Modeling Concept Activation in Working Memory during Online Sentence Processing

Patrick Plummer (pplummer@ucsd.edu)
Hsueh-Cheng Wang (hchengwang@gmail.com)
Yuhtsuen Tzeng (ttcytt@ccu.edu.tw)
Marc Pomplun (marc@cs.umb.edu)
Keith Rayner (krayner@ucsd.edu)

Department of Psychology, University of California at San Diego, USA

Department of Computer Science, University of Massachusetts at Boston,
100 Morrissey Boulevard, Boston, MA 02125 USA
Center for Teacher Education & Institute of Curriculum, National Chung Cheng University, Taiwan

Abstract

There have been several computational alternatives to the cloze task (Taylor, 1953) intended to approximate word predictability effects on eye movements during reading. In this study, we implement a computational model that instantiates each content word in a sentence as an input that activates semantic concepts in working memory. The predictability of a word is then determined by the extent to which its corresponding semantic representation is associated with the network of concepts already active in working memory from the preceding context. The computation of concept activation is based on a connectionist model (Landscape model, see van den Broek, 2010). Latent semantic analysis (LSA) is used to establish connections between words and simulate the long-term semantic associations among concepts (Landauer & Dumais, 1997). This model provides a means of investigating how language comprehension and eye movement behavior are affected by the activation of concepts in working memory.

Keywords: eye movements; reading; word predictability; latent semantic analysis; Landscape model.

Introduction

It has been well-established that eye movement behavior is affected by lexical variables such as frequency and predictability (Rayner, 1998; 2009). As such, the eye movement record provides an indication of language processing as it unfolds during normal reading. Rayner and Well (1996; see also Ehrlich & Rayner, 1981) found that the predictability of target words had a strong influence on eye movements during reading. In their experiment, subjects fixated unpredictable target words longer than either highly or moderately predictable target words; highly predictable words were also skipped more often than moderately predictable or unpredictable target words.

Accordingly, in the E-Z Reader model (Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999; 2003), word predictability within a given sentence context is considered in both first stage processing (i.e., L1, including identification of orthographic form and a familiarity check) and second stage processing (i.e., L2, including identification of phonological/semantic form and completion of lexical access). The model also maintains that the predictability effect is stronger in L2 than in L1.

Estimates of word predictability are typically derived from a modified cloze task procedure (Taylor, 1953) in which subjects are asked to guess the identity of a word when given the prior sentence context. Most reading studies utilize the cloze task to establish or confirm word predictability manipulations. These experiments use target words that differ substantially in cloze value (the probability with which subjects select the word), often with probabilities of .70 to .90 for highly predictable words and less than .10 for "low" predictability words. As an alternative to necessarily subjective cloze responses, several computational methods have been successfully utilized to approximate degrees of contextual constraint and predict the influence on eye movements during reading; including, transitional probabilities (McDonald and Shillcock, 2003; but see Frisson, Rayner, & Pickering, 2005), surprisal (Boston, Hale, Kliegl, Patil, and Vasishth, 2008; Levy, 2008), conditional co-occurrence probability (Ong and Kliegl, 2008). Additionally, Latent Semantic Analysis (LSA) (Landauer & Dumais, 1997) was used by Pynte, New, and Kennedy (2008) as well as Wang, Chen, Ko, Pomplun, and Rayner (2010), who reported that eye movement behavior during first-pass reading on content words could be predicted using LSA. McDonald and Shillcock (2003) and Wang et al. (2010) used the transitional probability (corpus-based statistical likelihood of encountering a word given the preceding or subsequent word) to categorize predictability conditions; both proposing that predictability effects could be accounted for using only the content word preceding a target. One limitation of these objective measures could be that prior context, before the immediately preceding lexical item, may affect processing of a word in many instances. Wang et al. (2010) also used *all* concepts in the preceding sentence context to compute contextual constraint for targets using the standard weighting from LSA. However, without a clearer understanding of working memory constraints

during comprehension it is difficult to make regarding semantic constraint.

