

frontiers
IN LANGUAGE SCIENCES

This Provisional PDF corresponds to the article as it appeared upon acceptance, after rigorous peer-review. Fully formatted PDF and full text (HTML) versions will be made available soon.

A Dual-Route Approach to Orthographic Processing

Jonathan Grainger & Johannes C. Ziegler

CNRS & Aix-Marseille University

Short title: dual-route orthographic processing

Corresponding author:

Jonathan Grainger

LPC-CNRS

Université de Provence

3 place Victor Hugo

13331 Marseille

jonathan.grainger@univ-provence.fr

Abstract

In the present theoretical note we examine how different learning constraints, thought to be involved in optimizing the mapping of print to meaning during reading acquisition, might shape the nature of the orthographic code involved in skilled reading. On the one hand, optimization is hypothesized to involve selecting combinations of letters that are the most informative with respect to word identity (diagnosticity constraint), and on the other hand to involve the detection of letter combinations that correspond to pre-existing sublexical phonological and morphological representations (chunking constraint). These two constraints give rise to two different kinds of prelexical orthographic code, a coarse-grained and a fine-grained code, associated with the two routes of a dual-route architecture. Processing along the coarse-grained route optimizes fast access to semantics by using minimal subsets of letters that maximize information with respect to word identity, while coding for approximate within-word letter position independently of letter contiguity. Processing along the fine-grained route, on the other hand, is sensitive to the precise ordering of letters, as well as to position with respect to word beginnings and endings. This enables the chunking of frequently co-occurring contiguous letter combinations that form relevant units for morpho-orthographic processing (prefixes and suffixes) and for the sublexical translation of print to sound (multi-letter graphemes).

The starting point of the present endeavor is the traditional dual-route model of reading aloud, that distinguishes between a lexical route and a non-lexical route for transforming print to sound (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Ellis & Young, 1988; Zorzi, 2010). The lexical route is often referred to as the direct route, whereby sublexical orthographic information makes direct contact with whole-word orthographic representations, which then provide access to whole-word phonology on the one hand, and higher-level semantic information on the other. Along the so-called indirect, non-lexical route, sublexical orthographic information is first transformed into a sublexical phonological code before making contact with phonological output units, whole-word phonological representations, and semantics. In its most recent form, the dual-route approach provides a comprehensive account of phenomena related to the process of reading aloud in skilled adult readers and dyslexics (Perry, Ziegler, & Zorzi, 2007; 2010).

This general approach was adopted for silent word reading in the bi-modal interactive activation model (BIAM - Diependaele, Grainger, & Ziegler, 2010; Grainger & Ferrand, 1994; Grainger & Ziegler, 2007; Jacobs, Rey, Ziegler, & Grainger, 1998). The BIAM can be seen as a localist implementation of the generic division of labor or “triangle” approach to visual word recognition, in which there are two routes from orthography to semantics - a direct route and an indirect route via phonology (Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg, & Patterson, 1996). The specific architecture of the BIAM allows it to account for a wide range of phenomena associated with visual word recognition, and in particular, the rapid involvement of phonological codes in the process of silent word reading (Braun, Hutzler, Ziegler, Dambacher, & Jacobs, 2009; Diependaele et al., 2010; see Grainger & Ziegler, 2007, for review).

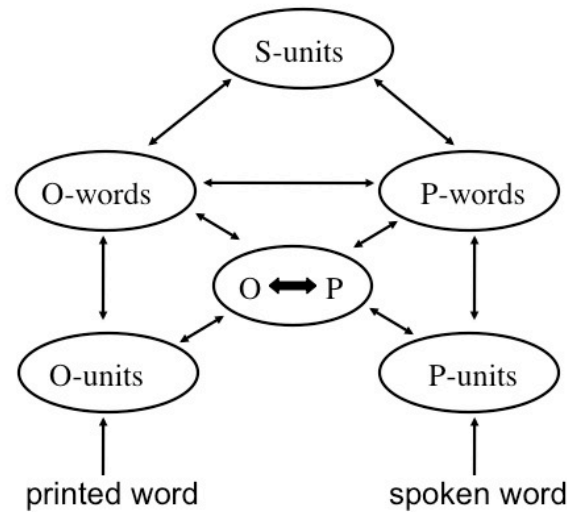


Figure 1. Generic architecture of the bi-modal interactive-activation model (BIAM). A distinction is drawn between sublexical and lexical orthographic (O-units, O-words) and phonological (P-units, P-words) representations that interact via a central interface ($O \leftrightarrow P$). Whole-word representations provide access to semantic representations (S-units).

