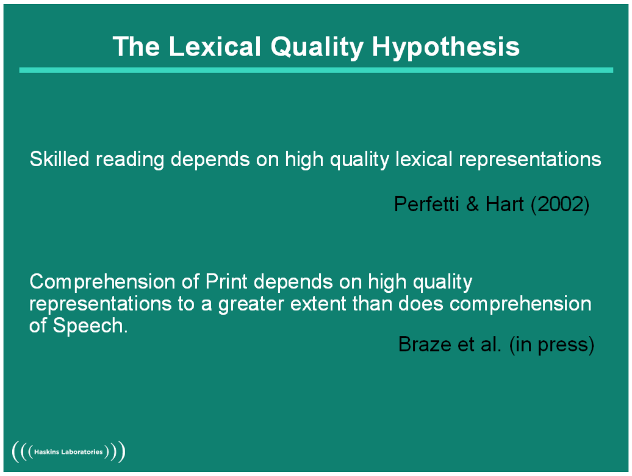
Conventional assessments of reading skill and its components tend to look at comprehension as a <u>product</u>. Our second goal, primary in the context of today's talk, is to understanding the <u>process</u> of comprehension, comprehension as it unfolds in time. We use the method of monitoring eye-movements over print to collect online indicators of cognitive and linguistic effort related to the comprehension process.



Slide **2**

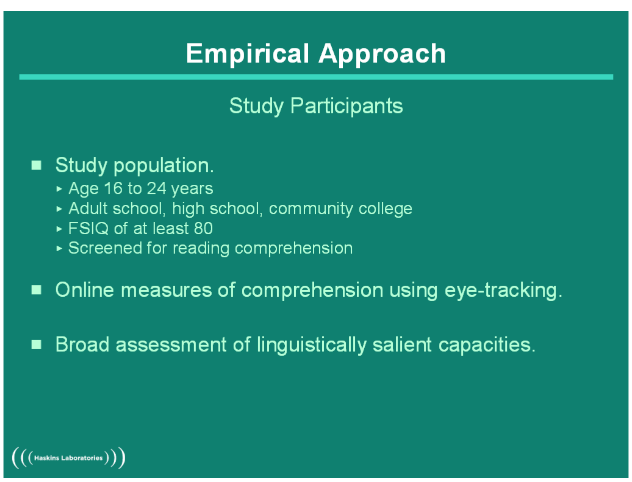
[slide 2] Those goals are guided by a couple of leading ideas: First, we ask whether Gough and Tunmer's (1986) Simple Model of Reading, as typically operationalized, is sufficient to explain all important variation in reading comprehension. The simple model aims to cope with the complexity of reading comprehension by partitioning the skills that contribute to it into two sets. Specifically, they proposed that reading comprehension (R) is the product of decoding skill (D) and the general capacity for language (L). L includes all those elements needed for the comprehension of speech, while D represents the ability to access a word's meaning by way of a print representation; D is the one new thing that must be learned to break into literacy. Yet, clearly both D and L are complex (as is R, for that matter) and the manner in which each is measured matters. We operationalize D as non-word reading skill (word-attack subtest Woodcock, McGrew, & Mather, 2001) and L as a sentence-picture matching task (details in Braze, Tabor, Shankweiler, & Mencl, in press, which describes a companion study of a separate cohort from the same population).

Moreover, we are interested in evaluating whether the simple model will be as helpful in explaining the comprehension process (evaluated through eye-movement measures) as well as its product (assessed through conventional reading comprehension measures).



