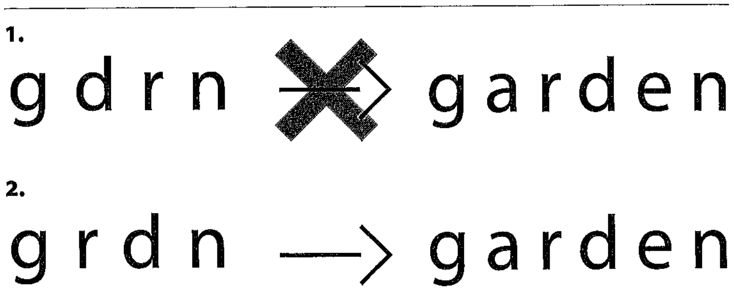
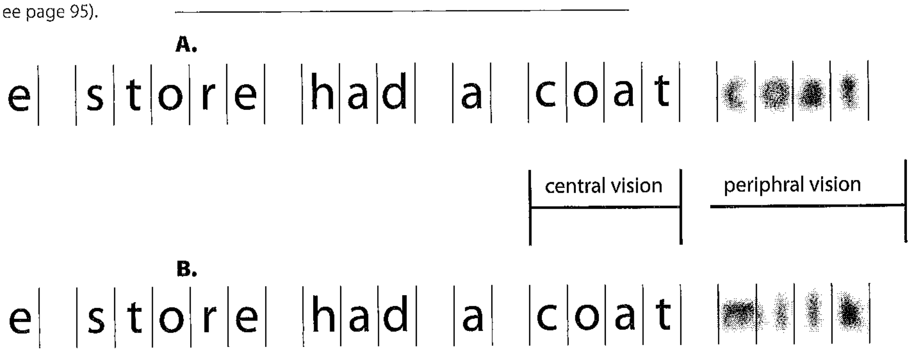
ɔosition priming:
ntification is
l by prior brief
:ion of a subset
get word's letters
ıtain their correct
osition,
'es word
on;
ɔt,
orts orthographic
ıg in reading
ɜ 90).



.phic processing
:o also be at work in
ıentences. Subjects
t" faster in A. than B.
ee page 95).



Orthographic Processing and Reading

Jonathan Grainger

Abstract

I will argue that processing letter identities and letter positions occupies a central interface between visual and linguistic processing during reading. This is primarily due to the fact that reading words in languages that use an alphabetic script is essentially letter-based. Information about letter identities and letter positions provides the gateway to whole-word written representations, to morphemes such as prefixes and suffixes, and to sound based representations. I will first summarize work on letter identification processes before describing mechanisms for parallel letter processing during single word reading. Finally, I will describe recent work demonstrating parallel processing of written information spanning several words during sentence reading.

Keywords

linguistic processing, letter identification, letter processing, letterform recognition, typography, orthographic, reading

Reading words:
m inkmarks to ideas[1]

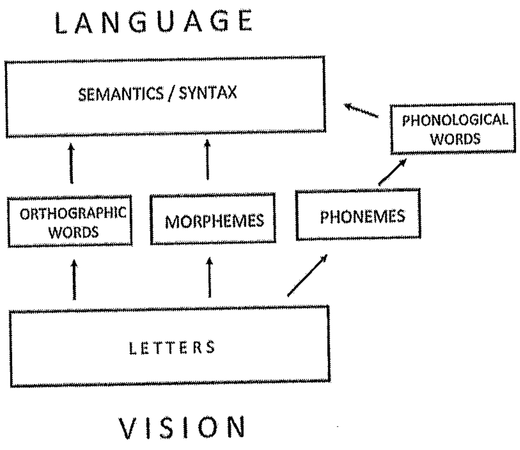
Words are the building blocks of reading in written languages that use word spaces, and in those languages that use an alphabetic script, letters are the building blocks of words. When reading, the eyes fixate the majority of words in the text, and typically only once. This implies that readers are getting a foveal glimpse (for about a quarter of a second) of most words in the text and that the essence of skilled reading behavior is contained in the processing that is performed during that glimpse. Therefore, quite understandably, explaining how literate adults read single words has been one of the major goals of experimental psychology since the very inception of this science (Huey, 1908).

The process of silent word reading (reading for meaning) minimally requires two types of codes: **orthography** (knowledge about letter identities and letter positions) and **semantics** (knowledge about the meanings of words). The process of reading aloud minimally requires an orthographic code and a **phonological** (knowledge about the sounds of words) code in order to generate a pronunciation. Although no more than two codes are necessarily required for each task, it has become increasingly clear that all three codes (orthography, semantics, and phonology) are involved in both silent reading and reading aloud. This has led to the development of a generic architecture for word recognition that emphasizes the key role for cross-code interactions (e.g., Grainger & Ziegler, 2008; Siedenberg & McClelland, 1989). Much research on single word reading to date has therefore focused on the processing of semantic, phonological, and **morphological** (knowledge of word parts that carry meaning like prefixes and suffixes) information, while largely ignoring orthographic processing. This research bias was also exaggerated by an undue focus on the process of reading aloud as opposed to silent reading for meaning. The last decade, however, has been to witness to a surge in interest for basic orthographic processing during reading; the present article aims to summarize some key findings from this recent research.

The importance of understanding orthographic processing for understanding reading in general can be best appreciated when considering the written word as both a visual object and a linguistic entity. From this perspective, single word reading is a combination of visual object identification processes and linguistic processing, with orthographic processing acting as the key interface between the two. Orthographic processing allows generic visual processing mechanisms to make contact with the linguistic processing that is specific to word stimuli compared with other kinds of visual object. This contact is established via three types of mapping: 1) letters - to - phonology - to - meaning; 2) letters - to - morphology - to - meaning; 3) letters - to - words - to - meaning (see *Figure 1*).

Title of an excellent book on the topic edited by Sally Andrews (2006).

Figure 1



Orthography as the interface between visual and linguistic processing.

2. Letter-based word recognition

There is a general consensus today among reading researchers that for languages that use an alphabetic script, visual word recognition is letter-based (see Grainger, 2008, for a summary of the arguments).[2] That is, visual feature information is used to obtain information about the word's component letters, and a word's identity is mainly derived from information about letter identities and letter positions as opposed to word shape information that might be gained, for example, from ascending and descending letters in lowercase text. There is one key computational argument against a major role for holistic word-shape information in reading: it is more efficient to solve shape invariance at the level of individual letters (N=26) than at the level of whole words (N≈30,000). Shape invariance refers to our ability to recognize words (and other kinds of visual objects) independently of the precise visual format in which they are presented (e.g., lowercase vs. UPPERCASE; `courier font` vs. *handwriting font*). The standard explanation for this ability is that we identify visual objects via abstract representations that enable different kinds of visual information to make contact with the same object identity. Figure 2, adapted from Grainger and Dufau (2012), illustrates the computational argument for letter-based word recognition.

