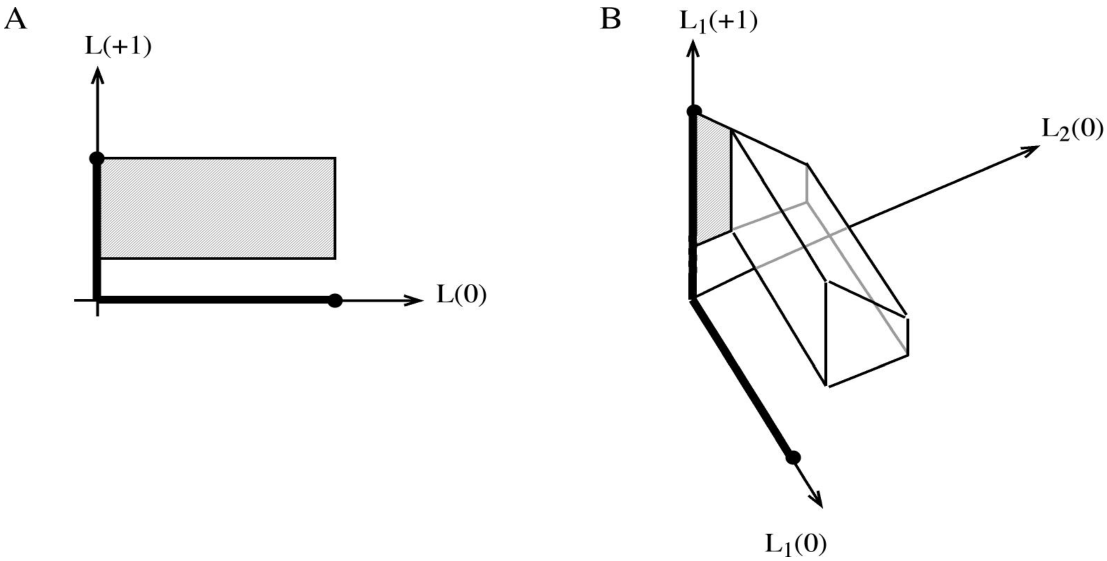

Figure 1 (Hanes & McCollum). A. Values of the pair (L(0), L(+1)) that can determine a fixation duration in Morrison's model. B. Values of the vector ($L_1(0)$, $L_2(0)$, $L_1(+1)$) that can determine a fixation duration in the E-Z Reader model. L(0) and L(+1), lexical processing time on fixated and following words, respectively; $L_1(0)$ and $L_1(+1)$, durations of first stage of lexical processing on fixated and following words, respectively; $L_2(0)$, duration of second stage of lexical processing on fixated word. Enclosed three-dimensional regions, shaded planar regions, and thick line segments are included in the graph.

Visual word recognition and oculomotor control in reading

Lynn Huestegge[a], Jonathan Grainger[b], and Ralph Radach[a]
[a] *Institute of Psychology, Technical University of Aachen, 52056 Aachen, Germany;* [b] *Université d'Aix-Marseille, LPC/CNRS, 13 621 Aix en Provence, France.* **lynn.huestegge@post.rwth-aachen.de**
grainger@up.univ-mrs.fr **ralph@psych.rwth-aachen.de**

Abstract: A central component in the E-Z Reader model is a two-stage word processing mechanism made responsible for both the triggering of eye movements and sequential shifts of attention. We point to problems with both the verbal description of this mechanism and its computational implementation in the simulation. As an alternative, we consider the use of a connectionist processing module in combination with a more indirect form of cognitive eye-movement control.

The E-Z Reader suite of computational models is characterized by the role played by lexical processing in the complex system that guides the eyes through text during reading. The word identification module of E-Z Reader 7 is seen as the "engine" that drives the whole process, determining both the dynamics of saccade generation and the assumed sequential shifting of attention. Although the section on word identification occupies only a modest proportion of the target article, there is no denying that this is a central component of the model.