Slide **3**

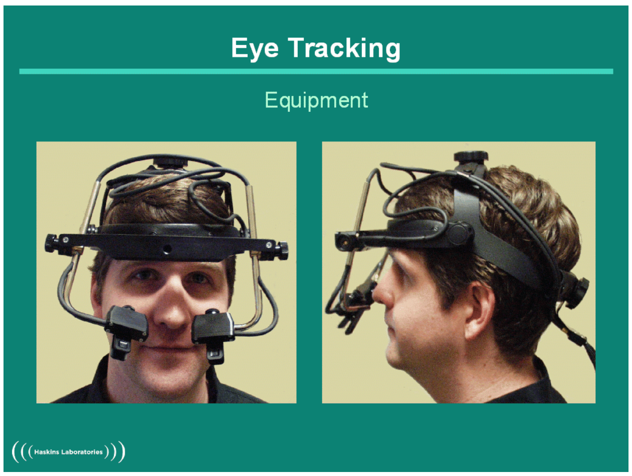
[slide 3] Relatedly, we ask whether a deeper appreciation for variation in vocabulary knowledge may help to understand individual differences in reading comprehension, either as process or as product, beyond what can be derived from the simple model (Perfetti & Hart, 2002). A high quality lexical representation is one in which the components of that representation, phonological, orthographic, semantic, are each well-represented and well-integrated with each other. Moreover, that word has clearly defined relationships with others in the lexicon. It has also been proposed that the comprehension of print depends on high quality representations to a greater extent than does the comprehension of speech (Braze et al., in press). While this is tangential to our purpose today, it suggests that comprehension differences stemming from differences in word knowledge may be sharper in the domain of print comprehension than in the domain of speech comprehension.



Slide **4**

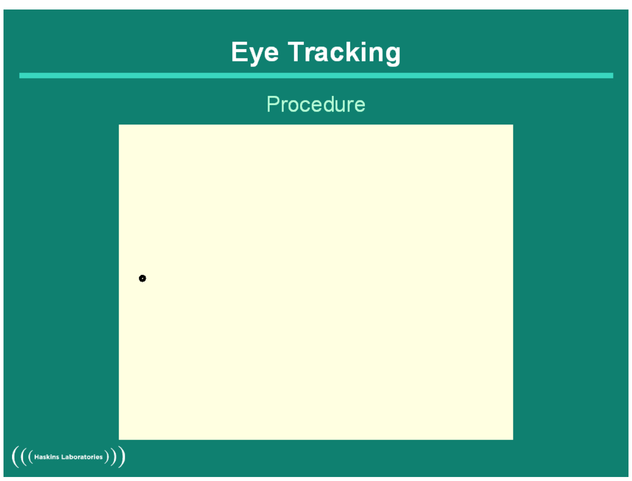
[slide 4] Our approach to these questions has three key components. First we recruit young adults with a wide-range of reading skill (N=46). Primarily, they have adult school, high school or community college educational backgrounds and a large proportion of our participants have literacy skills limited enough to constrain their educational or

occupational prospects. That said, participants must have a minimum full scale IQ of 80 (The Psychological Corporation, 1999) and the ability to read simple sentence materials with comprehension (details in Braze et al., in press). Second, we use eye-tracking to collect online measures of effort during the comprehension process. Finally, we collect a broad array of measures of language and language-related capacities. I'll elaborate on these second two points immediately, eye-tracking first.

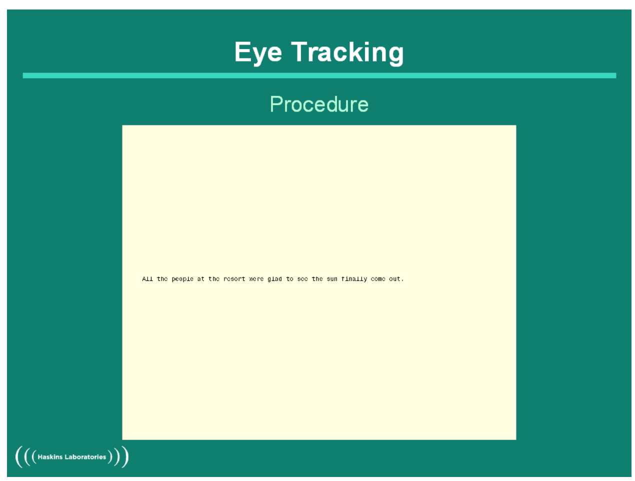


Slide **5**

[slide 5] We use a head-mounted video-based eye-tracking system made by SR Research. This is a binocular system with integrated head-tracking. Thus there is no need to use head-restraints like bite-bars or chinrests, making for a much more comfortable experience for participants.