Much of the success of this general approach lies in the application of the principle of nested incremental modeling (Grainger & Jacobs, 1996; Jacobs & Grainger, 1994; Perry et al., 2007). This principle encourages model development that builds on prior success and adjusts to prior failures. In this respect, the BIAM incorporates key aspects of McClelland and Rumelhart's (1981) interactive-activation model, as well as Grainger and Jacobs' (1996) extension of this model, which was put forward to account for a certain number of task-specific phenomena related to visual word recognition. Furthermore, it is important to note that two relatively independent lines of research, one focusing on silent reading for meaning (as in the present work), the other focusing on reading aloud (Perry et al., 2007; 2010), have converged on very similar proposals for a generic architecture of word recognition and reading aloud. It is this generic architecture that forms the basis of the present theoretical work.

One key feature of many models of visual word recognition, including the BIAM, is that there is a single type of sublexical orthographic code. Some form of word-centered letter-position code, such as the slot-coding used in the interactive-activation model (McClelland & Rumelhart, 1981), is typically applied in order to associate different letter identities with

different positions in the word. In dual-route models of reading, this unique sublexical orthographic code feeds activation forward to both whole-word orthographic representations (direct route) and sublexical phonological representations (indirect route).

In the present article, we describe a dual-route approach to orthographic processing that postulates the existence of two fundamentally different kinds of location-invariant, word-centered, sublexical, orthographic codes. These two types of orthographic codes are hypothesized to have developed as a result of the nature of the constraints that fashion the formation of orthographic representations during reading acquisition (Grainger & Dufau, 2011). The motivation for drawing such a distinction emerged from a consideration of the level of precision of letter position coding that is necessary for the successful sublexical conversion of print-to-sound on the one hand, and the growing evidence for the existence of a form of flexible, relatively imprecise sublexical orthographic code on the other. The latter form of evidence spurred the development of a number of letter position coding schemes, as alternatives to McClelland and Rumelhart's (1981) slot-based scheme and Seidenberg and McClelland's (1989) wickelgraph scheme (e.g., Davis, 2010; Gomez, Ratcliff, & Perea, 2008; Grainger & van Heuven, 2003; Whitney, 2001). All of the alternative schemes, including the one proposed by Grainger and van Heuven (2003) to be described below, involved an increased flexibility in the way letter identities are tied to within-word position.

In the following sections, we first describe the dual-route approach to orthographic processing, and how it emerged from a consideration of the constraints that arise when learning to map orthography onto semantics on the one hand, and orthography onto pre-existing sublexical morphological and phonological representations, on the other hand. We then describe how this approach can be integrated within the generic BIAM architecture, and discuss its consequences with respect to phonological and morphological influences during visual word recognition. Finally, we discuss the implications of this general approach for accounts of reading acquisition, and we discuss the possible role of attention in distinguishing between the learning of fine-grained and coarse-grained orthographic representations, a distinction that forms the backbone of our dual-route approach.

The hard problem of orthographic processing

The starting point of the vast majority of computational models of orthographic processing is a word-centered orthographic code. These models therefore avoid the hard problem of orthographic processing, that is, the transformation of location-specific retinotopic visual information into a location-invariant word-centered orthographic code. During reading, the eyes fixate the majority of words in the text, mostly just once, and information uptake from the fixated word is a function of fixation position in the word. The reader's brain therefore initially knows that the visual information associated with a given letter identity is at a particular location relative to eye fixation (i.e., retinotopic coordinates). However, identifying a unique orthographic word requires knowledge about where a given letter is in the word, not on the retina.