As argued by Grainger (2003), the time is ripe for a fruitful interaction between research on visual word recognition and research on eye movements in reading. Including a word identification component in a model of eye-movement control in reading is already a significant step in the right direction, and Reichle et al. are to be congratulated for their pioneering work towards this goal. However, there are several ways to go about generating such an interaction. One is to examine how a given model of visual word recognition, motivated by research using isolated word presentation techniques, could be integrated into a more global reading system that includes oculomotor control. Another way is to define a minimalist model of visual word recognition that, when coupled with a model of oculomotor control, optimally fits the data collected using eye-movement paradigms. Reichle et al. have adopted the latter approach, with some unfortunate consequences.

E-Z Reader 7 implements a two-stage approach to word identification. An early stage (L_1) is assumed to play a crucial role in the identification of the orthographic form of the word, whereas a second stage (L_2) is rather related to phonological and semantic processing. Completion of the first stage initiates the preparation of a saccade to the next word, and completion of the second stage initiates an attention shift to the next word. However, an examination of the verbal description of this part of the model and its mathematical implementation reveals a number of problems with this approach.

As the authors note, the major motivation for the two stages in the word identification component of E-Z Reader 7 is the decoupling of eye movements and attention in the model, and the fact that this allows E-Z Reader to capture a wide range of eye-movement data. In the mathematical implementation of E-Z Reader 7, it is apparent that only two empirically observed variables are used to model the various types of processing associated with each stage: word frequency and predictability (see equations for $t(L_1)$ and $t(L_2)$ below). This is because the model was designed to capture the influence of word frequency and predictability on the various measures obtained in eye movement recordings.

$$t(L_1) = [\beta_1 - \beta_2 \ln(\text{frequency})]\ (1\text{-}\theta \text{ predictability})$$

$$t(L_2) = \Delta[\beta_1 - \beta_2 \ln(\text{frequency})]\ (1\text{-predictability})$$

Two problems come immediately to mind. First, the only difference between these two processing stages concerns the relative weight assigned to frequency compared to predictability. In the first stage, the influence of predictability is reduced and that of word frequency enhanced, compared with the second stage. Second, the relationship between these variables and the orthographic, phonological, and semantic processing described in the verbal model, is left unspecified.

The verbal model is said to be partly motivated by the activation-verification model (AVM) of Paap et al. (1982). It is true that the AVM can be described as a two-stage (activation and verification) model of visual word recognition, but the analogy between

https://doi.org/10.1017/S0140525X03320102 Published online by Cambridge University Press

the AVM and the word identification component of E-Z Reader stops here. Worse, the precise structure of the AVM was generated to account for experiments showing no frequency effects on early orthographic processing in perceptual identification paradigms. E-Z Reader does just the opposite, by pronouncing frequency effects in the first stage of lexical access. The same is true for predictability. In the AVM, predictability as well as word frequency are thought to influence the second stage of processing exclusively. This is in contradiction to the implementation in E-Z Reader, where effects of predictability on stage L_1 are necessary to account for both word skipping and reduced fixation times for predictable words, since the completion of this early stage triggers eye-movement behavior. While it seems reasonable to assume that predictability may reflect top-down influences from higher-level sentential or text-level representations in the later phases of word identification, the authors need to provide independent evidence for such influences on early orthographic processing.

Furthermore, the option to put orthographic processing in one stage and phonological and semantic processing in the other stage appears to be totally arbitrary. Why not have three separate stages, or put orthography and phonology together (as representations of form), separate from semantics? Indeed, the authors' own presentation of parafoveal preview effects would motivate a regrouping of orthographic and phonological processing. The exclusion of phonology from early word processing is clearly not in harmony with the results obtained by the same group of authors using their fast priming paradigm (see Pollatsek et al. [2000] for a recent review of these issues).

In sum, it appears that there is a considerable gap between the verbal description of the model and its actual mathematical implementation – and no clear theoretical justification for the particular option adopted in either the verbal or the mathematical version of the model.

