Measuring word recognition in reading: eye movements and event-related potentials

Sara C. Sereno[1] and Keith Rayner[2]

[1]Psychology Department, University of Glasgow, Glasgow, Scotland, G12 8QB, UK
[2]Psychology Department, University of Massachusetts, Amherst, MA 01003, USA

The investigation of visual word recognition has been a major accomplishment of cognitive science. Two on-line methodologies, eye movements and event-related potentials, stand out in the search for the holy grail – an absolute time measure of when, how and why we recognize visual words while reading. Although each technique has its own experimental limitations, we suggest, by means of review and comparison, that these two methodologies can be used in complementary ways to produce a better picture of the mental action we call reading.

Reading – the visual comprehension of language – is an artificial ability not yet subject to the pressures of evolution. The complex structure and function of our visual system dates from our shared primate origins millions of years ago. The origin of spoken language is loosely estimated to begin ~100 000 years ago. Writing systems, on the other hand, emerged in cities in Mesopotamia and Egypt at the earliest around 3500 BC. Mass literacy in the developed world was not fully established until the second half of the nineteenth century. Even today, a large percentage of the world's population has only rudimentary levels of literacy. Yet the development of modern science and culture is predicated on written language. What is it about reading that makes it such a powerful tool? Central to the study of reading is understanding visual word recognition and answering such questions as: How does the visual form of a word activate its meaning? When exactly does this happen? What factors influence this process?

Words on a page enter the brain via the retina. The past three decades of neuroanatomical and neurophysiological work on monkeys has revealed much about the 30 or so visual areas of the brain, including their response characteristics and hierarchical organization into functionally distinct processing streams [1]. But monkeys do not read. In humans, visual word recognition has been investigated using various approaches – neuropsychological, behavioral, computational, and, most recently, neuroimaging (Box 1). To investigate on-line visual language processing, however, a methodology must heed the rapid rate of reading and also precisely capture the

Corresponding author: Sara C. Sereno (ssereno@psy.gla.ac.uk).

Box 1. Measures of visual word recognition

Our knowledge of how humans process written language has emerged from neuropsychological work on patient populations. The identification of focal brain lesions that resulted in language-processing deficits led to hypotheses about the normal functioning of such areas [35]. But the location and extent of lesions are often untidy, and generalizations across patients problematic. Standard behavioral testing over the past 35 years has produced a massive literature on word recognition. Typically in these studies, a variety of reaction-time tasks are used (e.g. lexical decision, semantic categorization, word naming, self-paced reading) under a variety of stimulus presentation conditions (e.g. words presented alone; words preceded or followed by a prime or mask; word-by-word sentence presentation). Results from this work have been used to generate and refine box-and-arrow models of word recognition, detailing stages of processing and their attendant properties. Running in a parallel stream, eye movement research over the past 25 years has also investigated word recognition, but in the context of normal reading using eye fixation duration as the response measure.

A gap in our knowledge separates psycholinguistic behavior from a level of neural implementation. One computational approach used to address this problem simulates patterns of response across a network of units ('neurons'), for example in parallel distributed models of word recognition [36,37]. More recently, neuroimaging methodologies have begun to chart the psycholinguistic territories of the brain [38,39]. The goal here is to determine the relationship between the psychology and function of the mind and the physiology and structure of the brain. Unlike single-cell recording, the resulting descriptions encompass an immense scale. Imaging techniques based on tracking blood flow to active sites in the brain, such as PET and fMRI, render general pictures of localized differences of activity. Like previous tools of cognitive science, they depend upon inference [40]. As the hemodynamic response is a metabolic consequence of electrical activity, it is invariably delayed. Developments based on the use of 'smart' contrast agents that track calcium flow or glucose uptake, for example, promise better temporal resolution. Other imaging techniques are based on measuring the electrical activity generated by the brain via densely mapped sensor arrays positioned on the scalp. These include magneto- and electroencephalogram (MEG and EEG) recording. Although both provide a millisecond-by-millisecond record of activity, MEG is superior in terms of localization of cortical sources. Localization of the neural generators in EEG is problematic; solving the 'inverse problem' (finding sources given surface recordings) requires multiple constraints [41]. Although the different imaging methodologies might hint at convergence [42], it is important to recognize the limits of each. Alignment of results using different techniques is most informative when the spatiotemporal scales are comparable.