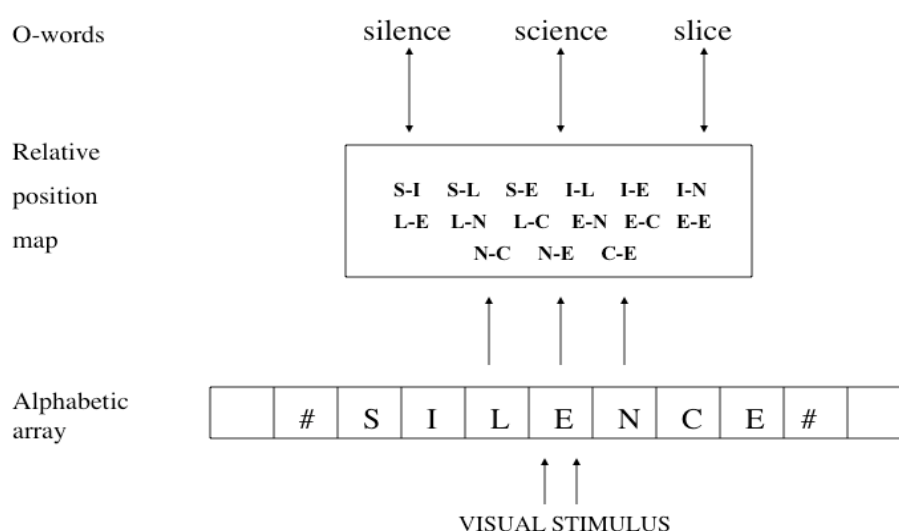


Figure 2. Grainger and van Heuven's (2003) model of orthographic processing. Location-specific letter detectors (alphabetic array) send information to sublexical, word-centered, orthographic representations (relative position map), which in turn activate whole-word orthographic representations (O-words).

Grainger and van Heuven (2003) proposed a solution inspired by the seminal work of Mozer (1987) and the subsequent development of this approach by Whitney (2001). In the Grainger and van Heuven model of orthographic processing (Figure 2), the alphabetic array codes for the presence of a given letter at a given location relative to eye fixation along the horizontal meridian. It does not say where a given letter is relative to the other letters in the

stimulus, since each letter is processed independently of all others. Thus, processing at the level of the alphabetic array is insensitive to the orthographic regularity of letter strings. However, for the purposes of location-invariant word recognition, this location-specific map must be transformed into a “word-centered” code such that letter identity is tied to within-word position (where a word is defined as a string of letters separated by spaces) independently of retinal location (cf. Caramazza and Hillis, 1990). In order to perform this transformation, Grainger and van Heuven (2003), following Mozer (1987) and Whitney (2001), proposed a mechanism they called “open-bigram” coding. In Grainger and van Heuven’s scheme, open-bigrams code for the presence of ordered pairs of letters independently of their contiguity. Therefore, exactly the same open-bigram representation (e.g., T-A) would be activated by words containing these two letters in that order independently of how many intervening letters there are (e.g., table, train, thrash).¹ In other words, this type of representation “knows” that a given pair of letters is present in the stimulus in a given order, but does not “know” whether the two letters are next to each other or not.

As pointed out by Goswami and Ziegler (2006), however, this kind of flexible orthographic code is unlikely to provide an appropriate input for the sublexical translation of orthography to phonology. In order to compute the identity of complex graphemes such as “sh”, you need to know that H immediately follows S. Knowing that S is somewhere before H (as in an open-bigram coding scheme) would generate too many false positive coding errors or binding errors (von der Malsburg, 1999), such as identifying the presence of the complex grapheme “sh” in the word “sahara”. This observation suggests that the initial phase of orthographic processing in the Grainger and van Heuven (2003) model, the alphabetic array, should feed into two different types of location-invariant sublexical codes. This is the basis of our dual-route approach to orthographic processing, to be described in more detail below. However, contrary to alternative dual-route accounts of orthographic processing (e.g., Whitney & Cornelissen, 2005; 2008), the motivation behind the two routes in our approach is not based on the traditional distinction between direct “orthographic” and indirect “phonological” pathways in reading, although, as we show, there is a clear link with this tradition. As argued above, the principal motivation behind the two types of orthographic

¹ Although Grainger and van Heuven (2003) imposed a limit of a maximum of two intervening letters, the number of intervening letters (0, 1, or 2) had no influence on bigram activation, which is a key difference with respect to Whitney’s (2001) SERIOL model and Dehaene, Cohen, Sigman, and Vinckier’s (2005) local combination detector (LCD) model.

code in our dual-route approach arises from a consideration of the different types of constraint that operate during the course of learning to read and the development of location-invariant sublexical orthographic representations.² It is hypothesized that these constraints push the system to develop diagnostic features (letter combinations) for word identity on the one hand, and clusters of highly co-occurring letters of functional significance, on the other. These arguments are considered in more detail in the following section.

A dual-route approach to orthographic processing in skilled readers

With the focus on silent word reading, the general goal of our modeling efforts is to account for how, given the constraints on letter-in-string visibility, plus the temporal constraints imposed by reading rate (about 250 ms per word), the skilled reader optimizes uptake of information from the printed word stimulus in order to recover the appropriate semantic information necessary for text comprehension. The dual-route approach acknowledges that two different types of constraints affect processing along the two routes. Both types of constraints are driven by the frequency with which different combinations of letters occur in printed 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 highly 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. 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. This dual-route approach to orthographic processing is illustrated in Figure 3.

