ing a given fixation, the point of maximum salience dynamically changes to be highest at the saccade word target before the saccade execution. In interactive activation models (McClelland & Rumelhart 1981), the processing systems (as lexical access) are controlled by the connections among different interconnected units (features, letters, and words) and are not capacity limited.

Please stop using word frequency data that are likely to be word length effects in disguise

Marc Brysbaert[a] and Denis Drieghe[b]

[a]*Department of Psychology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom;* [b]*Department of Experimental Psychology, Ghent University, B-9000 Ghent, Belgium.*
marc.brysbaert@rhul.ac.uk **denis.drieghe@rug.ac.be**
http://psyserver.pc.rhbnc.ac.uk/staff/mbrysbaert.html
http://allserv.rug.ac.be/~ddrieghe/

Abstract: Reichle et al. claim to successfully simulate a frequency effect of 60% on skipping rate in human data, whereas the original article reports an effect of only 4%. We suspect that the deviation is attributable to the length of the words in the different conditions, which implies that E-Z Reader is wrong in its conception of eye guidance between words.

A computational model is as good as the data it simulates. This is why Reichle et al. rightly pride themselves about the good fit of the model's outcome with human data. The human data predominantly come from a reading study, reported by Schilling et al. (1998), in which 30 college students read 48 sentences. According to Figure 6 in the target article, the observed frequency effects in this study were roughly 70 msec for gaze duration, 30 msec for first fixation duration, and a 60% for word skipping rate. What Reichle et al. did not mention is that the Schilling et al. study was originally designed to look at the word frequency effect under very controlled circumstances (i.e., with words that were matched on all other variables except for word frequency, and with sentence context that constrained the target words equally). Each participant saw a number of sentences with low frequency words (2 per million) and a number of sentences with high frequency words (141 per million). These frequencies probably coincide with the frequency classes 1 and 5 of Figure 6 in the target article. If we look at the data reported by Schilling et al. for these particular stimuli, we obtain a frequency effect of 67 msec for gaze duration and 35 msec for first fixation duration, but only 4% for skipping rate ("Subjects fixated on HF words 89% of the time and on LF words 93% of the time" – Schilling et al., p. 1,272). That is, for this particular subset of well-controlled stimulus words, in Schilling et al., the effects for gaze duration and first fixation duration agree well with the overall data used by Reichle et al., but this is not true for the skipping rate. How come E-Z Reader "correctly" simulates a 60% difference in skipping rate between low-frequency and high-frequency words, whereas in the human data there was only a 4% difference attributable to word frequency?

After a review of all previously published word skipping data, Brysbaert and Vitu (1998) concluded that the frequency effect on word skipping is 4% on average (i.e., exactly the effect reported by Schilling et al., as well), and that the effect was 9% for contextual predictability (i.e., very predictable words in a sentence are skipped, on average, 9% more often than unpredictable words). In addition, they observed a 60% difference attributable to word length: 2-letter words are skipped more than 60% of the time, whereas 10-letter words are virtually never skipped in first-pass reading. To us, these data strongly suggest that what Reichle et al. simulate in the lower part of Figure 6 is not so much a frequency effect on skipping rate but a word-length effect on skipping rate. The authors themselves are clearly aware of this problem, because in Rayner et al. (1998c, p. 256, footnote 3), they wrote:

> In our modelling, to minimize the number of parameters, we did not distinguish between frequency and word length effects. Thus "frequency effects" in our model are really a combination of frequency and word length effects because the two are highly correlated in our sample of text as in printed English in general.

For this reason, we were very surprised to see that in the present article they still refuse to report the data separately for word length and word frequency, even though the current model is supposed to have a mechanism to deal with the effects of the length of the parafoveal word (see Equation 1 of the target article). What we ask is that Reichle et al. give us a figure in which the word-skipping rates of the Schilling et al. corpus are shown as a function of word length and word frequency, together with the predictions of E-Z Reader. If these provide a good fit, we will rest our case. However, we strongly suspect that the model will largely overestimate the effect of frequency and underestimate the effect of word length. For this reason, until proven wrong, we still believe that E-Z Reader is fundamentally flawed in its conception of interword behaviour in general and word skipping in particular.