http://tics.trends.com 1364-6613/$ - see front matter doi:10.1016/j.tics.2003.09.010

dynamics of concurrent mental processing. We believe that eye movement (EM) and event-related potential (ERP) recording represent the best of what is currently available to identify real-time markers of word recognition. Although the EM and ERP techniques have developed largely independent of each other, we suggest that their alliance can help reveal a more precise time course of word recognition in reading.

Eye movements (EMs)

An important contribution to understanding on-line visual language processing has come from EM research where reading behavior is measured via the position, duration, and sequence of eye fixations in text. The EM technique offers certain advantages over traditional behavioral techniques. EMs are a normal part of reading. First, subjects do not have to make decisions about words they read, or name them aloud – procedures that disrupt the flow of reading when used as a secondary task. Second, the processing of a word in text is, in fact, reflected in its fixation time [2–5]. For example, words that are short in length, regular in their spelling–sound pattern, frequent in their occurrence, or semantically or syntactically predictable from a previous context are fixated for less time than those that are not. Finally, because fixation time is quite brief (on average, ~200–250 ms), it sets temporal constraints for processing (Box 2).

Whereas most words are typically fixated once, sometimes words are immediately refixated or entirely skipped. Before a word is foveated, it has generally been viewed

Box 2. Constraints that eye movements place on word recognition and ERPs

A typical reading rate for college students of 300 words per minute averages out to 200 ms per word. Reading, however, involves more than identifying individual words. Because eye movements during skilled reading reflect moment-to-moment cognitive processes, the average fixation duration of ~250 ms constrains the amount of time for lexical processing. The time course of events within a single fixation in reading can be delimited (Figure I). At the beginning of a fixation, it takes ~60 ms for information about the fixated word to travel to higher cortical areas where lexical processing begins. Before the end of a fixation, oculomotor latency (the time needed to program an eye movement) limits the interval during which a sufficient degree of lexical processing must be achieved [43]. Fixation duration will vary as a function of lexical difficulty, but processing must be well under way within the first 100–200 ms of the fixation to meet a deadline for programming the next eye movement. However, fixation time on a particular word does not pinpoint when during the fixation the word is actually recognized. ERPs can elucidate the stages of processing within a fixation but, to do so, such effects must occur within the temporal confines of a typical eye fixation (Figure I).

What drives the eyes forward in text has not been fully resolved. EM research has sought to specify: (1) the type of information acquired (e.g. featural, orthographic, phonological, semantic); (2) when this information is acquired (e.g. parafoveally, foveally, before or after initiation of saccadic programming); and (3) how decisions of when and where to move the eyes are made. The accumulation of evidence favors higher-level, cognitive oculomotor control in executing EMs [44–46]. It seems that a certain level of lexical access is necessary for initiating EMs. Moreover, an attentional mechanism must be involved in normal reading to explain such effects as the parafoveal preview benefit, the size and asymmetry of the perceptual span (effective field of vision in reading), and word skipping. Finding electrophysiological correlates of these processes could provide independent verification.