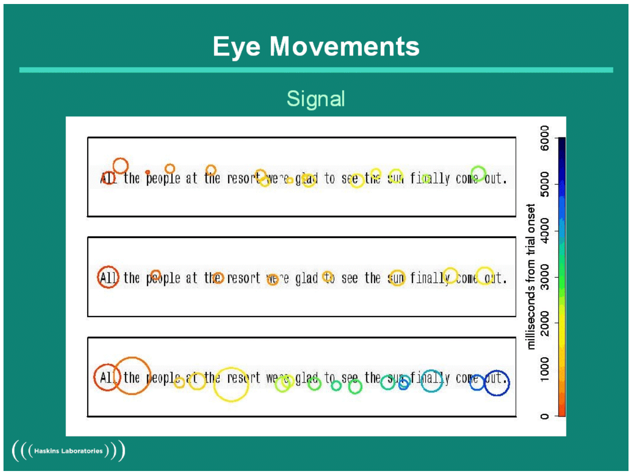
Slide **6**



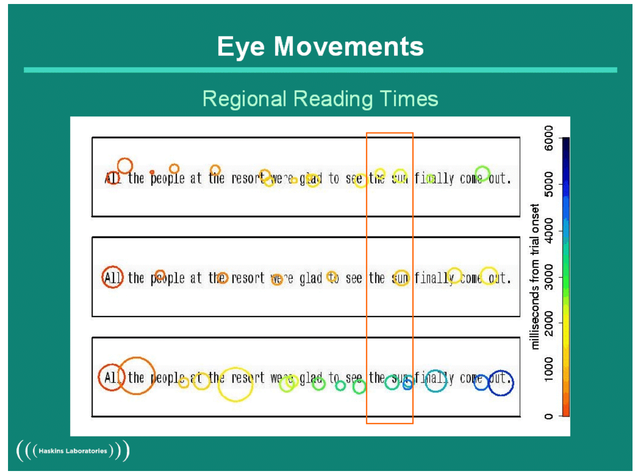
Slide **7**

[slide 6] Sentence materials are presented one sentence at a time, each on a single line. Sentences are preceded by a fixation target. While fixating the target, the participant clicks a button

to bring up a sentence, which replaces the target [slide 7]. The participant reads the sentence, clicks a button to indicate completion, at which point the sentence is replaced by another fixation target or, on some trials, a comprehension question (not shown). During the trials, gaze position is sampled at 250hz. Samples are then converted to fixations. Fixation patterns form the basic data.



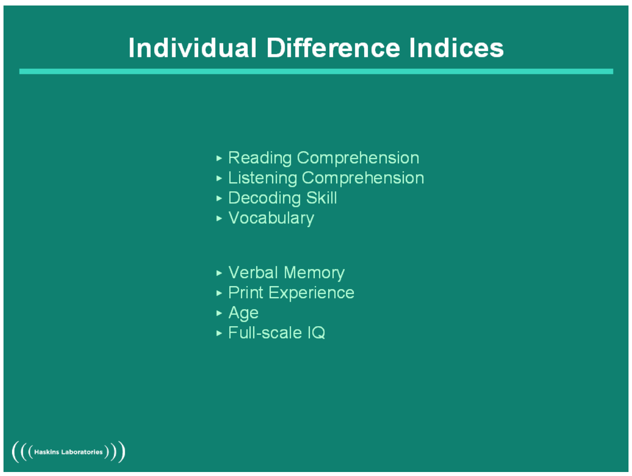
Slide **8**



Slide **9**

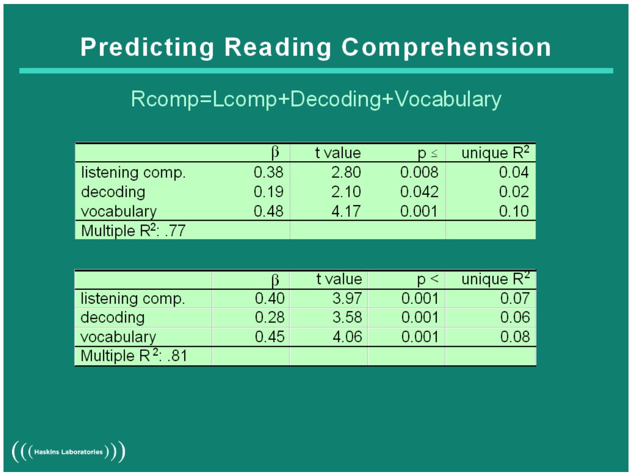
[slide 8] Here, we are looking at the fixation patterns of three different readers over the same sentence (a foil trial in our experiment). The center of each circle is a fixation point and the size of the circles is scaled to fixation duration. The color of the circles indicates the fixation starting time as an offset from the beginning of the trial. I want to emphasize that there is considerable variation among readers in the pattern (density) of fixations, in their durations, and in overall reading times. There is obviously a lot of complexity in the eye-movement record. [slide 9] A conventional way to simplify that complexity is by defining regions of interest within the sentences. In this example we might look at the summed fixation durations on a single noun phrase, “the sun”.