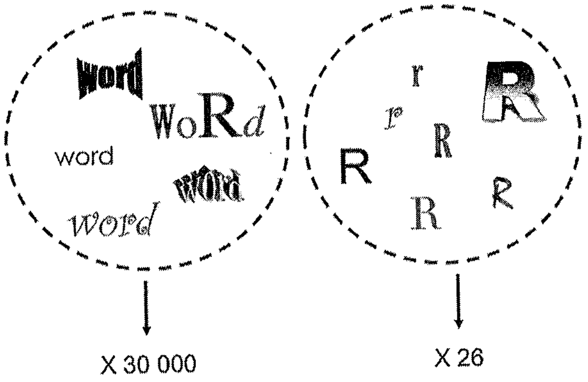
It could, however, be argued that storing different exemplars for lowercase and uppercase words is not a major computational cost and that the vast majority of fonts used in printed text vary little in terms of overall word shape. Our ability to read words in a very unusual and unfamiliar format (i.e., under extreme distortions of word shape) is there-

---

[2] It should be noted that ever since Cattell's (1886) observation that word naming is easier than letter naming (a "word superiority effect"), it was generally thought that written words were identified using holistic word-shape information, because it was not obvious how word recognition could be letter-based if it is harder to read letters than to read words (I will refer to this as "Cattell's conundrum"). This theoretical position was instrumental in erroneously guiding educational practice for teaching reading for the better part of the 20th century.

re 2

shape-invariance
r-based word
cation. Rather than
izing the different
ormats of about
different words,
re economical to
ize the different
ormats of 26 letters
cognize words via
t, shape-invariant,
epresentations.

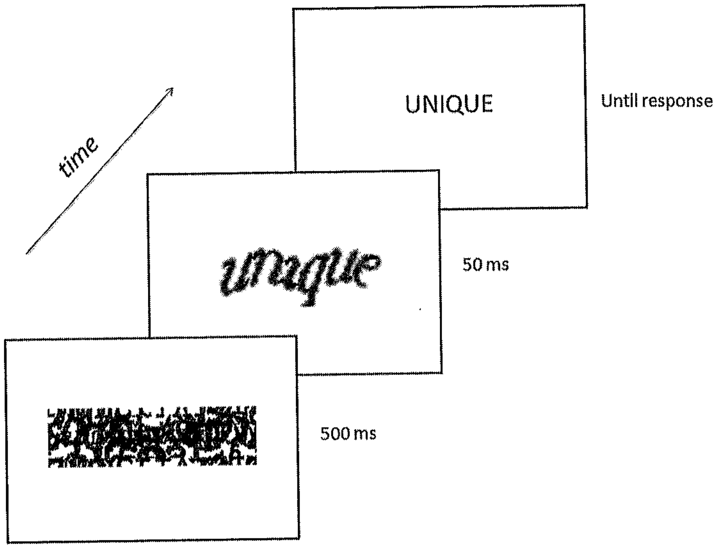


fore a key finding in this debate. One example here is our ability to read CAPTCHAs (Completely Automated Public Turing test to tell Computers and Humans Apart) when prompted to do so on an Internet site that is checking whether we are a human or a machine. These distorted versions of words are easily read by humans (see *Figure 3* for an example), but discourage brute force explorations into databases given the very large number of possible distortions. Most important is that we have shown that our ability to solve such extreme cases of shape distortion is achieved automatically without resort to slow inferential processes. This was demonstrated in a study by Hannagan, Ktori, Chanceaux and Grainger (2012), where we asked subjects to perform a word/nonword classification task with undistorted targets preceded by subliminal CAPTCHA primes, that could be the same word as the target or not (*Figure 3*). Repetition priming (i.e., faster and more accurate responses to target words preceded by primes that are the same word vs. a different word) was found for CAPTCHA primes. This finding suggests that shape invariant orthographic representations are being computed automatically and very rapidly and are therefore in line with our proposal that shape invariance is solved at the level of abstract letter representations, which would be less affected by the CAPTCHA distortions than hypothetical word-shape representations would be (see Chauncey, Holcomb & Grainger, 2008, for converging evidence obtained with masked priming and electrophysiological recordings, and Gil-López, Perea, Moret-Tatay & Carreiras, 2011, for a similar result with handwritten words).

There is, nevertheless, some empirical evidence that word shape information might influence skilled word reading in certain situations.[3] Thus, for example, Perea and Rosa (2002) found an advantage for lowercase compared with uppercase words in a simple lexical decision task, but only for relatively unfamiliar words. Another example was provided by

Figure 3

Masked priming with CAPTCHA primes and normal print targets (Hannagan et al., 2012). CAPTCHA prime stimuli are briefly presented and preceded by a pattern mask (formed of a random combination of segments extracted from different CAPTCHA stimuli for more effective masking of these stimuli). Target words (and nonwords) are presented in normal format and remain on the screen until participants press a response key to indicate whether the target is a word or not (lexical decision task).



Lété and Pynte (2003) who manipulated the "shape frequency" of written words, defined as the number of other words that shared the same ordering of ascending (A), descending (D), and neutral (N) letters. Thus a word like elephant would be coded as NANDANNA, and its shape frequency would correspond to the number of other words with the same shape code (see Walker, 1987). Lété and Pynte (2003) found an effect of shape frequency on lexical decision latencies to a set of relatively long (7-9 letters) low-frequency French words, such that words with rare shapes were easier to recognize than words with frequent shapes (which were composed uniquely of neutral letters).[4] This analysis points to a possible explanation for the lowercase advantage reported by Perea and Rosa (2002) in reading Spanish words, given that information about consonant-vowel status (ascenders and descenders can only be consonants) might be particularly useful for reading in a syllabically structured language like Spanish. In other words, prior research claiming to provide evidence for a role for word shape information in reading, might actually have been showing how letter shape information can facilitate certain **sublexical** (smaller than a word) processes such as consonant-vowel classification (see Chetail & Content, 2012, for a demonstration of such influences on visual word recognition).

In the remainder of this article we will assume that most of the information used by skilled readers to silently read words for meaning concerns information about abstract (i.e., case and font independent) letter identities and information about letter positions (i.e., orthographic information). However, before beginning our examination of letter-based reading, it should be noted that the solution to Cattell's conundrum (how can we read words via their constituent letters if it is harder to read individual letters than

---

Evidence for so-called "logographic" reading in beginning readers is one example of the use of word shape rmation. However, I follow the general consensus in seeing this as a transitional phase that is rapidly abandoned as nt orthographic processing develops and reading vocabulary increases (Share, 1995). Nevertheless, I acknowledge that ain brand names could be examples of such logographic reading in adults (see Perea et al., 2015).