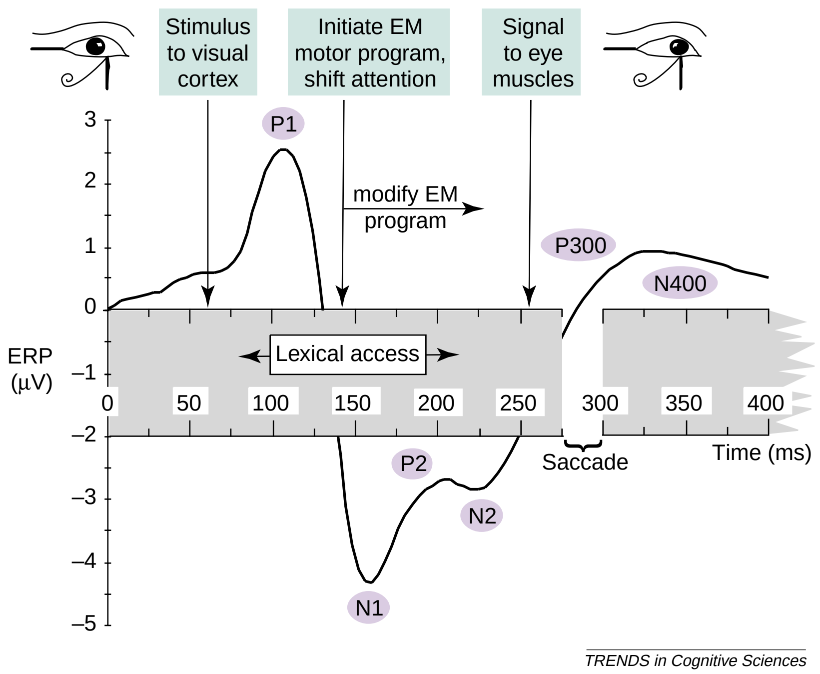

Figure I. Time line of processing a word in reading. A slightly inflated average fixation time (275 ms) is used to compensate for lack of parafoveal preview. After the saccade from the word, a new fixation begins. On the same time line, an ERP waveform is shown, representing the averaged response to a foveal visual word (average of 40 subjects presented 288 words each). Such waveforms are recorded from an electrode over the left occipito-temporo-parietal region of the scalp [25]. ERP-component nomenclature is a combination of the following: (1) the polarity of the component, positive- or negative-going (P or N, respectively) and (2) either the ordinal number within a given polarity (e.g. 1, 2) or the approximate latency (in ms) of its peak amplitude (e.g. 300, 400). So, for example, the N1 here could also be termed the N150.

parafoveally on the previous fixation. This parafoveal preview assists the later (foveal) processing of the word, conferring a benefit of around 20 ms or more, an effect demonstrated when such information is denied [6–8]. Because of the imperfect coupling between a given fixation and the direct processing of a word, EM research has sought out additional fixation duration and eye position measures to characterize the data, and there is evidence that the many components of word recognition can be indexed by such measures [3]. We can infer, for example, from word frequency effects in fixation duration, that lexical access is well under way before the end of the fixation. Because the terminal portion of a fixation must be reserved for oculomotor latency (the lag between programming and making an EM), any processing that affects fixation time must occur roughly within the first 150 ms or so of fixation (Box 2). What remains uncertain, however, is pinpointing in real time precisely when such effects occur.

Event-related potentials (ERPs)

ERPs are stimulus-locked averages of the electroencephalogram (EEG) across many presentations of stimuli. They provide a continuous millisecond-by-millisecond record of electrical changes related to on-going perceptual and cognitive processing and can thus index changes related to word recognition in real time [9,10]. ERPs recorded at the scalp via high-density (e.g. 128-channel) sensor arrays provide a measure of processing with fine-grained (external) spatial as well as temporal resolution [11].

Traditional ERP studies of written language have focused on later, endogenous components of the waveform, such as the N400 (a negative-going wave occurring ~400 ms after stimulus onset), a component thought to be sensitive to semantic relationships [12,13]. A challenge to this research is the apparent discrepancy between what a normal reading rate dictates in terms of the speed of lexical and semantic events and what the ERP record has mainly shown. That is, many significant ERP differences seem to occur too late. In reading, for example, by 400 ms the eyes have already moved onto the next word (see Box 2). The onset of the N400 is not fixed at 400 ms, as its name might imply, but can begin as early as 200 ms and last as long as 600 ms post-stimulus. Nevertheless, in the literature, the amplitude and latency of the N400 are typically examined between 300–500 ms post-stimulus and this time window is often considered to encompass the process of interest.