The predictability of a given word can, in large part, be conceptualized as the degree to which the semantic concept represented by the word is associated with the context of preceding lexical items. By treating incoming lexical items as semantic concepts that interactively influence working memory processes, prior context for a word can be represented as inputs which influence the activation of associated concepts and have the potential to facilitate or inhibit the processing of upcoming words. As a result, the higher the activation of a concept when it is encountered, the more processing of the concept is facilitated. Importantly, individuals can allocate their processing attention to only a finite number of linguistic items at a given moment. Thus, any model of language processing and working memory must establish limits to the number of lexical-semantic concepts that can be simultaneously active and exert an appreciable influence on the processing of upcoming lexical inputs.

A Connectionist Model for Sentence Reading

This study proposes a computational model to monitor the activations of concepts in working memory. The computation of concept activation is derived from a connectionist model (the Landscape model, see van den Broek, 2010). The current model is not connectionist in the sense of having distributed semantic representations; rather, words are represented as localized semantic "concepts" with weighted connections to a network of additional concepts. The semantic connections among concepts in the simulation are computed using LSA cosine values based on the default 300 dimension semantic space, "general reading up to 1st year college", available at the LSA@CU Boulder website (http://lsa.colorado.edu/). LSA represents word meaning and computes associations by applying a linear algebra method, singular value decomposition (SVD), to a large corpus of text (see Landauer & Dumais, 1997).

The Landscape model is a connectionist approach to instantiating comprehension using psychologically plausible algorithms that can potentially be used to model several aspects of text comprehension (see van den Broek, 2010; Tzeng, van den Broek, Kendeou, & Lee, 2005). The architecture of the conventional Landscape model assumes that as a reader proceeds through a text in reading cycles (with each cycle roughly corresponding to the reading of a new sentence), concepts fluctuate in activation as a function of four sources of information: the current processing cycle, the preceding cycle, the current episodic text representation, and reader's background knowledge. With the reading of each cycle, particular concepts are activated and added as nodes to the episodic memory representation of the text. If a concept is already part of the text representation and is reactivated, its trace is strengthened. Furthermore, co-

activation of concepts leads to the establishment (or strengthening) of connections between those concepts. The resulting network representation influences subsequent activation patterns. This phenomenon is called the *cohort effect*. These cyclical and dynamically fluctuating activations lead to the gradual emergence of an episodic memory representation and discourse model of the text, in which textual propositions and inferences are connected via semantic relations (such as causal and referential links). Thus, the model captures the fluctuations of concepts during reading (Linderholm, Virtue, Tzeng, van den Broek, 2004), as well as readers' memory representation of text (Tzeng, 2007). As such, this model has prescribed mechanisms that can link the iterative and reciprocal relations between fluctuations of activations and the episodic text representation. However, there are necessary differences with regard to how readers generate and update active discourse representations for the comprehension of an individual sentence, compared to the processing of a longer narrative or expository text. For the comprehension of an individual sentence, a reader must primarily rely on establishing connections between relevant concepts in working memory and pre-existing long-term semantic representations. For a longer text, on the other hand, readers are often able to take advantage of more extensive and detailed context and presumably a more enriched discourse model. Thus, the current computational approach adapts the Landscape Model to a connectionist framework more suitable for capturing sentence reading. The current model utilizes LSA in order to represent pre-existing connections between semantic representations stored in long-term memory (i.e., background or world knowledge).

In the current model, as with the Landscape Model, text inputs are represented by an *input matrix* and each is indexed as a *Mention* (concepts being read from the text). The conventional Landscape model also defines other sources of activation including *Referential* (for building referential coherence), *Causal*, and *Enabling* (for the causal explanation of the current statement), but those activations are as of yet, not implemented here. The input matrix for example sentence: "*The knight uses his sword to fight the dragon*" is shown in Table 1.

Table 1: Input matrix for the *Knight* example.

cycle	knight	use	sword	fight	Dragon
1	1	0	0	0	0
2	0	1	0	0	0
3	0	0	1	0	0
4	0	0	0	1	0
5	0	0	0	0	1

Initially, the sentence is segmented into component concepts: "knight", "use", "sword", "fight", and "dragon"; as, currently, only content words are considered as concepts. The model assumes that each word is fixated and processed

sequentially. In each cycle, the concept of *Mention* receives 1 unit of activation. In addition to the sequential activation of concepts, the influence of semantic knowledge and pre-existing lexical associations between concepts is established using LSA corpus-learned associations. Table 2 presents the *connection matrix* for the example sentence. The values are always between -1 and 1, but are rarely below 0 because of LSA's high-dimensional space.