I want to make two points here. First, following much research, we assume that the pattern of eye-movements over a text reflects the immediate ebb and flow of effort expended in apprehending that text (for a review see Rayner, 1998). Second, we assume that much of the variation among individuals' eye-movement patterns (indexing moment by moment expenditures in effort) are explainable in terms of each person's linguistically relevant capacities. Others have looked at process measures like eye-movements as a function of underlying capacities, but the focus has been almost exclusively on measures of verbal memory (Daneman & Carpenter, 1980, and much subsequent work). We aim to cast a wider net.



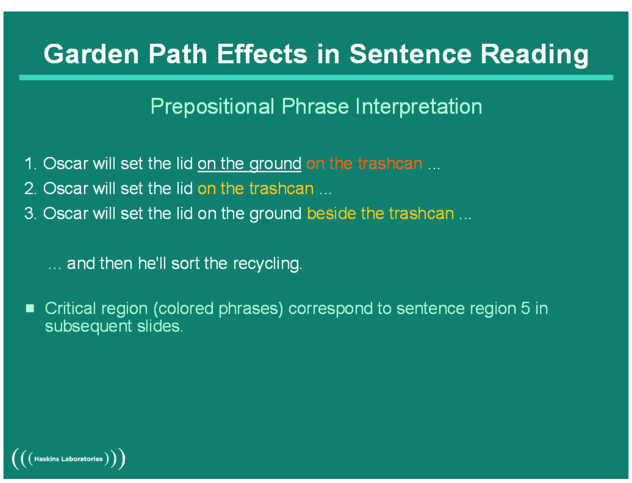
Slide **10**

[slide 10] So, let me say a little bit about the skill measures that we collected in addition to the eye-movement data. The most relevant are measures of reading and listening comprehension, decoding skill and vocabulary. Vocabulary was assessed through the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997) and decoding through the Woodcock Johnson III word-attack subtest (Woodcock et al., 2001). Reading comprehension was operationalized through a composite score derived from tasks of sentence-picture matching and comprehension of short narrative passages, while listening comprehension was assessed through a sentence picture matching task (details of the comprehension measures can be found in Braze et al., in press). Beyond these basic measures we also collected data on of verbal memory (sentence/listening span, Daneman & Carpenter, 1980), experience with print (magazine recognition test, Stanovich & Cunningham, 1992), full-scale IQ (The Psychological Corporation, 1999) and age, as well as others that I'm not going to get into today.