[4] However, see Paap, Newsome, and Noel (1984) for a failure to find an effect of shape frequency.

to read words?) was provided by theoretical advances (e.g., McClelland & Rumelhart, 1981) showing how a word can be identified from the combination of partial information available at the level of each of its constituent letters (see Grainger, 2008, and Grainger & Dufau, 2012, for further details about the "word superiority effect" and its interpretation).

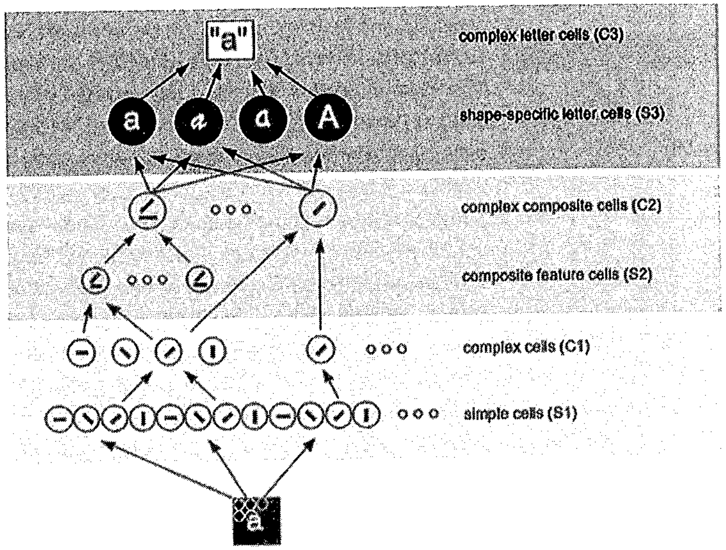
Letter identification: From pixels pandemonium

Letter-based word recognition requires processing of letter identities and letter positions. In this section, I first review current knowledge with respect to the processes involved in letter identification before examining work on letter position coding in the following section. According to Grainger, Rey, and Dufau (2008), it is the seminal work of Oliver Selfridge (Selfridge, 1959; Selfridge & Neisser, 1960) that laid the foundations for a cognitive theory of letter perception. In Selfridge's "pandemonium" model, letter identification is achieved by hierarchically organized layers of feature and letter detectors. Support for such a hierarchical organization was provided at that time by neurophysiological studies of the cat visual cortex (Hubel & Wiesel, 1962), and over the years, a general consensus has developed in favor of a generic feature-based approach to letter perception. One key guiding principle here is that isolated letter perception is just a simplified case of visual object recognition (e.g., Pelli et al., 2006). Therefore, our knowledge of visual object perception, much of which has been derived from neurophysiological studies of non-human primates, should help constrain our knowledge of letter perception in humans. This general principle is exemplified in the model presented in Figure 4. This figure shows a blueprint for a model of letter perception (Grainger et al., 2008) adapted from a classic account of object recognition (Riesenhuber & Poggio, 1999; see Dehaene, Cohen, Sigman, & Vinckier, for an extension of this approach to visual word recognition). What is the evidence in favor of such an approach, and what might be the nature of the sub-letter features involved in letter identification?

The confusion matrix is the traditional method used to hunt for features. In a typical experiment used to generate a confusion matrix, isolated letters are presented in data-limited conditions (brief exposures and/or small visual angle and/or low luminance and/or masking), and erroneous letter reports are noted. Error rate (e.g., reporting F when E was presented) is hypothesized to reflect visual similarity driven by shared features. An analysis of the pattern of letter confusions was therefore expected to reveal the set of features used to identify letters. There are more than 70 published studies on letter confusability (see Mueller & Weidemann, 2012, for a review), and some have formed the basis of concrete proposals of lists of features for letters of the Roman alphabet, mainly consisting of lines of different orientation and curvature (Gibson, 1969; Geyer & DeWald, 1973; Keren & Baggen, 1981). Two more recent studies have applied arguably improved methodologies for measuring the complete similarity space of Roman letters (Courrieu, Farioli, & Grainger, 2004; Mueller & Weidemann, 2012).

Figure 4

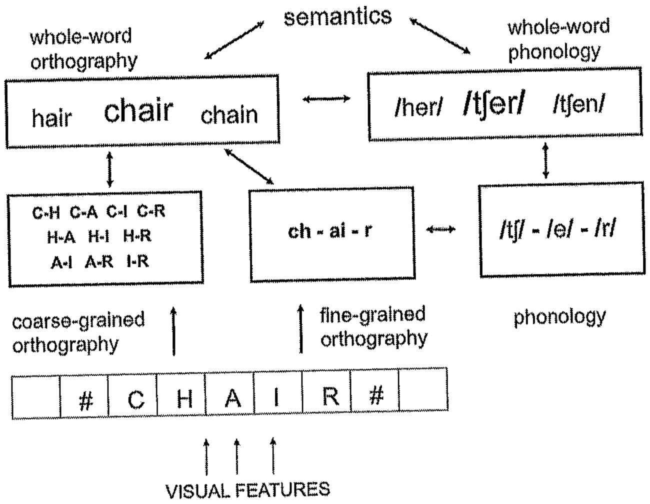
Adaptation of Riesenhuber and Poggio's (1999) model of object identification to the case of letter perception (Grainger et al., 2008). Information about simple visual features (lines of different orientation at precise locations in the visual field) extracted from the visual stimulus is progressively pooled across different locations (complex cells) and feature combinations (composite cells) as one moves up the processing hierarchy.



Another line of research has applied Gosselin and Schyns' "bubbles" technique (2001) to explore the nature of the critical features for letter perception. The classification images obtained by Fiset et al. (2007) for 26 lowercase and 26 uppercase Roman letters in Arial font revealed several important pieces of evidence. First, on average only 32% of the printed area of uppercase and 24% of lowercase letters was used by observers to identify letters, and the greatest proportion of useful information was apparent in the 2-4 cycles per letter frequency band, in line with estimates from critical-band masking studies (Solomon & Pelli, 1994). Second, the analysis revealed that terminations were by far the most diagnostic piece of information for letter identification, with intersections and horizontal lines providing further significant sources of information for uppercase letters. For example, the letter W was mainly distinguished from other letters by the presence of two terminations, one in the upper left corner and the other in the upper right corner. Finally, computational modeling has revealed that the diagnostic features used by human observers closely match those extracted by a simple two-layered associative network trained to identify letters from a pixel input (Hannagan & Grainger, 2013).