Table 2: Connection matrix for the example.

.	knight	use	sword	fight	dragon
knight	1	.01	.64	.15	.28
use	.01	1	.03	-.02	.06
sword	.64	.03	1	.20	.40
fight	.15	-.20	.20	1	.13
dragon	.28	.06	.40	.13	1

The activation values for each concept are represented in an m x n *activation matrix*, where m represents the number of concepts in the sentence and n represents the number of *cycles*. Each column in the matrix thus represents the status of each concept. The *activation matrix* takes each column of the *input matrix* as raw input and processes it row by row. In our model, the activation during the current reading cycle is defined by Equation (1):

$$A_i^{cycle} = \sum_{j=1} \delta A_j^{cycle-1} \sigma(S_{ij}) + \sum_{j=i} input_i^{cycle} \sigma(S_{ij}) \qquad (1)$$

A_i^{cycle} is the activation of concept *i* during the current cycle. Starting from the summation (Σ) term in Equation (1), for all activated concepts in the previous reading cycle, each activation value is multiplied by a transformation function σ of connection strength (S_{ij}) and by the cohort activation parameter δ. S_{ij} is the strength of the relation from concept *j* to *i*. For the current cycle, $input_i^{cycle}$ is the activation of concept *i* in the *input matrix*. The sum of the net inputs for these *m* concepts is multiplied by the transformation function σ of connection strength (S_{ij}).

The conventional Landscape model uses a sigmoid function as the transformation function σ to control the possible linear growth of spreading of activation and limit the effect of cohort activation to those strongly related to the concept. Since S_{ij} is usually between 0 and 1, a linear function with absolute value is used in this model. The value of the cohort activation parameter, δ, directly determines the amount of cohort activation and in the future can be used to mimic individual differences in the spreading of activation. Our model assumes that for any concept, its cohort activation can never exceed its input activation. For this reason the model will take the largest of the input and cohort activation values, and *Mention* is the maximum activation a concept can receive. Furthermore, the system parameter- *Activation Threshold* sets any activation below a set threshold to zero.

The working memory constraint is implemented by a parameter *WMC (Working Memory Capacity)*. When the actual sum of activation exceeds the value of WMC, the activation of each concept is scaled down using Equation (2):

$$A_i^{cycle} = A_{i,Actual}^{cycle} \cdot \frac{WMC}{\sum_{i=1}^{m} A_{i,Actual}^{cycle}} \qquad (2)$$

For the example sentence, the activation matrix is shown in Figure 1. For the 1st cycle (in which "*knight*" is processed during first-pass reading), the activation of "*knight*" is 1, from the *Mention* input. There is no cohort effect for the first reading cycle since no previous cycle exists. The activations for "*use*", "*sword*", "*fight*", and "*dragon*" are established by multiplying their connections, .01, .64, .15, and .28 respectively, and the input of "*knight*" (1). The activation of "*use*" does not reach the threshold (set to 0.1) and as a result receives an activation of 0. For the 2nd cycle when "*use*" is being processed, the activation of each concept is calculated according to Equation (1). Figure 1 illustrates that the activation of "*dragon*" stays relatively high from cycle 1 to cycle 4 because of relatively strong connections to "*knight*", "*sword*", and "*fight*." Conversely, the activation of "*use*" decreases from cycle 2 to 5 because of relatively weak connections to "sword", "fight", and "dragon" (less than .06).
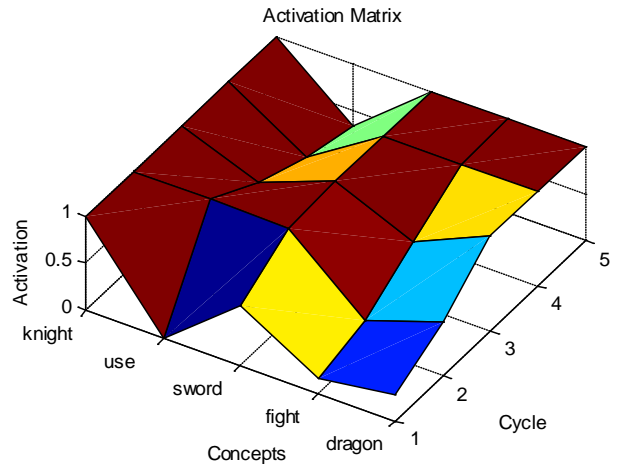

Figure 1: The "landscape" of the activation matrix for the *Knight* example.