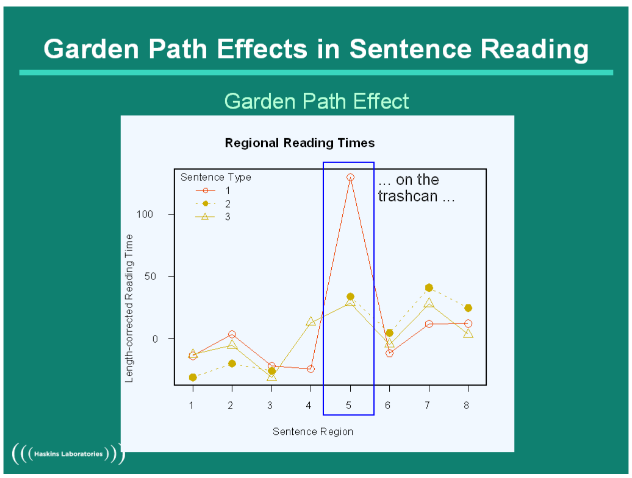
Slide **11**

[slide 11] Given that set of skill measures, I want to turn back to our question about the adequacy of the simple model, or rather, the way that it's typically operationalized. Our data indicate that, while there is considerable explanatory force in components L and D with regard to the reading comprehension score, word knowledge does pick up additional unique variance in reading comprehension, as predicted by the lexical quality hypothesis. The top table here shows those relationships for the present study, and the bottom table corresponds to a different cohort (N=44) from the same population (Braze et al., in press). So, the relationships seem to be pretty stable. I want to emphasize that this vocabulary supplemented simple model is the best model of reading comprehension given our entire battery of measures. Listening comprehension, decoding skill and vocabulary pick up all non-random variance in a <u>product</u>-measure of reading comprehension. However, the central question for today is how much variance in online reading behavior -- the eye-movement based <u>process</u> measure -- can we account for using these same predictors.

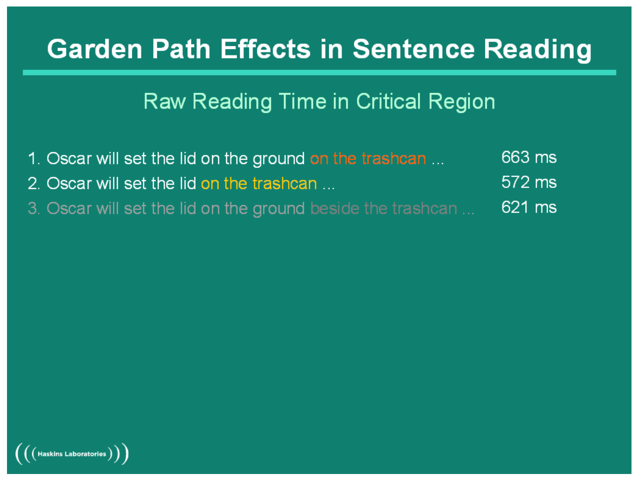


Slide **12**

[slide 12] The approach we take is to build sentence materials with carefully controlled levels of difficulty. We chose to use sentences with temporary ambiguities, garden-path sentences, for this purpose. Garden-path sentences contain an ambiguity where the initial interpretation turns out to be wrong and must be revised. Sentence type 1 is a garden-path sentence in which the underlined phrase “on the ground” is initially misconstrued to be the place where the lid is “set”, the goal of the verb. Upon encountering the phrase “on the trashcan” the reader becomes aware of the error. The need to revise that initial misinterpretation would typically result in longer reading times at that second prepositional phrase, relative to the similar phrases in the non-garden-path control sentences 2 and 3. But, note the length difference for the critical prepositional phrases (“beside” in sentence type 3). For this reason our initial analysis will rely on length corrected regional reading time. We expect longer reading times on the prepositional phrase that includes “the trashcan” in sentence type 1 than in sentence types 2 and 3.