² Another key difference with respect to Whitney and Cornelissen's (2005, 2008) dual-route approach, is that the two types of sublexical location-invariant orthographic representation postulated in our model, are seen as two alternative means to derive location-invariance from lower-level location-specific letter representations.

Fundamentally different types of orthographic processing are performed by the two routes of our dual-route approach, since they are geared to use frequency of occurrence in diametrically opposite ways. The two routes differ notably in terms of the level of precision with which letter position information is coded. In one route, a coarse-grained orthographic code is computed in order to rapidly home in on a unique word identity and the corresponding semantic representations (the fast track to semantics). Given variations in visibility across letters in a string, the key hypothesis here is that the best way to optimize performance is to adapt processing to the constraints imposed by variations in letter visibility and variations in the amount of information carried by different letter combinations. That is, the strategy of this route is to code for combinations of the most visible letters that best constrain word identity.

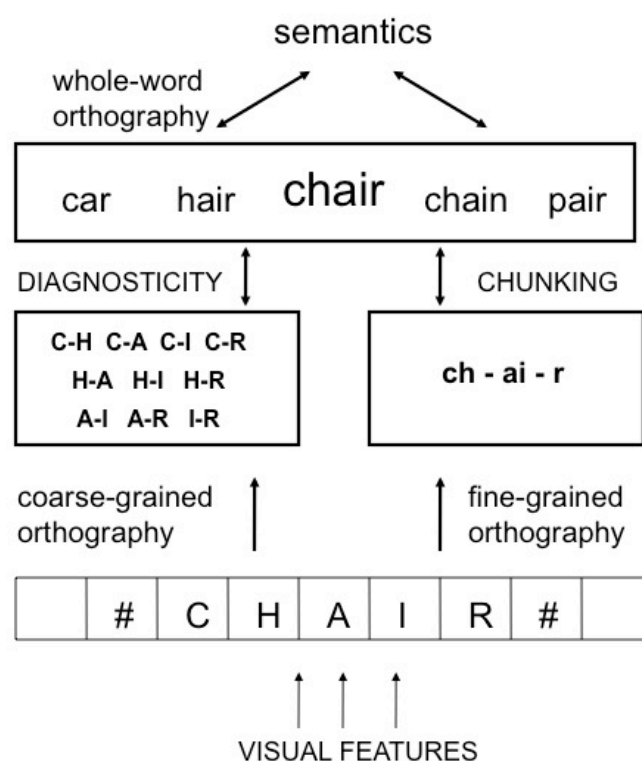


Figure 3. A dual-route approach to orthographic processing. A bank of location-specific letter detectors send activation forward to two types of sublexical location-invariant orthographic representations: 1) *coarse-grained* representations that code for the presence of informative letter combinations in the absence of precise positional information, and 2) *fine-grained* representations that code for the presence of frequently co-occurring letter combinations (multi-letter graphemes, affixes). The coarse-grained code optimizes the mapping of orthography to semantics by selecting letter combinations that are the most informative with respect to word identity (*diagnosticity*), irrespective of letter contiguity. The fine-grained code optimizes processing via the *chunking* of frequently co-occurring contiguous letter combinations.

Coding for contiguous and non-contiguous letter combinations in Grainger and van Heuven's (2003) model of orthographic processing (so-called "open-bigram" coding: Dehaene et al., 2005; Grainger & van Heuven, 2003; Grainger & Whitney 2004; Whitney, 2001; see Dandurand, Grainger, & Dunabeitia, 2011, for further discussion), provides one means of implementing this specific strategy. The key idea behind this proposal is that given variations in letter visibility across a string of letters (e.g., Stevens & Grainger, 2003), the most efficient means of obtaining a fast guess at word identity is to compute order and identity information for the most visible letters. This is not intended to be a foolproof means of knowing what word is present, but a means to provide rapid bottom-up activation of whole-word representations that can be combined with contextual constraints in order to home-in on the correct word meaning during reading comprehension. In the absence of such additional top-down constraints, processing along the slower fine-grained route will also provide disambiguating information when necessary.³