The conventional Landscape model updates the connection strengths in its episodic memory using a learning algorithm in order to adjust active discourse representations for the comprehension of a longer text. In this study, we assume that the background knowledge (represented by the connection matrix) is not altered during sentence reading.

In summary, by assuming (a) that words in a sentence are read and processed sequentially, and (b) long-term memory

representations (i.e., background knowledge) are not affected during comprehension; we propose a computational model of sentence processing which takes advantage of an existing discourse comprehension model designed to take into account contextual effects. The proposed model allows us to examine several factors that affect sentence comprehension; namely, (1) semantic activation in working memory, (2) background knowledge, and (3) working memory capacity. To assess the model's ability to reflect linguistic processing we will compare its performance to the cloze task.

Experiment: Reanalysis of Previous Data

The key objective for this implementation is to disambiguate high from low semantic constraint in sentence contexts. Another objective of this implementation is to demonstrate that the LS model surpasses previously utilized methods as an alternative to the cloze task. In order to demonstrate that the proposed computational model is capable of matching cloze results more accurately than previous approaches, i.e., Wang et al. (2010), we re-analyzed the materials in Gollan, Slattery, Goldenberg, Van Assche, Duyck, and Rayner (2011), in which predictable/unpredictable target words were determined by a norming cloze task. We estimated predictability of a target word by (1) the previous content word, (2) all words in prior context, and (3) the estimates of the proposed connectionist model in this endeavor. We expect that our model can outperform other predictors on differentiating high- and low-constraint contexts and generate higher correlation to cloze values.

Participants. Twenty undergraduate students at the University of California, San Diego, participated. All of them were native speakers of English.

Materials. There were 90 target words; all target words were embedded in either a high-constraint (HC) or low-constraint (LC) sentence. For example, "*the hockey player moved on the ice on his _____*" (S1) was considered HC while "*The little girl was very happy when she unwrapped her brand new _____*" (S2) was LC for the target "skates". The target words in HC context were generated 87% of the time, whereas the ones in LC context were generated less than 3% of the time.

Procedure. Participants were presented with the sentences up to the target words, and asked to provide one-word continuations for each sentence.

Analysis. The first estimate of predictability of each target word was derived by extracting the LSA connection weight to the previous content word (PreCont) for each target, e.g., the previous content word of S1 is "ice," while the one of S2 is "new." The second approach computed the LSA cosine value using all words in the previous context (AllW). The final estimate was derived from Landscape

model of sentence processing described above in the previous section (LS). We manually segmented the sentence into concepts and removed function words such as "a", "the", "in", etc., for instance, "hockey / player / moved / ice" for S1. The parameters of our model were set as following: $\delta = .7$, *Mention* = 1, *Activation Threshold* = .1, and *WMC* = 7. The averages and standard deviations of Cloze, PreCont, AllW, and LS for HC and LC are described in Table 4.

Table 4. The averages and standard deviations (in parentheses) of Cloze, PreCont, AllW, and LS for HC and LC conditions.

	Cloze	PreCont	AllW	LS
HC	.87 (.13)	.17 (.16)	.21 (.16)	.66 (.29)
LC	.03 (.03)	.05 (.11)	.04 (.07)	.13 (.20)

Results

As shown in Figure 2, an operating characteristic (ROC) analysis demonstrates that the area under the curves (AUC) of Cloze, PreCont, AllW, and LS are 1, .70, .87, .91, respectively. The LS model obtains a higher AUC than AllW or PreCont. Furthermore, a correlation analysis demonstrates that the Pearson correlations between Cloze and PreCont, AllW, and LS are .39, .56, and .70, respectively.
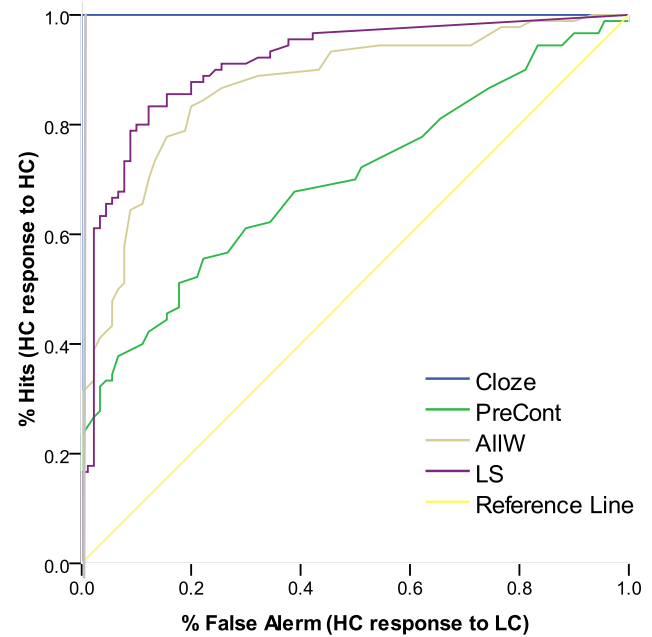