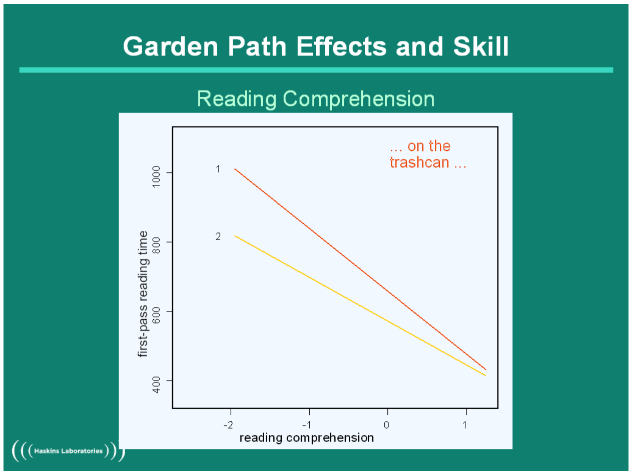
Slide **13**



Slide **14**

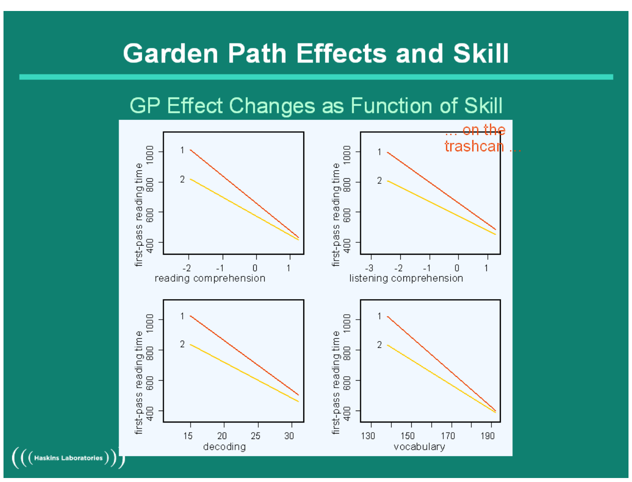
[slide 13] The data support this prediction. The critical phrase is plotted as sentence region 5. It’s clear to see that regional reading times for garden-path sentences are much longer than for the control sentences. The garden-path effect is robust even if we use uncorrected reading times as the dependent measure, with the garden-path condition having longer reading times than both control conditions (marginally so for the difference between sentence types 1 and 3) [slide 14].

Of course, our primary interest is in the relationship between the garden-path effect shown here, and our predictors of reading comprehension, L, D and V. For this purpose, we'll focus on the difference between sentence types 1 and 2 only, which are identical in the critical region, and use uncorrected regional reading times in subsequent analyses.



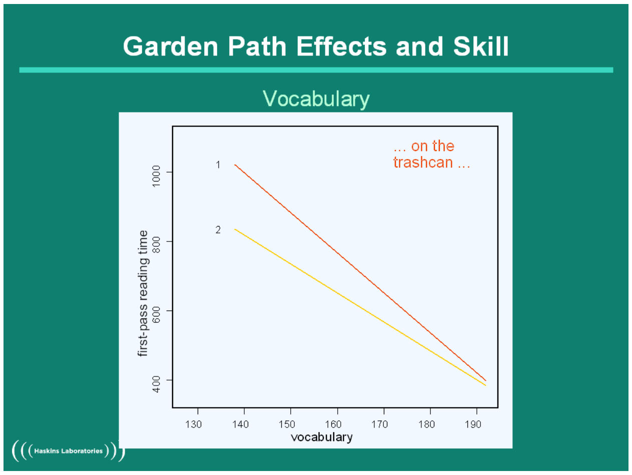
Slide **15**

[slide 15] Here we see the relationship between the garden-path effect and our index of reading comprehension. Regional reading time for the critical prepositional phrase is on the y axis. Two things are apparent from this plot. First, and unsurprisingly, poor comprehenders read more slowly; there is a robust main effect of comprehension level. More interesting is the interaction between comprehension and the strength of the garden-path effect, the difference in reading times between garden-path sentences and non-garden-path sentences. Poor comprehenders show a larger garden-path effect than good comprehenders.



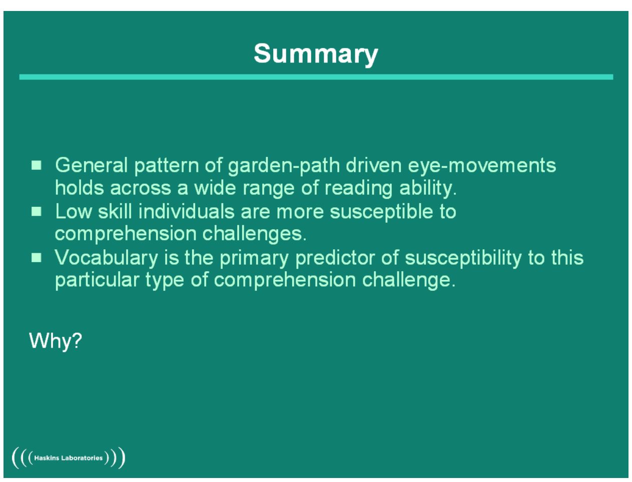
Slide **16**

[slide 16] Of course, we want to see what components of the vocabulary supplemented simple model of reading are actually responsible for this relationship. Here we see reading times for each sentence type plotted against indices of reading comprehension (same as slide 15), listening comprehension, decoding skill and vocabulary knowledge. Each dimension of reading skill has a similar relationship to overall reading times at the critical phrase as well as showing a similar interaction with the garden-path effect; better scores along any of these dimensions correspond to a smaller garden-path effect. Ultimately, the question we want to address is where among the skill measures do we find the most potent predictors of online reading behavior -- our measure of the comprehension process.