Empirical evidence in favor of this type of coarse orthographic coding has been obtained using the masked priming paradigm in the form of robust priming effects with transposed-letter primes (e.g., gadren-GARDEN: Perea and Lupker 2004; Schoonbaert and Grainger 2004), and subset and superset primes (e.g., grdn-GARDEN, gamrdsen-GARDEN: Grainger et al., 2006; Peressotti & Grainger 1999; Van Assche & Grainger 2006; Welvaert, Farioli & Grainger, 2008; see Grainger, 2008, for review). Not only does open-bigram coding provide a natural explanation for these empirical demonstrations of flexible orthographic processing, but when combined with the constraints of letter visibility and informativity it can also account for more subtle variations in orthographic priming effects as a function of the position of orthographic overlap and the precise letters involved (e.g., consonants vs. vowels – Dandurand et al., 2011). We nevertheless point out that open-bigram coding is only one possible implementation of the coarse-grained orthographic processing route of our dual-route approach. Replacing this mechanism with an alternative coding scheme (e.g., Gomez et al., 2008) would not be detrimental to our approach, as long as the new mechanism exhibits the

³ An obvious analogy can be made with Bar et al.'s (2006) model of object recognition, where a fast-guess to object identity is provided by low spatial frequency information (approximate shape), and this fast-guess mechanism is backed-up by the slower processing of high spatial frequency information (fine details) enabling disambiguation of the fast-guess. However, we do not think that low spatial frequency information can provide the same constraints on word identity as they provide for object identity, although this remains a key issue for future research (see Grainger & Dufau, 2011).

key properties of flexibility, and maximizing the use of diagnostic information given constraints in letter visibility.

On the right-hand side of Figure 3, the fine-grained orthographic code provides more precise information about the ordering of letters in the string. This fine-grained code enables the coding of multi-letter graphemes and their precise ordering in the string. These graphemes then activate the corresponding phonemes, which in turn lead to activation of the appropriate whole-word phonological representation and the corresponding semantic representations (see Perry et al., 2007, for a specific implementation of a graphemic parser). However, in the present theoretical approach, the fine-grained orthographic processing route is not limited to the case of processing grapheme representations. Here, we hypothesize that this route is more generally dedicated to precisely specifying the within-string positions of letter identities in order to facilitate the chunking of frequently co-occurring contiguous⁴ letter combinations, such as complex graphemes (e.g., TH, CH, OO) and small morphemes (i.e., affixes such as RE, ED, ER, ING). This allows orthographic information to make contact with pre-existing sublexical phonological and morphological representations. Furthermore, under the hypothesis that phonemes are a good candidate for such sublexical phonological representations, single-letter representations are required in order to make contact with the numerous phonemes that are written with a single letter. So, how might letter position information be coded in the fine-grained orthographic processing route?

One solution would be to adopt slot-coding with both beginning and end anchor points, such as proposed by Jacobs et al. (1998), a variant of which has been adopted in recent work on letter position coding in written language production (Fischer-Baum, McCloskey, & Rapp, 2010). Fischer-Baum et al. propose that letter position is coded both from the beginning to the end, *and* from the end to the beginning of the string (a scheme referred to as “both ends” letter position coding). This proposal applies the same anchor points as in the Jacobs et al. scheme, but extends coding in both directions to cover all letters, such that each letter in the string is represented twice – with its location relative to word beginning on the one hand, and its position relative to the end of the word on the other. Fischer-Baum et al. showed that this type of coding scheme provided a superior account of the patterns of spelling errors in

⁴ We note here that contiguity is not deemed to be a necessary condition for chunking, with perhaps the best example of non-contiguous chunking provided by Semitic morphology. We return to discuss this issue in the sections on morphology and learning to read.

two dysgraphic persons, compared with alternative letter position coding schemes. Again, we point out that this is one possible implementation of fine-grained orthographic coding in our dual-route approach. Replacing this specific mechanism with an alternative mechanism would not be detrimental to our approach, as long as the new mechanism exhibits the key property of precise within-word position coding that is hypothesized to facilitate the chunking of frequently co-occurring contiguous letter combinations.

Why would a human brain exposed to print adopt this dual-route approach to orthographic processing? As suggested by Grainger and Holcomb (2009), it is possible that this dual-route architecture emerges from the fact that visual word recognition is a mixture of two worlds: one whose main dimension is space - the world of visual objects; and the other whose main dimension is time - the world of spoken language. Skilled readers might therefore have learned to capitalize on this particularity, using structure in space in order to optimize the mapping of an orthographic form onto semantics, and using structure in time in order to optimize the mapping of an orthographic form onto phonology. However, even in the domain of spoken language one can draw a distinction between diagnosticity and chunking. Therefore, a more general answer to the above question would be that optimization of identification processes (for any kind of object) involves extraction of diagnostic information on the one hand, and a simplification of processing via data compression on the other.⁵ A dual-route approach follows naturally from this distinction to the extent that fundamentally different mechanisms are likely to be involved in learning these two types of code. The important consequence of this distinction for models of visual word recognition, is that we should not be looking for a single type of sublexical orthographic code, but different types of orthographic codes that are developed to perform very different functions.