Parallel independent ter processing

Although the debate is still ongoing, another general consensus that has arisen among reading researchers over the years is that orthographic processing of written words by skilled readers is performed in parallel across all letters of the word, within the limits imposed by visual acuity and crowding (e.g., Adelman, Marquis, & Sabatos-DeVito, 2010; McCelland & Rumelhart, 1981). Parallel processing of letter identities is therefore thought to be the basis of efficient orthographic processing and reading, but how is this achieved? One solution is to align a set of individual letter detectors, such as described in the previous section, in order to form a horizontally arranged bank of letter detectors that can operate in parallel. This is the starting point of Grainger and van Heuven's (2003) model of orthographic processing, and has been retained in more recent developments of this approach (Grainger & Ziegler, 2011; Grainger, Dufau, & Ziegler, 2016). This account of orthographic processing during single word reading is shown in Figure 5.



ure 5

al-route approach to graphic processing nger & Ziegler, 2011). nk of location-specific, -centered letter ctors (bottom) send ation forward to two s of sublexical location- iant orthographic esentations: 1) *coarse- ned* representations code for the ence of informative r combinations in absence of precise tional information, 2) *fine-grained* esentations that e for the presence of uently co-occurring er combinations.[1] coarse-grained code imizes the mapping rthography to antics by selecting er combinations that the most informative h respect to word ntity, irrespective etter contiguity. fine-grained code imizes processing via *chunking* of frequently occurring contiguous er combinations such complex graphemes affixes that are used access phonological d morphological resentations respectively orphology is not shown re to avoid clutter – see ure 1).

Figure 5 describes the mapping of visual information onto whole-word orthographic representations. From there, the whole-word orthographic representations connect to the meaning of words. Sublexical orthographic representations between visual representations and whole-word representations differ in terms of how positional information is encoded for letters. Indeed, one central hypothesis in this approach is that the initial encoding of letter position information is achieved using gaze-centered coordinates. That is, when looking at the word *table* with eye fixation on the letter |b|, the letter |t| is coded as being located two letter positions to the left of fixation. As noted by Grainger and Ziegler (2011), the "hard problem" in orthographic processing is therefore to understand how such location-specific gaze-centered orthographic representations are transformed into location-invariant word-centered representations. In order to read the word *table* it is important to know where the letter |t| is in the word, not where it is on the retina.

There are several different ways that one can code for within-word letter position (see Grainger, 2008, for a review, and Davis, 2010, Gomez, Ratcliff, & Perea, 2008, for different approaches). Here we build on a solution first proposed by Mozer (1987) and further developed by Whitney (2001) and Grainger and van Heuven (2003). In this particular solution, within-word letter position is coded by an unordered set (a bag) of **n-grams** (ordered letter combinations) while allowing for non-contiguous combinations that respect relative position in the word (such as C-A, C-I, in the word "chair", see *Figure 5*). In the simplest version of this approach (i.e., a bigram model), the contiguous and non-contiguous ordered letter combinations are referred to as open-bigrams. In the approach to orthographic processing described in Figure 5, there are two different types of constraints that affect processing along the two orthographic processing routes. Both types of constraints are driven by the frequency with which different combinations of letters occur in written words. On the one hand, frequency of occurrence determines the probability with which a given combination of letters belongs to the word being read. Letter combinations that are encountered less often in other words are more diagnostic of the identity of the word being processed. In the extreme, a combination of letters that only occurs in a single word in the language, and is therefore a rarely occurring event when considering the language as a whole, is completely informative with respect to word identity. On the other hand, frequency of co-occurrence enables the formation of higher-order representations (chunking) in order to diminish the amount of information that is processed, via data compression (i.e., explaining away). Letter combinations that often occur together can be usefully grouped to form higher-level orthographic representations such as multi-letter graphemes (<th>, <ch>) and morphemes (ing, er), thus providing a link with pre-existing phonological and morphological representations during reading acquisition (see *Figures 1 & 5*).

The coarse-grained orthographic representations in this approach bring a certain amount of flexibility to the way that within-word letter position information is represented. Indeed, the concept of open-bigrams was initially developed (Grainger & van Heuven, 2003; Whitney, 2001) to account for specific phenomena observed in the behavior of skilled readers, that pointed to the need for flexible coding of letter position information (see Grainger, 2008, for a review). One such phenomenon, observed using the popular masked priming technique (Forster & Davis, 1984), is referred to as relative-position priming, whereby word identification is improved by the prior brief presentation of an orthographically related

---

The terms arse-grained" and ne-grained" refer to the ecision with which letter- -word order information encoded: either flexibly arse-grained) or precisely ne-grained).

prime stimulus formed of a subset of the target word's letters that maintain their correct relative position in the stimulus (e.g., "grdn" as a prime for "garden").[5] Crucially, transposing the two inner letters of the prime stimulus (e.g., "gdrn") cancels the priming effect measured relative to a completely unrelated prime stimulus (Peressotti & Grainger, 1999). Furthermore, providing absolute position information (e.g., "g-rd-n") does not increase priming effects (Grainger et al., 2006; Grainger & Holcomb, 2009).

The kind of coarse orthographic coding shown in Figure 5 accounts for these findings by the fact that in prime stimuli like "grdn", all of the prime's bigrams are contained in the target "garden", whereas a prime stimulus like "gdrn" provides evidence for one bigram (D-R) that is not present in the target. It is interesting to note, however, that while this research was being performed and the notion of flexible coding of letter position information being developed, an interesting email started to circulate in 2003. This is the "Cambridge University" email[6], according to which "it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe." There are two important points to note with respect to this anecdotal evidence. First our ability to read such text was a key prediction of models that use flexible within-word coding of letter position, such as the open-bigram scheme (Grainger & Whitney, 2004), and this prediction has been supported by hundreds of laboratory experiments run since then (e.g., Perea & Lupker, 2004; Rayner, White, Johnson, & Liversedge, 2006; Schoonbaert & Grainger, 2004; see Grainger, 2008, for a review). Second contrary to the claims of the Cambridge University email, our ability to recover word identity from such transposed-letter stimuli constitutes key evidence for letter-based word recognition and evidence against the use of more holistic information.

. Letter-specific processing?

Within the general framework of neuronal recycling theory (Dehaene & Cohen, 2007), learning to read involves the adaptation of general purpose visual processing mechanisms to the specificities of written words. That is, the mechanisms employed to identify everyday objects, such as tables and chairs, must be adapted to the special nature of written words as visual objects that also need to be identified for the purposes of efficient print-to-meaning translation during skilled reading. Within the account of orthographic processing described in section 3, there are two issues at odds with basic visual object processing: 1) the hypothesized specialized bank of gaze-centered letter, and digit, detectors and 2) the mechanism used to code for within-word letter positions. In this section I will examine the evidence in favor of such letter-specific processing.