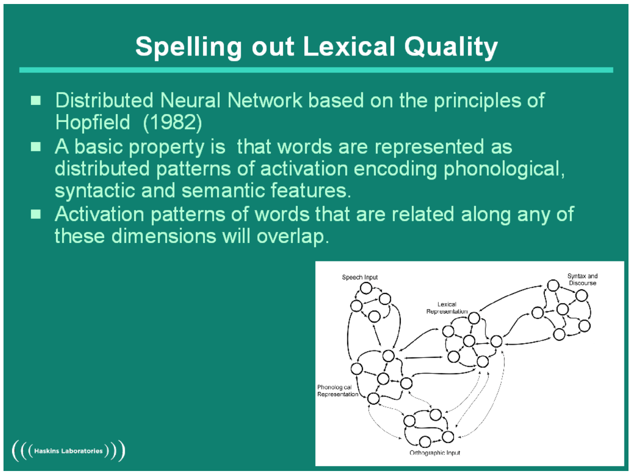
Slide **17**

[slide 17] If we consider the <u>partial</u> contributions of listening comprehension, decoding skill and vocabulary to reading times on the critical prepositional phrase, each after taking the other two into account, we find that neither listening comprehension nor decoding skill accounts for any unique variance in the interaction with sentence type, while the partial effect of vocabulary does have a marginally significant interaction with sentence type. Vocabulary is the sole predictor of unique variance in the magnitude of the garden-path effect. Further, this is true even if we partial verbal memory as well as decoding and listening comp.



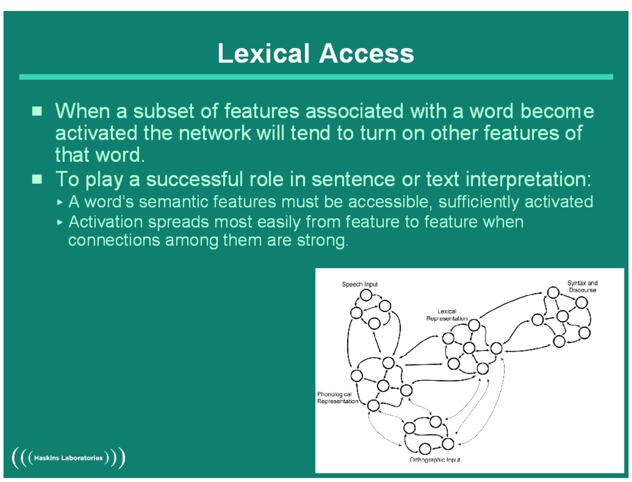
Slide **18**

[slide 18] So to sum up, we can report that the pattern of garden-path driven eye movements holds for individuals across a wide range of reading ability. What's more, low-skill individuals show a larger garden-path effect than high-skill individuals; they are more susceptible to this particular type of comprehension challenge. Finally, vocabulary seems to be the most potent predictor of susceptibility to this particular type of comprehension challenge. How can the lexical quality hypothesis help us to understand these facts?