ACKNOWLEDGMENT
Denis Drieghe is a research assistant of the Fund for Scientific Research (Flanders, Belgium).

Reading the scene: Application of E-Z Reader to object and scene perception

Peter De Graef and Filip Germeys

Laboratory of Experimental Psychology, University of Leuven, Leuven, B-3000, Belgium. **peter.degraef@psy.kuleuven.ac.be**
filip.germeys@psy.kuleuven.ac.be
www.psy.kuleuven.ac.be/labexppsy/top/peterweb
www.psy.kuleuven.ac.be/labexppsy/top/filipweb

Abstract: We discuss five basic principles of E-Z Reader in terms of their potential for models of eye-movement control in object and scene perception. We identify several obstacles which may hinder the extrapolation of the E-Z Reader principles to nonreading tasks, yet find that sufficient similarities remain to justify using E-Z Reader as a guide for modeling eye-movement control in object and scene perception.

Eye-tracking has provided vision science with a powerful tool to unobtrusively monitor on-line perceptual and cognitive processing. Unfortunately, eye movement records generate a host of overt measures which all may (or may not) reflect some aspect of covert processing, leading to much debate about which measure would be most appropriate (e.g., Inhoff & Radach 1998). The most promising solution to this debate is to consider multiple eye movement measures simultaneously (Henderson et al. 1999). However, to do this, an integrated model is required that specifies the relations between the various overt measures as well as their correspondence to covert processes. This is precisely what Reichle et al. have achieved with E-Z Reader.

As users of eye-tracking methodology in object and scene perception, we can be only envious of this situation, yet at the same time Reichle et al. inspire some optimism with their suggestion that the basic principles of E-Z Reader may apply to other visual information processing tasks (sect. 4.9). We would like to evaluate the grounds for such optimism by examining five basic principles of E-Z Reader to determine whether and how they can be applied to the study of eye-movement control in object and scene perception.

First, according to E-Z Reader, the main engine of eye movements in reading is serial word identification. This makes sense given (a) the importance of individual word order and meaning to understand the whole sentence, and (b) the ease with which individual words can be segregated from a sentence. In scene perception, neither of these conditions is fulfilled. It is quite possible to achieve scene interpretation without identifying any of the objects in the scene, let alone identifying them in a certain order (Oliva & Torralba 2001). Moreover, object-background segregation in scenes can be so computationally demanding that object identification precedes figure-ground organization (Vecera & Farah 1997). Thus, object fixation times could reflect processing of object identity as well as background and figure-background relations. Consequently, in scene perception we may have to adopt a more general principle than object identification as the engine behind eye movements. One possibility is that the visual system simultaneously monitors the rate of activation buildup in an object lexicon and in an object localization module: As soon as both rates drop below a criterion level, the current fixation position is deemed to be suboptimal and an eye movement is planned.

Second, E-Z Reader assigns an important role to pre-attentive processing of the upcoming word. Given the rigid serial structure in which information needs to be acquired in reading, this implies that pre-attentive processing is restricted to the word that is about to become the saccade target. In scene exploration, however, there is no inherent spatial order in which objects need to be processed. Therefore, pre-attentive processing occurs for saccade bystanders as well as for the saccade target (Germeys et al. 2002). It will therefore be necessary to determine the spatial and temporal windows within which pre-attentive processing of a saccade bystander can influence that object's fixation duration or skipping probability once it has finally become the saccade target.

Third, E-Z Reader posits that attention shifts and eye movements are decoupled. We wonder whether the sparse and serial stimulus structure inherent in sentences may not be a necessary prerequisite for such a decoupling. In other words, autonomous attention shifts and eye movements may be possible only because the next relevant stimulus component is always easily discriminated on the basis of rudimentary boundary information. In scene perception, however, the next relevant stimulus component could be anywhere in the visual field; a more sophisticated process is required to mark the location of the next saccade target. As demonstrated by behavioral data (Deubel & Schneider 1996) as well as single-cell recordings in LIP (Colby et al. 1996), spatially selective shifts of visual attention appear to be that process, indicating a strong coupling between attention and eye movements. This implies that in scene perception we must invoke different mechanisms when our eye-movement records indicate refixations, spillover effects, and foveal-on-parafoveal effects.