First of all, the constraints of parallel letter processing are thought to impose changes at the level of location-specific letter representations in order to reduce crowding and optimize information uptake. Following Tydgat and Grainger (2009), we would argue that letters and digits are processed alike at this level, via a horizontally aligned bank of letter/digit detectors (Grainger & van Heuven, 2003). This is because reading words and numbers can be optimized by parallel processing of the component letters/digits. It is this level of processing that is affected by visual factors such as visual acuity and crowding; the evidence suggests that letters and digits are indeed processed in the same way at this level. This evidence was obtained in experiments where subjects were asked to identify a single character in a string of 5 characters with eye fixation on the central character (Tydgat & Grainger, 2009). In Tydgat and Grainger's (2009) study, the 5 characters were presented very briefly (200 ms), and subjects had to indicate the identity of a single character that had just been presented at a specified location (see *Figure 6*). In experiments like this, accuracy is typically highest at the first, central, and final positions for letters and digits. Symbols and simple shapes, on the other hand, show maximum performance at the central position, and performance tends to decrease from the center outwards (see also Mason, 1982; Hammond & Green, 1982). This pattern of results is shown in Figure 6.

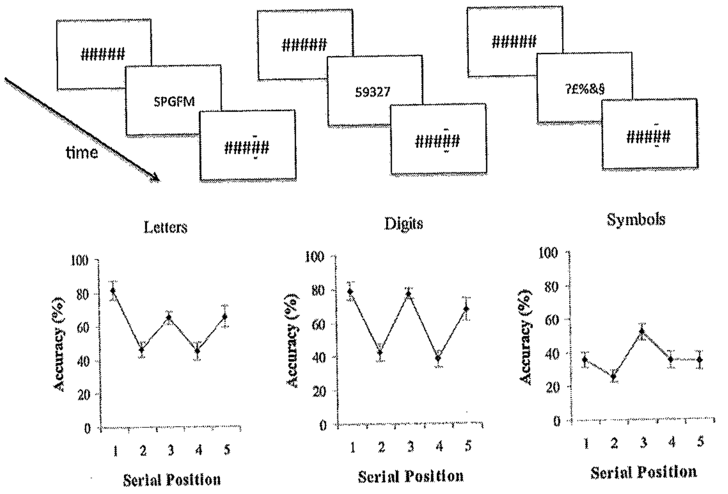


Figure 6

Results of Tydgat and Grainger's (2009) study comparing identification of letters, digits, and symbols within a string composed of the same elements. Subjects see a briefly presented string in between two pattern masks (#####) and have to indicate which character was present in the string at the location cued by the horizontal bars (4th position in the examples).

The different serial position functions shown in Figure 6 can be explained by differences in the way crowding affects letters and digits compared with other types of visual stimuli. First of all, the fact that accuracy is higher for the central position compared with the 2nd and 4th positions

---

See Carreiras, Duñabeitia, and Molinaro (2009) for a discussion of the role of consonant-vowel status in lative-position priming effects.

This email claimed to report on research performed at Cambridge University, but no such research had in fact een carried out at that university at that time.

in strings can be accounted for by differences in visual acuity as a function of eccentricity (i.e., distance from fixation). This would operate identically for all kinds of stimuli. Key differences arise at the first and last positions in the string, with only letter and digit stimuli showing greatly improved identification relative to the 2nd and 4th positions. This specific pattern can be explained by greater crowding for symbol stimuli, such that a single flanking element suffices to generate almost maximum crowding for symbol targets at the outer positions in the string, whereas letter and digit stimuli would benefit from reduced crowding at these locations (Grainger, Tydgat, & Isselé, 2010). As argued above, the reduced crowding for letters and digits arises from adaptation to the hyper-crowding imposed by the parallel processing of such characters.

Recent research with beginning readers points to a special status of the first letter in words. Indeed, the advantage for outer letters in letter-in-string identification is almost always accompanied by a further advantage for the first letter compared with the last letter in the string (Tydgat & Grainger, 2009; see *Figure 6*). In an unpublished developmental study, we have shown that it is performance in identifying the first letter in strings of random consonants that improves as a function of reading ability, and this increase in performance to initial letters contrasts sharply with the lack of change in identifying simple familiar shapes at the first position in a string of shape stimuli. On the basis of this finding, we have argued that the first-letter advantage seen in skilled adult readers (e.g., Marzouki & Grainger, 2014; Scaltritti & Balota, 2014) results from adaptive mechanisms operating during the process of learning to read in order to prioritize processing of the first letter in words.

Strings of letters and digits might therefore be processed similarly at the level of location-specific character detectors, since visual factors such as acuity and crowding have a similar impact on these stimuli. However, differences in the way these two kinds of stimuli are processed emerge at the next level of processing, where positional information is coded relative to the object (word or number) and independently of where the object is (i.e., knowing that there is a "T" at the beginning of the word, or the digit "5" at the beginning of a number, independently of where the stimuli are in the visual field). Given that only letter stimuli are systematically associated with higher-level familiar objects, we hypothesize that only letter stimuli develop the kind of flexible coding of positional information such as provided by ordered combinations of contiguous and non-contiguous elements in a string. This is simply because such approximate position coding is good enough to know a word's identity with a relatively high probability (Dandurand et al., 2011), but it is not very good for obtaining magnitude information from a number. For the latter, one requires more precise order information, such that the identity of the different digits can be accurately associated with each position in the number. The precision required for number processing is like the precision required for the sublexical translation of print-to-sound during reading aloud. It is therefore the potential use of a more flexible object-centered position code that distinguishes letter-strings from numbers. This leads us to hypothesize, somewhat counter-intuitively, that position coding for strings of letters might in certain conditions (i.e., when object-centered coding is required for the task) be less precise[7] than for strings of less familiar stimuli. Is there any evidence for this?

One paradigm that has proved useful for comparing position coding for different kinds of visual stimuli is the same-different judgment task. In this task, two stimuli are briefly presented in rapid succession, and subjects have to decide as rapidly and as accurately as possible if the stimuli are the same or are different (see *Figure 7*). This paradigm has been recently applied to examine similarities and differences in processing strings of letters, digits, and symbols. In a typical experiment, a subject will see a string of characters such as PGFM for 300 ms, which is immediately replaced one line below by a second string such as PFGM again for 300 ms, and the subject presses one response key for a "same" response and another response key for a "different" response. Recent research has specifically examined response times and error rates to respond "different" to pairs of characters differing by a transposition of two characters (PGFM - PFGM) or differing by the substitution of two characters (PGFM – PDRM). It has been shown (Duñabetia, Dimitropolou, Grainger, Hernandez & Carreiras, 2012) that detecting a transposition change is harder than detecting a substitution change, and that this effect is greater for letter strings compared with both digit strings (e.g., 3842 - 3482) and symbol strings (%&?# - %?&#). The behavioral results of this study are shown in Figure 7 (see Massol, Duñabeitia, Carreiras & Grainger, 2013, for further evidence obtained with the same paradigm).