Figure 2: ROC curves for Cloze, PreCont, AllW, and LS.

The results suggest that the LS model can simulate much of the linguistic processing subjects perform when

producing cloze responses (and presumably during normal reading). The current objective is not to match cloze probabilities per se, but to successfully demonstrate the model's ability to differentiate highly constrained and unconstrained sentence contexts as well as the conventionally used cloze task. The LS model also demonstrates superiority over objective measures that utilize only the prior content word or LSA connections between content words exclusively.

Discussion

The current implementation of the model has demonstrated that it is an effective measure of contextual constraint; in that it differentiates high and low-constraint sentence contexts better than previously employed alternatives to the cloze task. Furthermore, model activations for target words correlate with cloze responses more highly than previous objective methods of measuring contextual constraint. We believe this is an initial step toward the ultimate objective of representing both the fluctuating activation of lexical-semantic concepts in working memory during online sentence processing and how the processing of upcoming words can be facilitated by prior context. Discourse-mediated spreading activation across lexical-semantic representations has been proposed as an appreciable source of predictability effects during reading (Morris, 1994; Pynte et al., 2008; Traxler, Foss, Seely, Kaup, & Morris, 2000). Thus, modeling the process whereby linguistic inputs activate concepts in long-term memory and continuously influence working memory operations during sentence comprehension is an important endeavor in psycholinguistics.

As shown by the comparison to standard cloze responses, the current model can be used to reliably derive predictability of word *n* given the preceding context. The model generates a specific level of activation for word *n*, assuming that each word in the preceding context has been identified and all associated concepts have been engaged in working memory. As demonstrated above, this predicted level of activation correlates to cloze probabilities for a target word (*n*).

Critically, the LS Model is able to reliably differentiate high and low constraint sentence fragments. Moreover, when using the LS model, in many cases the level of activation for word *n* will provide a more psychologically realistic measure of word processing difficulty when compared to cloze proportions, especially in neutral or unconstrained contexts. For instance, referencing cloze scores alone, there is no distinction between words that are plausible, yet not highly-predictable, and those that are completely implausible or anomalous given the preceding sentence context. In fact, it is quite feasible for plausible target words in unconstrained sentence frames to receive cloze probabilities at or around zero; however, low cloze probabilities are not necessarily indicative of potential processing difficulty. The manner in which the cloze task is conventionally used produces binary measurements (to the

extent that non-target responses are ignored). In this way, the current computational model may produce a more accurate representation than cloze scores with regard to indexing online word processing difficulty. This is particularly true for low constraint sentence frames. As such, the next logical step is to assess the LS model's goodness-of-fit to reading times and other eye movement data.

By modifying the framework of the conventional Landscape Model to reduce the size of text segments being processed during a reading cycle and situating activated concepts within limited working memory resources, we have attempted a psychologically plausible computational model of semantic effects on sentence comprehension. Crucially, the fluctuating activation of within sentence concepts is not determined merely by summing its cumulative activation across all preceding words; rather, the interactive and co-dependent influence of the prior sequence of words determined the extent to which the prior sentence context results in activation for a particular lexical-semantic concept.

The model is also a useful tool for investigating the number of semantic entities that are generally active in working memory. As well as the upper limits for the number of lexical-semantic entities simultaneously activated. Computationally examination of working memory limitations during reading could provide insight into what linguistic constructions are likely to elicit processing difficulty, result in longer fixation times, and lead to inter-word regressions during sentence reading. Model outputs can also be used to make predictions as to which concepts are likely to maintain relatively high levels of activation in working memory.