Human electrophysiological studies assume that ERP components from 50–150 ms post-stimulus represent measures of signaling through a hierarchical visual organization. Research combining ERP and fMRI has confirmed this sequential progression [14]. There is a developing argument that the speed of processing and information flow through the visual system is more rapid than has traditionally been assumed [15]. The first afferent volley reaches frontal cortex 80 ms post-stimulus and continues through the top-down feedback loops that modulate further processing in sensory areas [16].

Findings based on monkey intracranial recordings indicate a brief time frame for signal transmission [17]. In one model [18], stimulus activation of the visual system produces a rapid fast-forward sweep followed by a slower set of recurrent interactions operating both within an activated area and backwards to lower levels of the system. By implication, the traditional ERP components might be indicative of recurrent feedback-driven processes rather than the first information sweep through the system [19]. Importantly, the neurophysiological evidence for quick, distributed activation of the visual system provides a plausible framework for the rapid completion of perceptual and cognitive tasks.

Comparisons of EMs and ERPs

Eye fixation time in reading sets constraints for processing but without sufficient temporal resolution. ERPs can provide greater resolution because they are time-locked on a millisecond-by-millisecond basis to brain events during the presentation of a word. For ERP research to advance knowledge about the timing of visual word recognition, however, it must show sensitivity within earlier components. Early, exogenous ERP components have not readily exhibited processing differences arising from lexical variables. However, a few ERP studies have demonstrated early effects – those occurring before 200 ms post-stimulus – commensurate with EM data [20–25]. For example, effects of word frequency and contextual constraint have been demonstrated in the N1 (first negative component, which peaks around 150 ms post-stimulus) beginning as early as 132 ms post-stimulus [24,25].

To date, a handful of studies have used EM and ERP techniques in a complementary way. One approach has recorded the EEG during reading but used the electrooculogram (EOG) to measure EMs [26,27]. Eye position in the text is estimated and the EEG data are averaged from saccade offset (not from stimulus onset as with ERPs), producing a saccade-related potential. Another approach has directly compared EM and ERP data by using similar materials in separate studies [8,24,25,28,29]. Effects of word frequency, contextual predictability, lexical ambiguity, and rereading have been examined in this way.

Tracking the word frequency effect across behavioral and eletrophysiological paradigms is particularly relevant because its presence is considered a marker for successful lexical access (Box 3). Word frequency effects have also been used in examining how the meaning of an ambiguous word is resolved within a context [24,29]. A high-frequency ambiguous word such as 'bank' has a highly dominant sense (money) and one or more weaker subordinate senses (river). In a context that biases the subordinate interpretation, the ambiguous target is at once a high-frequency word form and a low-frequency word meaning. Although EM studies have been somewhat equivocal regarding the relative timing of the activation of alternative meanings, the electrophysiological record seems to favor early lexical selection of the subordinate, low-frequency sense. If word frequency effects do indeed index lexical access, the N1 (the earliest electrophysiological marker of frequency) could become a functional watershed separating early, lexical access from later, post-lexical integration stages of processing. Knowing the precise locus of an effect has important implications for models of word recognition and

Box 3. Measuring lexical access via word frequency

The word frequency effect represents the difference in responses to high-frequency (HF) words that are most commonly used and low-frequency (LF) words that occur much less often. The presence of word frequency effects indicates that lexical access has occurred [47]. Using similar stimuli across different experimental measures – lexical decision, EMs and ERPs – allows for a comparison of elicited responses on a single temporal scale (Figure I).