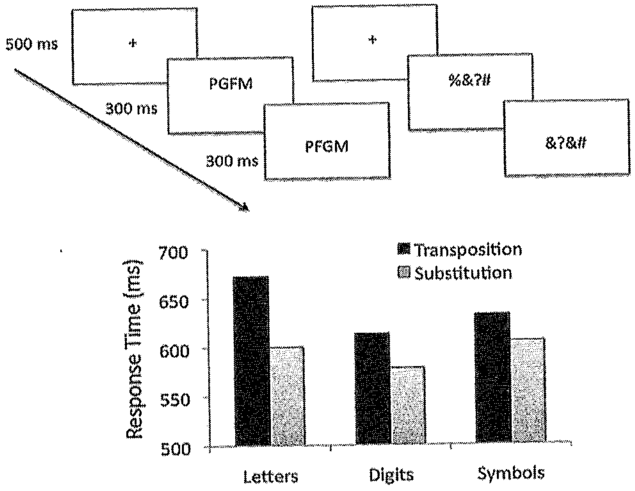


Figure 7

Behavioral results of Duñabetia et al. (2012) and the procedure of the same-different judgment task (decide as rapidly as possible whether the two strings are the same or not) illustrated for letters and symbols (the study also included digit stimuli). The figure shows that it is harder (longer response times) to indicate that two strings are different when the difference is induced by transposing two characters compared with substituting two characters. Most important is that letter strings show significantly greater transposition costs (i.e., the difference between the transposition and substitution conditions) than the other two types of stimuli.

---

[7] An important distinction must be drawn between positional flexibility and positional noise. The hypothesis here is that orthographic processing endows a greater flexibility in position coding for an equivalent amount of noise in the system.

I have argued that the greater transposition cost seen with letter stimuli arises because an object-centered positional code is used to inform responses in the same-different judgment task and that letter stimuli are coded with a more flexible position coding mechanism than other kinds of stimuli. This leads to the somewhat paradoxical situation whereby the most familiar stimuli (most of us read much more than we do arithmetic) generate the poorest performance. The explanation we offer for this pattern of results is cast within the dual-route framework for orthographic processing (Grainger & Ziegler, 2011) shown in Figure 5. Only letter stimuli use relative character position coding, since numbers require precise position coding in order to accurately retrieve magnitude information from a string of digits. The model also predicts, however, that letters and digits are processed by the same machinery at the level of location-specific character detectors, and evidence in favor of this has been provided by experiments using character-in-string identification (e.g., Tydgat & Grainger, 2009; see *Figure 6*).

. Orthographic processing and sentence reading

In this final section I will examine how basic mechanisms of orthographic processing, as described in the preceding sections, might happen during sentence reading in preparation for the semantic and syntactic processing that is necessary for sentence-level comprehension.

Grainger et al. (2016) described a theoretical framework for parallel orthographic processing during sentence reading inspired by recent evidence in favor of the spatial integration of orthographic information spanning multiple words. More specifically, this theoretical framework, shown in Figure 8, was motivated by recent findings suggesting that orthographic information extracted from several words in parallel is integrated into a single processing channel. These findings were from eye tracking studies of the influence of parafoveal stimuli on the processing of the fixated foveal stimulus. In one such study, participants read sentences for meaning, and their eye movements were recorded. Unbeknownst to participants, while their eyes move from one word to the next, a key word in a sentence was changed from one word to another just before fixating on the word of interest. Changing text on screen is not easily detected by participants because the change happens during a saccade when it isn't noticed; this is referred to as the *boundary technique* because the word changes when the participant's eye crosses an unseen boundary (Rayner, 1975). The initial word can be manipulated to have certain characteristics or to share certain features with the word it changes to or not. The word it changes to is always the correct word for the sentence.

Two sentence reading studies using the boundary technique have shown that information extracted from the to-be-fixated parafoveal initial word influences the processing of the word before the unseen boundary (Angele, Tran, & Rayner, 2013; Dare & Shillcock, 2013). This finding is in line with prior observations of *parafoveal-on-foveal* effects during reading (e.g., Dimigen, Kliegl, & Sommer, 2012; Vitu, Brysbaert, & Lancelin, 2004). The key difference with respect to older studies is that the recent studies have shown that orthographic relatedness of the word before the boundary and the initial word after the boundary affects processing of the changed word. In Angele et al.'s (2013) study, participants read sentences such as "the store had a coat / coat …", where the 2nd occurrence of "coat" is replaced by the word "sale" when the eyes leave the 1st occurrence of "coat". Participants were faster at reading "coat" in this context compared with a sentence like "the store had a coat / milk …" (with "milk" being replaced by "sale" when the eyes leave "coat").[8] Using the same method, Dare and Shillcock (2013) found facilitation from parafoveal nonword stimuli after the boundary formed by transposing two letters of the foveal word (e.g., the store had a coat / caot …) compared with a double-substitution control condition (e.g., the store had a coat / ceit …). These results clearly suggest that orthographic information extracted in parallel from the fovea and the parafovea collectively influences the process of foveal word recognition.