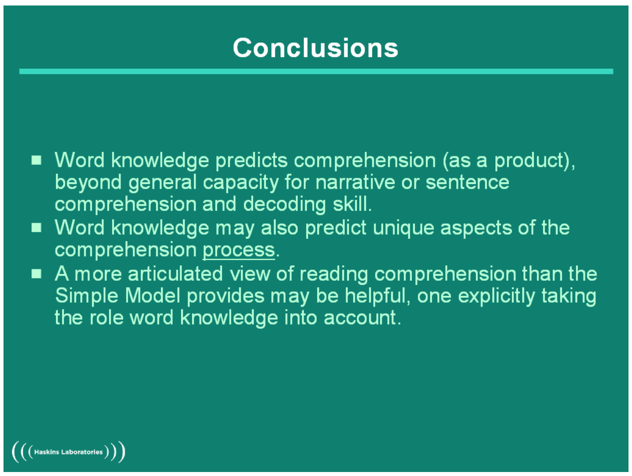
Slide **19**

[slide 19] In earlier work, we proposed a specific implementation of lexical quality which views the lexicon as a neural network (Hopfield, 1982). This can be seen as a variation on the Seidenberg and McClelland "Triangle Model" of lexical representation (Seidenberg & McClelland, 1989). A basic property of this kind of model is that words are represented as distributed patterns of activation encoding phonological, syntactic and semantic features. The activation patterns of words that are related along any of these dimensions will overlap. In this slide, bold lines indicate strong connections between auditory input and phonology, while faint lines indicate weaker connections between orthographic input and phonology, a situation that we take to hold for many readers, perhaps all but the most accomplished.



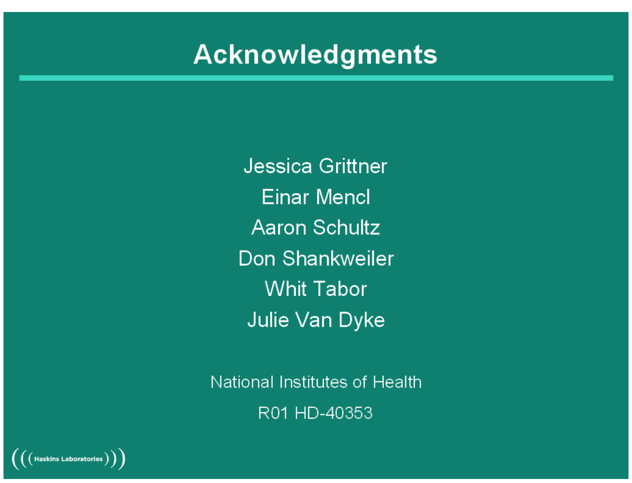
Slide **20**

[slide 20] When a subset of features associated with a word become activated, phonological features for example, the network will tend to turn on other features of that word, its semantic features for instance. In order for a word to play a useful role in sentence interpretation, its semantic features have to be accessible. In connectionist terms, the word must be sufficiently activated. Our focus here is on the contribution of well-developed lexical knowledge to lexical access in the face of less than optimal external activation of lexical representations. This will include words taken in by eye by less than fully accomplished readers. Marginal activation of words impacts not only the apprehension of word meaning, but also the ability to integrate those words into representations of phrase, sentence and text meaning.



Slide **21**

[slide 21] Vocabulary, word knowledge, is a potent predictor of static measures of reading comprehension, the product of comprehension. This doesn't necessarily undermine the conceptual validity of the simple model. But, it does suggest that no one index of L, those capacities necessary to the comprehension of speech, is sufficient to operationalize that component of the simple model. Word knowledge also predicts, perhaps uniquely, aspects of the comprehension process related to the integration of meaning across words. A broader assessment of L than is typical, including at least a vocabulary measure, may be helpful both in the context of the simple model of reading and in attempting to understand the process of comprehension.



Slide **22**

[slide 22] Finally, I want to acknowledge my collaborators on this project. It is very much a team effort.

References

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

Daneman, M., & Carpenter, P. A. (1980). Individual Differences in Working Memory and Reading. Journal of Verbal Learning and Verbal Behavior, 19(4), 450-466.

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

Gough, P. B., & Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. RASE: Remedial & Special Education, 7(1), 6-10.

Hopfield, J. J. (1982). Neural Networks and Physical Systems with Emergent Collective Computational Abilities, Proceedings of the National Academy of Sciences (79, pp. 2554-2558).

Perfetti, C. A., & Hart, L. (2002). The Lexical Quality Hypothesis. In L. Verhoeven (Ed.), Precursors of Functional Literacy (pp. 189-213). Philadelphia, PA: John Benjamins.

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. Psychological Bulletin, 124(3), 372-422.

Seidenberg, M. S., & McClelland, J. L. (1989). Visual Word Recognition and Pronunciation: A Computational Model of Acquisition, Skilled Performance, and Dyslexia. In A. M. Galaburda (Ed.), From Reading to Neurons (Issues in the Biology of Language and Cognition Vol. pp. 255-303). Cambridge, MA: Mit Press.

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.

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

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