In a representative lexical decision (word/non-word) task, reaction time was significantly faster to HF words (490 ms) than LF words (553 ms) [25]. However, word recognition does not take as long as this. The mental chronometry of a lexical decision includes, in addition to word identification, various decision and motor preparation stages [48].

A subset of lexical decision words was placed in the context of sentences in an EM reading task in which the parafoveal preview of HF and LF targets was varied [8]. Parafoveal preview was either valid (the normal reading situation) or invalid (a letter string other than the upcoming target word was displayed parafoveally; as in a single-word presentation paradigm, the target is only viewed foveally). First fixation duration (FFD) on the target was measured. FFD is the average of all first fixations made on a word disregarding any refixations (typically, words are fixated once). When the parafoveal preview was valid, FFDs were significantly shorter on HF words (259 ms) than LF words (275 ms). Similarly, when the parafoveal preview was invalid, FFDs were significantly shorter on HF words (280 ms) than LF words (293 ms). The difference in latencies between valid and invalid conditions provides an estimate of the preview benifit (~20 ms).

ERPs were recorded for the identical set of HF and LF words used in the aforementioned lexical decision task [25]. The amplitude of the N1 component of the waveform from 132–164 ms post-stimulus was significantly less for HF than LF words. In another experiment, ERPs were recorded for a different set of HF and LF words that were sentence-final targets in a word-by-word sentence presentation paradigm [24]. Significant amplitude differences (from 132 to 192 ms post-stimulus) were again demonstrated in the N1.

In summary, behavioral measures such as lexical decision are useful in identifying the stages of word recognition. EM measures, because of their shorter duration, produce a working estimate of the time window within which lexical access must occur, namely around 100–200 ms post-stimulus. Finally, ERPs are capable of specifying precisely when such effects emerge in real time.

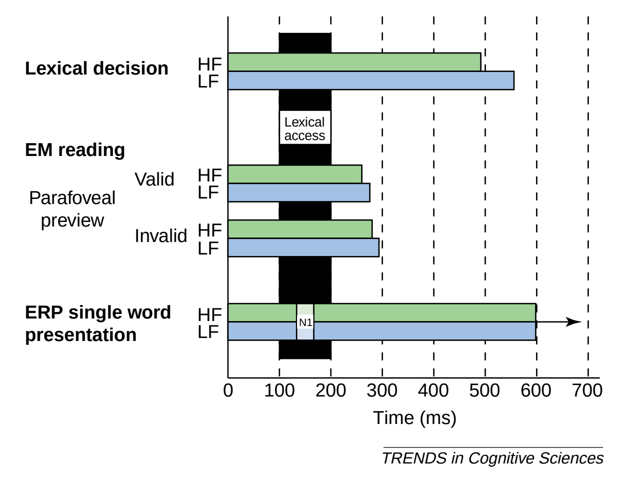

Figure I. Temporal comparison of word frequency effects across experimental paradigms. Responses to high-frequency (HF) and low-frequency (LF) words using similar stimuli across three different methodologies are represented on a time scale. Significant word frequency effects are present in the lexical decision reaction time, the first fixation duration in EM reading tasks, and in the ERP N1 component. The vertical black bar represents the time period (100–200 ms) during which lexical access is estimated to occur.

for determining whether a process is driven by bottom-up or top-down mechanisms.

The way forward

EM research, guided by the temporal contours of the reading task, has introduced innovations that use eye contingent display changes – such as the 'moving window' [30], 'boundary' [7], 'fast priming' [31], and 'disappearing text' [32] techniques – to explore the intricacies of reading. In ERP research, the increased use of dense-mapping arrays, the application of statistical techniques such as principal component analysis [20,24,33] and independent component analysis [34], corroboration from MEG studies [23], and evidence for rapid visual and higher-order processing have stimulated a welcome interest in the early components that are crucial for establishing exact time measures.