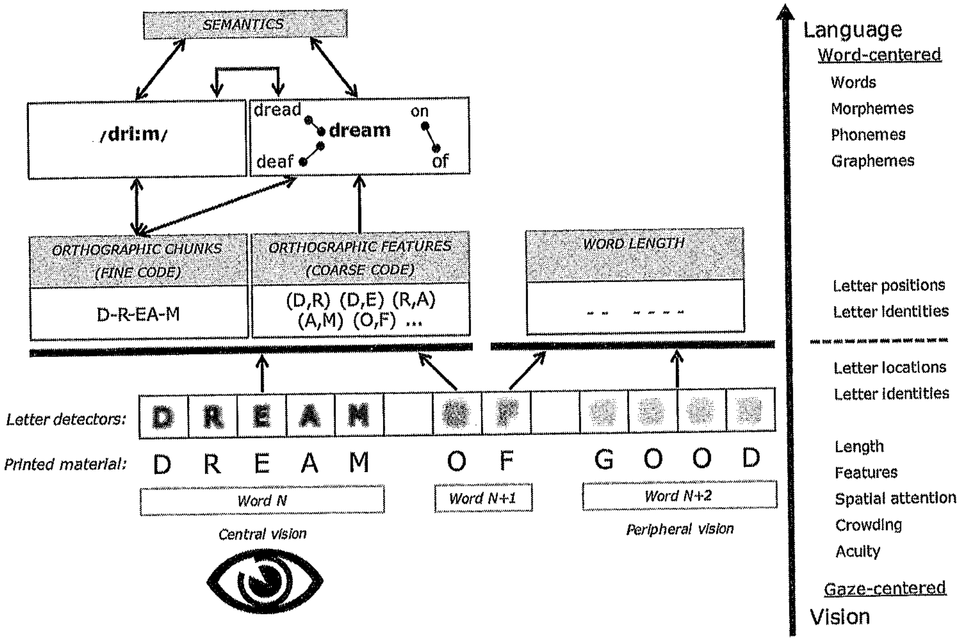


Figure 8

Architecture for orthographic processing during sentence reading proposed by Grainger et al. (2016). Gaze-centered letter detectors process visual information extracted from several words in parallel within the limits imposed by visual acuity, crowding, and spatial attention. These letter detectors feed-forward information into a single pool of word-centered orthographic representations that enable one word representation to emerge as the best bet given the incoming evidence. This winner-take-all mechanism is implemented via lateral inhibitory connections (lines and filled circles) between co-activated word representations.

---

[8] It is crucial to understand the distinction between these parafoveal-on-foveal effects and parafoveal preview benefits. In parafoveal preview experiments, it is the influence of a parafoveal "prime" stimulus on processing of the word after the unseen boundary. The observed effects therefore reflect ***temporal integration*** of information associated with the same spatiotopic location. Parafoveal-on-foveal effects, on the other hand, reflect the ***spatial integration*** of information before and after the unseen boundary on the word before the boundary.

In order to account for these and related findings, I and others have proposed that there is some form of spatial integration of orthographic information that is extracted in parallel from several words (Angele et al., 2013; Grainger, Mathôt, & Vitu, 2014). Thus, when fixating a word, orthographic information is extracted in parallel from that word and the next word, and this information pooled such that orthographic overlap across the two words facilitates processing of the word being fixated (see *Figure 8*). Further crucial evidence for such spatial integration of orthographic information has been obtained from the novel *"flanking letters lexical decision"* (FLLD) task. In this paradigm, centrally located word and nonword stimuli are flanked by letters located to the left and to the right and separated from the central stimulus by a space. Participants are asked to decide whether the central stimulus is a word or not and can therefore ignore the flanking stimuli. In the first study to use the FLLD task, Dare and Shillcock (2013) found faster lexical decision times to central targets when the flanking letters were the same as in the target: "RO ROCK CK" vs. "DA ROCK SH". More surprisingly, however, they found that the order of the shared bigrams did not matter. Thus lexical decisions to the word ROCK were the same in the following conditions: "RO ROCK CK" and "CK ROCK RO". This key finding rules out an explanation couched in terms of letter migrations induced by positional noise, since if this were the case, priming effects should have been greater with bigrams in the correct order. This result points to spatial integration of orthographic information across word boundaries into a single channel for orthographic processing, as illustrated in Figure 7. Furthermore, Grainger et al. (2014) showed that although bigram order does not impact on flanking letter effects, hence replicating Dare and Shillcock (2013), the order of letters within a bigram does matter. Thus, there was greater facilitation in the "RO ROCK CK" condition than in the "OR ROCK KC" condition that they tested.

Within the framework proposed by Grainger et al. (2016), visibility constraints operating of gaze-centered letter detectors ensures that the most activated word in the single channel is indeed the word being fixated. Nevertheless, there is evidence that skilled readers are capable of keeping track of the spatial locations of different word identities in parallel. Therefore, although orthographic information might be initially pooled into a single channel, as illustrated in Figure 7, the system must be able to keep track of which letters/bigrams belong to which words. This will enable the orthographic processor to output word identities that are tied to a particular position in the phrase/sentence. Sentence comprehension requires access to semantic and syntactic information from the different words in the sentence (when available) and information about the positions of the words in the sentence (or sentence constituent). Whole-word orthographic representations provide access to the semantic and syntactic information associated with words. This orthographic processing module therefore outputs the three key ingredients for higher-level processing: semantic information, syntactic information, and word-in-phrase position.

References

Adelman, J.S., Marquis, S.J., & Sabatos-DeVito, M.G. (2010). Letters in words are read simultaneously, not in left-to-right sequence. *Psychological Science, 21*, 1799-1801.

Andrews, S. (2006). *From inkmarks to ideas: Current issues in lexical processing*. Hove, UK: Psychology Press.

Angele, B., Tran, R., & Rayner, K. (2013). Parafoveal-foveal overlap can facilitate ongoing word identification during reading: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 526-538.

Carreiras, M., Duñabeitia, J. A., & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. Cerebral Cortex, 19, 2659-2670.

Cattell, J. (1886). The time it takes to see and name objects. *Mind,* 11, 53-65.

Chauncey, K., Holcomb, P.J., & Grainger, J. (2008). Effects of stimulus font and size on masked repetition priming: An ERP investigation. *Language and Cognitive Processes*, 23, 183-200.

Chetail, F. & Content, A. (2012). The internal structure of chaos: Letter category determines visual word perceptual units. *Journal of Memory and Language*, 67, 371-388.

Courrieu, P., Farioli, F., & Grainger, J. (2004). Inverse discrimination time as a perceptual distance for alphabetic characters. *Visual Cognition, 11*, 901-919.

Dandurand, F., Grainger, J., Duñabeitia, J.A., & Granier, J.P. (2011). On coding non-contiguous letter combinations. *Frontiers in Cognitive Sciences*, 2:136. doi: 10.3389/ fpsyg.2011.00136

Dare, N., & Shillcock, R. (2013). Serial and parallel processing in reading: Investigating the effects of parafoveal orthographic information on nonisolated word recognition. *The Quarterly Journal of Experimental Psychology, 66*, 417-428.

Davis, C.J. (2010). The spatial coding model of visual word recognition. *Psychological Review, 117*, 713-758.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*, 384-398.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences, 9*, 335-341.

Dimigen, O., Kliegl, R., & Sommer, W. (2012). Trans-saccadic parafoveal preview benefits in fluent reading: A study with fixation-related brain potentials. *NeuroImage*, 62, 381-393.

Duñabeitia, J.A., Dimitropoulou, M., Grainger, J., Hernández, J.A., & Carreiras, M. (2012). Differential sensitivity of letters, numbers and symbols to character transpositions. *Journal of Cognitive Neuroscience, 24,* 1610-1624.

Fiset, D., Blais, C., Ethier-Majcher, C., Arguin, M., Bubu, D., & Gosselin, F. (2007). Features for uppercase and lowercase letter identification. *Psychological Science*, in press.

Geyer, L.H., & DeWald, C.G. (1973). Feature lists and confusion matrices. *Perception & Psychophysics, 14,* 471-482.