Fourth, E-Z Reader elegantly limits the number of factors influencing fixation duration to visual acuity, word frequency, and word predictability. In object and scene perception these factors are also likely to play a role, although some may not be easy to estimate (e.g., what would constitute a good estimate of object frequency?). The relative importance of these factors is likely to be different in scene perception than in reading. Specifically, because scene identity is available early on in scene exploration (Biederman 1981) more subsequent fixations may show predictability effects than in reading where context develops more gradually. In addition, the list of factors influencing fixation times probably also needs to be extended. For example, ease of object identification has been argued to be a function of object orientation (Boutsen et al. 1998), object size (Theios & Amrhein 1989), and object camouflage (De Graef et al. 1990), all of which may have effects on eye-movement measures.

Fifth and finally, E-Z Reader capitalizes on the incorporation of very task-specific constraints in the model, such as the preferred saccade length in reading English. One could argue that this limits the generality of the model, but we feel such parameters are justified when they accurately reflect eye-movement behavior in the task under study. Moreover, while the parameter value is obviously task-dependent, the parameter itself may not be. Specifically, that preferred saccade length in reading English is estimated to be seven characters may be linked to the fact that the perceptual span for word encoding in reading English extends about eight characters to the right of fixation (Rayner et al. 1982). In other words, readers prefer to saccade to the edge of their perceptual span, a principle which may also apply to much less constrained tasks such as scene exploration (Shioiri & Ikeda 1989). Other task-specific constraints derived from reading data may be less suitable to extrapolate. For example, E-Z Reader assumes that all fixation times are sampled from a unimodal distribution. However, in other tasks, fixation time distributions may be multimodal, raising the question of whether fixation times in the various component distributions can all be modeled in the same fashion (De Graef 1998).

In summary, it would be unwise to extrapolate E-Z Reader to object and scene perception without careful consideration of task-specific differences in the interplay between visual processing, processing goals, attention, and oculomotor control. However, E-Z Reader does provide a valuable framework for thinking about the best design principles for a model of eye movements in object and scene perception.

ACKNOWLEDGMENTS

This work was supported by Concerted Research Effort Convention GOA 98/01 of the Research Fund K.U. Leuven and the Research Training Network "Perception for Recognition and Action" (RTN-2001-00107), under the direction of the European Commission.

Are there two populations of refixations in the reading of long words?

Karine Doré-Mazars, Dorine Vergilino-Perez, and Thérèse Collins

Laboratoire de Psychologie Expérimentale, UMR 8581 CNRS Université René Descartes Paris 5, France. **dore@psycho.univ-paris5.fr**
vergilino@psycho.univ-paris5.fr **http://d.vergilinoperez.free.fr**

Abstract: This commentary focuses on the limitations of the E-Z Reader model in its attempt to explain refixation saccades in reading. Listing factors that influence probability of refixating leads the model to assume two sorts of refixations. However, taking into account data on the metrics of refixation saccades allows us to propose an alternative explanation for empirical observations reported in the literature.

In reading, the probability of refixating a word – that is, to read it with two fixations – is known to increase with word length, first fixations landing far from the word center and the decrease of word familiarity (McConkie et al. 1989). The E-Z Reader model accounts for these empirical data by making the following assumptions. The rapid parafoveal integration of the length of the to-be-fixated word results in the preparation of a refixation program. Although it remains implicit, this assumption suggests that the *decision* to refixate long words is taken before landing on the word. The computation of the refixation saccade is then initiated once the eyes land in the word. As the ability of the saccadic system to modify or cancel previous motor plans is now well documented in the literature (since the famous Becker & Jürgens [1979] study; see also Vergilino-Perez & Beauvillain, in press), the target model proposed that the refixation saccade program can be canceled during the first fixation on the word. The cancellation of a refixation saccade program would be more likely in a high frequency word than in a low frequency word because the progression of the first stage of the lexical processing is faster on the former than on the latter. Such an assumption is an elegant explanation for the word frequency effect on refixation probability (Inhoff & Rayner 1986; McConkie et al. 1989).

However, such a scenario does not fit with the classical interpretation of the effect of the first landing position on refixation probability. When the first fixation position was imposed at different locations in an isolated word (e.g., O'Regan et al. 1984), refixation probability increased when locations were far from the word center, a location usually called the optimal viewing position