The current challenges impinge on both EM and ERP research. It is important to show that factors affecting fixation time can be reliably demonstrated in early ERP components. It then becomes possible to use ERP data to determine, for example, whether the temporal locus of an EM effect is lexical or post-lexical. The prospect of simultaneous EM and ERP recording also opens the door to possibilities such as finding the electrophysiological marker of saccade generation during normal reading. We know that word recognition occurs very rapidly – normal rates of reading with comprehension demonstrate this. We are confident that EM and ERP methodologies can be combined more effectively than they have been to date.

Acknowledgements

We thank Patrick O'Donnell, Jeffrey Bowers, and the anonymous reviewers for their helpful comments. Preparation of this article was supported by a British Academy grant 30315 to SCS and grants HD17246 and HD26765 to KR.

References

1 Felleman, D.J. and Van Essen, D.C. (1991) Distributed hierarchical processing in the primate visual cortex. *Cereb. Cortex* 1, 1–47

2 Liversedge, S.P. and Findlay, J.M. (2000) Saccadic eye movements and cognition. *Trends Cogn. Sci.* 4, 6–13

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

4 Rayner, K. and Sereno, S.C. (1994) Eye movements in reading: psycholinguistic studies. In *Handbook of Psycholinguistic Research* (Gernsbacher, M.A., ed.), pp. 57–81, Academic Press

5 Starr, M.S. and Rayner, K. (2001) Eye movements during reading: some current controversies. *Trends Cogn. Sci.* 5, 156–163

6 Blanchard, H.E. *et al.* (1989) The acquisition of parafoveal word information in reading. *Percept. Psychophys.* 46, 85–94

7 Rayner, K. (1975) The perceptual span and peripheral cues in reading. *Cogn. Psychol.* 7, 65–81

8 Sereno, S.C. and Rayner, K. (2000) Spelling-sound regularity effects on eye fixations in reading. *Percept. Psychophys.* 62, 402–409

9 Picton, T.W. *et al.* (2000) Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37, 127–152

10 Rugg, M.C., Coles, M.G.H. eds, (1995) *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition* Oxford University Press

11 Srinivasan, R. *et al.* (1998) Estimating the spatial Nyquist of the human EEG. *Behav. Res. Methods Instrum. Comput.* 30, 8–19