Gibson, E.J. (1969). *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts.

Gil-López, C., Perea, M., Moret-Tatay, C., & Carreiras, M. (2011). Can masked priming effects be obtained with handwritten words? *Attention, Perception, & Psychophysics, 73,* 1643–1649

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115,* 577-601.

Gosselin, F., & Schyns, P.G. (2001). Bubbles: A technique to reveal the use of information in recognition. *Vision Research, 41,* 2261-2271.

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes, 23,* 1-35.

Grainger, J. & Dufau, S. (2012). The front-end of visual word recognition. In J.S. Adelman (Ed.) *Visual Word Recognition Vol. 1: Models and Methods, Orthography and Phonology.* Hove, UK: Psychology Press.

Grainger, J., Dufau, S., & Ziegler, J.C. (2016). A vision of reading. *Trends in Cognitive Sciences*, 20, 171-179.

Grainger, J., Granier, J.P., Farioli, F., Van Assche, E., & van Heuven, W.J. (2006). Letter position information and printed word perception: the relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 865-884.

Grainger, J. & Holcomb, P.J. (2009). An ERP investigation of orthographic priming with relative-position and absolute-position primes. *Brain Research*, 1270, 45-53.

Grainger, J., Mathôt, S., Vitu, F. (2014). Tests of a model of multi-word reading: Effects of parafoveal flanking letters on foveal word recognition. *Acta Psychologica, 146,* 35-40..

Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: from pixels to pandemonium. *Trends in Cognitive Sciences, 12,* 381-387.

Grainger, J., Tydgat, I., & Isselé, J. (2010). Crowding affects letters and symbols differently. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 673-688.

Grainger, J., & van Heuven, W. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *The Mental Lexicon.* New York: Nova Science Publishers (pp. 1-24).

Grainger, J., & Whitney, C. (2004). Does the huamn mind raed wrods as a wlohe? *Trends in Cognitive Sciences, 8,* 58-59.

Grainger, J., & Ziegler, J. (2008). Cross-code consistency effects in visual word recognition. In E. L. Grigorenko & A. Naples (Eds.) *Single-word reading: Biological and behavioral perspectives* (pp. 129-157). Mahwah, NJ: Lawrence Erlbaum Associates.

Grainger, J., & Ziegler, J.C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology*, 2:54. doi: 10.3389/fpsyg.2011.00054.

Hammond, E.J., & Green, D.W. (1982). Detecting targets in letter and nonletter arrays. *Canadian Journal of Psychology, 36,* 67-82.

Hannagan, T. & Grainger, J. (2013). Learning diagnostic features: The delta rule does bubbles. *Journal of Vision*, 13(8): 17, 1-11.

Hannagan, T., Ktori, M., Chanceaux, M., & Grainger, J. (2012). Deciphering CAPTCHAs: What a Turing test reveals about human cognition. *PLoS ONE*, 7(3), e32121.

Hubel, D., & Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology of London, 160,* 106-154.

Keren, G., & Baggen, S. (1981). Recognition models of alphanumeric characters. *Perception & Psychophysics, 29,* 452-466.

Lété, B. & Pynte, J. (2003). Word-shape and word-lexical-frequency effects in lexical decision and naming tasks. *Visual Cognition, 10,* 913-948.

Marzouki, Y. & Grainger, J. (2014). Effects of stimulus duration and inter-letter spacing on letter-in-string identification. *Acta Psychologica, 148,* 49–55.

Mason, M. (1982). Recognition time for letters and nonletters: Effects of serial position, array size, and processing order. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 724-738.

Massol, S., Duñabetia, J.A., Carreiras, M., & Grainger, J. (2013). Evidence for letter-specific position coding mechanisms. *PLoS ONE*, 8(7): e68460. doi 10.1371/journal.pone.0068460

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.

Mozer, M. (1987). Early parallel processing in reading: A connectionist approach. In M. Coltheart (Ed.) *Attention and Perfomance XII: The Psychology of Reading*. (pp. 83-104). Hillsdale, NJ: Lawrence Erlbaum.

Mueller, S. T., & Weidemann, C. T. (2012). Alphabetic letter identification: Effects of perceivability, similarity, and bias. *Acta Psychologica*, *139*, 19-37.

Pelli, D.G., Burns, C.W., Farrell, B., & Moore-Page, D.C. (2006). Feature detection and letter identification. *Vision Research*, *46*, 4646-4674.

Perea, M., & Lupker, S.J. (2004). Can caniso activate casino? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, *51*, 231–246.

Perea, M., & Rosa, E. (2002). Does "whole-word shape" play a role in visual word recognition? *Perception & Psychophysics*, *64*, 785-794.

Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception and Psychophysics*, *61*, 691–706.

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

Rayner, K., White, S.J., Johnson, R.L., & Liversedge, S.P. (2006).Raeding wrods with jubmled lettres: there is a cost. *Psychological Science*, *17*, 192-193.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*, 1019-1025.

Scaltritti, M. & Balota, D.A. (2013). Are all letters processed equally in parallel? Further evidence of a robust first-letter advantage. *Acta Psychologica*, *144*, 397-410.

Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, *19*, 333–367.

Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, develop- mental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Selfridge, O.G., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American*, *20*, 60-68.

Selfridge, O.G. (1959). Pandemonium: A paradigm for learning. In D. V. Blake and A. M. Uttley (Eds.), *Proceedings of the Symposium on Mechanisation of Thought Processes*, (pp 511-529). London : H. M. Stationary Office.

Solomon, J.A., & Pelli, D.G. (1994). The visual filter mediating letter identification. *Nature*, *369*, 395-397.

Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits and symbols. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 480-498.

Vitu, F., Brysbaert, M. & Lancelin, D. (2004). A test of parafoveal-on-foveal effects with pairs of orthographically related words. *European Journal of Cognitive Psychology*. 16, 154–177

Walker, P. (1987). Word shape as a cue to the identity of a word: An analysis of the Kucera and Francis (1967) word list. *The Quarterly Journal of Experimental Psychology*, *39A*, 675-700.

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: the SERIOL model and selective literature review. *Psychonomic Bulletin and Review*, *8*, 221–243.

Whitney, C. & Cornelissen, P. (2008). SERIOL reading. *Language and Cognitive Processes*, *23*, 143–164.

Author

Jonathan Grainger graduated from Manchester University, UK, before obtaining his PhD in Experimental Psychology in Paris, France, in 1986. He became a CNRS research scientist at the Laboratoire de Psychologie Expérimentale, Paris, in 1988, and founded and directed (2000-2011) the Laboratoire de Psychologie Cognitive in Marseille. His research focuses on understanding basic processes in written word comprehension during reading.