While among the most sophisticated computational frameworks in the field of cognitive science, current models of eye movement control during reading do not focus on how prior words render specific words predictable. The more well-developed models of oculomotor behavior and language comprehension represent the predictability of a given word in a sentence using only its cloze probability (Engbert, Nuthmann, Richter & Kliegl, 2005; Reichle et al., 1999; 2003). Our model successfully attempts to represent the cognitive processes that are sensitive to semantic constraint. Future implementations of the LS model will be capable of more thoroughly examining aspects of language processing and eye movement behavior. The *connection matrix* in the LS model can operationalize a variety of linguistic characteristics stored represented in long-term memory. Semantically-based connection weights can be modified to accommodate mediation by parafoveal preview information. In addition, the *connection matrix* could be modified to capture morphological, orthographic, and phonological relationships between lexical items. Currently, the LS model is a computational alternative to the cloze that is sensitive to both strong and subtle changes in contextual semantic constraint; as shown by the reasonable activation of plausible words in low constraint sentence

frames. Ultimately, the model will be expanded in an effort to achieve more comprehensive measurement of lexical-semantic predictability as it affects reading behavior.
.

Acknowledgments

References

Boston, M. F., Hale, J., Kliegl, R., Patil, U. & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, 2(1):1, 1-12.

Deerwester, S., Dumais, S., Furnas, G., Landauer, T., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Sciences*, 41, 391–407.

Ehrlich, S.F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*, 641-655.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, 112, 777-813.

Gollan T. H., Slattery T. J., Goldenberg D., Van Assche E., Duyck W., Rayner K. (2011). Frequency Drives Lexical Access in Reading but Not in Speaking: The Frequency-Lag Hypothesis. *Journal of Experimental Psychology: General*,140, 2, 186–209

Jones, M. N. & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, 114, 1-37.

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,* 211–240.

Landauer, T. K., McNamara, D. S., Dennis S., & Kintsch W. (2007). *Handbook of Latent Semantic Analysis*, Lawrence Erlbaum Associates.

Levy, R. (2008). Expectation based syntactic comprehension. *Cognition 106*, 1126- 1177.

Linderholm, T., Virtue, S., Tzeng, Y., & van den Broek, P. W. (2004). Fluctuations in the Availability of Information during Reading: Capturing Cognitive Processes using the Landscape Model. *Discourse Processes*, 37(2), 165-186.

McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science, 14*, 648–652.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of experimental psychology: Learning, Memory, &Cognition*, *20*, 92–103.

Ong, J. K. Y. & Kliegl, R. (2008). Conditional co-occurrence probability acts like frequency in predicting fixation durations. *Journal of Eye Movement Research*, 2(1):3, 1-7

Pynte, J., New, B. & Kennedy, A. (2008). A multiple regression analysis of syntactic and semantic influences in reading normal text. *Journal of Eye Movement Research*, 2(1):4, 1-11

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

Rayner, K. The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention during reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62, 1457-1506.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3*, 504–509.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105, 125-157.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26,* 445–476.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research*, 39, 4403-4411.

Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly*, 30, 415–433.

Traxler, M. J., Foss, D. J., Seely, R. E., Kaup, B., & Morris, R. K. (2000). Priming in sentence processing: Intralexical spreading activation, schemas, and situation models. *Journal of Psycholinguistic Research, 29*, 581-594.

Tzeng, Y. (2007). Memory of narrative texts: How parts of Landscape model work. *Chinese Journal of Psychology*, *49* (3), 1.-25.

Tzeng, Y., van den Broek, P., Kendeou, P., & Lee, C. (2005). The computational implementation of the landscape model: Modeling inferential processes and memory representations of text comprehension. *Behavior Research Methods*, 37(2), 277-286.

van den Broek, P. (2010). Using texts in science education: cognitive processes and knowledge representation. *Science*, 328, 453.

Wang, H. C., Pomplun, M., Ko, H. W., Chen M. L., & Rayner, K. (2010). Estimating the effect of word predictability on eye movements in Chinese reading using latent semantic analysis and transitional probability, *Quarterly Journal of Experimental Psychology*, 63, 1374-1386.