12 Kutas, M. and Federmeier, K.D. (2000) Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470

13 Kutas, M. and Van Petten, C.K. (1994) Psycholinguistics electrified. In *Handbook of Psycholinguistic Research* (Gernsbacher, M.A., ed.), pp. 83–143, Academic Press
14 Martínez, A. *et al.* (1999) Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat. Neurosci.* 2, 364–369
15 Thorpe, S. *et al.* (1996) Speed of processing in the human visual system. *Nature* 381, 520–522
16 Foxe, J.J. and Simpson, G.V. (2002) Flow of activation from V1 to frontal cortex in humans: a framework for defining 'early' visual processing. *Exp. Brain Res.* 142, 139–150
17 Schroeder, C.E. *et al.* (1998) A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cereb. Cortex* 8, 575–592
18 Lamme, V.A.F. (2003) Why visual attention and awareness are different. *Trends Cogn. Sci.* 7, 12–18
19 Buchner, H. *et al.* (1997) Fast visual evoked potential input into human area V5. *Neuroreport* 8, 2419–2422
20 Dien, J. *et al.* (2003) Parametric analysis of event-related potentials in semantic comprehension: evidence for parallel brain mechanisms. *Brain Res. Cogn. Brain Res.* 15, 137–153
21 Neville, H.J. *et al.* (1992) Fractionating language: different neural subsystems with different sensitive periods. *Cereb. Cortex* 2, 244–258
22 Nobre, A.C. and McCarthy, G. (1994) Language-related ERPs: scalp distributions and modulation by word type and semantic priming. *J. Cogn. Neurosci.* 6, 233–255
23 Pulvermüller, F. *et al.* (2001) Neuromagnetic evidence for early semantic access in word recognition. *Eur. J. Neurosci.* 13, 201–205
24 Sereno, S.C. *et al.* (2003) Context effects in word recognition: evidence for early interactive processing. *Psychol. Sci.* 14, 328–333
25 Sereno, S.C. *et al.* (1998) A time-line of word recognition: evidence from eye movements and event-related potentials. *Neuroreport* 9, 2195–2200
26 Joyce, C.A. *et al.* (2002) Tracking eye fixations with electroocular and electroencephalographic recordings. *Psychophysiology* 39, 607–618
27 Marton, M. and Szirtes, J. (1988) Context effects on saccade-related brain potentials to words during reading. *Neuropsychologia* 26, 453–463
28 Raney, G.E. and Rayner, K. (1993) Event-related brain potentials, eye movements, and reading. *Psychol. Sci.* 4, 283–286
29 Sereno, S.C. *et al.* (1992) The effect of meaning frequency on processing lexically ambiguous words: evidence from eye fixations. *Psychol. Sci.* 3, 296–300
30 McConkie, G.W. and Rayner, K. (1975) The span of the effective stimulus during an eye fixation in reading. *Percept. Psychophys.* 17, 578–586
31 Sereno, S.C. and Rayner, K. (1992) Fast priming during eye fixations in reading. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 173–184
32 Rayner, K. *et al.* (2003) Reading disappearing text: cognitive control of eye movements. *Psychol. Sci.* 14, 385–388
33 Spencer, K.M. *et al.* (1999) A componential analysis of the ERP elicited by novel events using a dense electrode array. *Psychophysiology* 36, 409–414
34 Jung, T-P. *et al.* (2001) Analysis and visualization of single-trial event-related potentials. *Hum. Brain Mapp.* 14, 166–185
35 Martin, R.C. (2003) Language processing: functional organization and neuroanatomical basis. *Annu. Rev. Psychol.* 54, 55–89
36 Kello, C.T. and Plaut, D.C. (2003) Strategic control over rate of processing in word reading: a computational investigation. *J. Mem. Lang.* 48, 207–232
37 Seidenberg, M.S. and McClelland, J.L. (1989) A distributed, developmental model of visual word recognition and naming. *Psychol. Rev.* 96, 523–568
38 Gernsbacher, M.A. and Kaschak, M.P. (2003) Neuroimaging studies of language production and comprehension. *Annu. Rev. Psychol.* 54, 91–114
39 Price, C.J. and Devlin, J.T. (2003) The myth of the visual word form area. *Neuroimage* 19, 473–481
40 Donders, F.C. (1969) On the speed of mental processes. *Acta Psychologia* 30, 412–431
41 Dale, A.M. and Sereno, M.I. (1993) Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *J. Cogn. Neurosci.* 5, 162–176
42 Posner, M.I. *et al.* (1999) Neuroanatomy, circuitry and plasticity of word reading. *Neuroreport* 10, R12–R23
43 Rayner, K. *et al.* (1983) Latency of sequential eye movements: implications for reading. *J. Exp. Psychol. Hum. Percept. Perform.* 9, 912–922
44 Rayner, K. *et al.* (1996) Eye movement control in reading: a comparison of two types of models. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 1188–1200
45 Reichle, E.D. *et al.* The E-Z Reader model of eye movement control in reading: comparisons to other models. *Behav. Brain Sci.* (in press)
46 Sereno, S.C. (1992) Early lexical effects when fixating a word in reading. In *Eye Movements and Visual Cognition: Scene Perception and Reading* (Rayner, K., ed.), pp. 304–316, Springer-Verlag
47 Balota, D.A. (1990) The role of meaning in word recognition. In *Comprehension Processes in Reading* (Balota, D.A. *et al.*, eds), pp. 9–32, Erlbaum
48 Balota, D.A. and Chumbley, J.I. (1984) Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 340–357