As mentioned by the authors in their detailed discussion of alternative models, a different approach to the integration of linguistic processing and eye-movement control has been taken by Reilly and Radach (2003). They used a letter- and word-processing module that implements a well established type of interactive activation model as developed in research on single word recognition (Grainger & Jacobs 1996; 1998). The results of word processing are continuously fed back to a spatial salience vector that serves as an arena to integrate visual and linguistic processing in the selection of words as saccade targets. The trigger for the execution of an eye movement comes from an independent fixate center that is codetermined by the level of activity in the word processing module (see Engbert et al. [2002] for a similar approach). In this architecture, there is no distinct processing event that triggers saccade programming, and no shifting of attention. Hence, there is also no need to divide word processing into stages, and some of the problems originating from this division can be avoided. However, although the Glenmore model appears to represent a promising theoretical alternative, it has still to be tested in simulations with a realistic corpus of reading data. It remains to be seen whether it can then match the impressive performance of E-Z Reader in accounting for a wide range of eye-movement phenomena in reading.

Future challenges to E-Z Reader: Effects of OVP and morphology on processing long and short compounds

Jukka Hyönä and Raymond Bertram

Department of Psychology, University of Turku, Turku, FIN-20014, Finland.
hyona@utu.fi rayber@utu.fi http://users.utu.fi/hyona/

Abstract: We argue that although E-Z Reader does a good job in simulating many basic facts related to readers' eye movements, two phenomena appear to pose a challenge to the model. The first has to do with word length mediating the way compound words are identified; the second concerns the effects of initial fixation position in a word on eye behavior.

As Reichle et al.'s target article convincingly demonstrates, the research on eye guidance has reached a stage where the accumulation of empirical data has paved the way for attempts to model and simulate the basic findings related to eye movements in reading. Without any doubt, E-Z Reader has been the most influential model in this respect.

For the information-processing models, it is desirable that they (1) are transparent; (2) are psychologically and neurally plausible; (3) account for the basic empirical facts; and (4) make novel, testable predictions. On the one hand, we think E-Z Reader fares well with respect to the first three requirements (with some limitations, mentioned below). On the other hand, although E-Z Reader may also be capable of producing novel empirical predictions, we are not told what these might potentially be (with one exception concerning a previous version of the model). We are left wondering whether the model is indeed restricted to predicting only the effects that it is designed to simulate, or whether the authors have not yet fully exploited its capacity to generate novel predictions.

In what follows, we take up two empirical phenomena that pose a challenge to E-Z Reader, namely, the role of morphology in the processing of compound words and the effects of fixation location in a word on eye behavior.

Many languages (e.g., German, Dutch, and Finnish) depart significantly from English in having highly frequent compounding. For example, in Finnish, more than 50% of all existing words are compounds. If we are to understand the basic reading processes in these languages, we need to acquire a good insight into how compounds are processed. In a recent study, we (Bertram & Hyönä 2003) demonstrated that word length mediates this process: For relatively long compounds (12–14 characters) the recognition process starts with lexical access of the first constituent and not of the whole word (see also Hyönä & Pollatsek 1998; Pollatsek et al. 2000), whereas the opposite is true for the relatively short compounds (7–8 characters). The potential challenge these findings pose to E-Z Reader is that word length appears to determine whether word or constituent frequency affects the initial fixation on a word.

An initial attempt to model compound word processing with E-Z Reader demonstrates that the job is not trivial (Pollatsek et al. 2003). Pollatsek et al. showed that a version that fitted the data best on reading long compounds was the one in which word identification was assumed to appear serially via the constituents (i.e., accessing first the initial constituent, then the second one, then gluing the two together). However, such a model runs into problems in accommodating our finding (Bertram & Hyönä 2003) that short compounds are recognized via the whole word form and that the first constituent does not seem to get activated. What might be needed are letter-level representations feeding activation to word-level nodes, where letter-level activation varies as a function of eccentricity from the center of fixation point (cf. the Glenmore model of Reilly & Radach [2003]). In the case of short compounds this is not enough, but there should also be a mechanism that gives priority to word-level nodes over compound word constituents. Without such a mechanism, there would be faster and more pronounced activation of the first constituents, because they

https://doi.org/10.1017/S0140525X03320102 Published online by Cambridge University Press