Orthography and phonology

Our dual-route approach to orthographic processing can be easily integrated within the more general framework of a bi-modal interactive-activation model of visual word recognition. Figure 4 describes a multiple-route model of printed word recognition that is basically an extension of the bi-modal interactive-activation model (BIAM) that incorporates

⁵ Note that minimizing the number of diagnostic features is also a means to compress data (e.g., Mel & Fiser, 2000).

the distinction drawn between two types of sublexical orthographic code. Like the BIAM, our multiple-route model of word recognition has many similarities with dual-route models of reading aloud (Coltheart et al., 2001) and particularly the CDP+ model (Perry et al., 2007, 2010). The model proposed in Figure 4 provides a more explicit description of the processing involved in getting from print to meaning via orthographic representations alone, and draws a key distinction between location-specific and location-invariant (word-centered) orthographic codes.

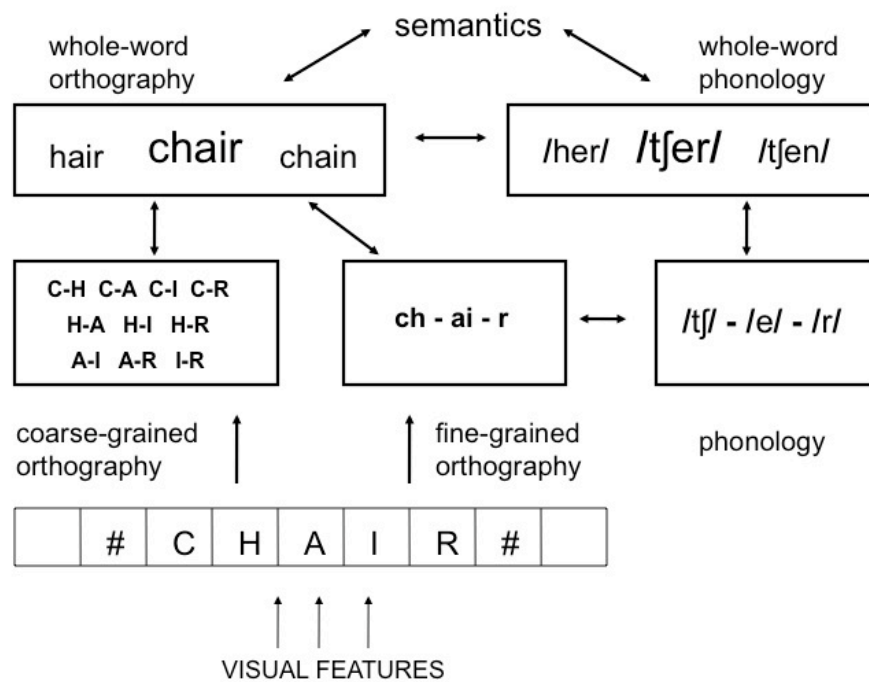


Figure 4. A multiple-route model of word comprehension in silent reading that integrates the principle of two types of location-invariant sublexical orthographic code within a generic bi-modal interactive-activation model (BIAM). The fine-grained orthographic code provides the level of precision in position coding that is necessary to interface with sublexical phonological representations. Note that the distinction between “direct” orthographic and indirect “phonological” pathways in traditional dual-route models is extended here with the distinction between the two orthographic pathways.

According to the model depicted in Figure 4, the route from print to meaning via sublexical phonological representations, involves fine-grained orthographic processing. That is, the system needs to know precisely the ordering of the different letter identities in the stimulus word (Goswami & Ziegler, 2006). This is particularly important for extracting contiguous letter combinations that form multi-letter grapheme representations, such as the “ch” and “ai” in the word “chair”. It is also thought to be the mechanism responsible for extracting other types of highly recurring contiguous letter combinations, such as affixes.

Pseudo-homophone effects represent one key empirical signature of fine-grained orthographic processing, since it is generally agreed that the processing of such stimuli involves some form of sublexical conversion of print-to-sound. Pseudo-homophones are nonwords that can be pronounced like a real word, such as the letter string “brane” pronounced as the word “brain”. These stimuli are harder to reject as nonwords in a lexical decision task (e.g., Goswami, Ziegler, Dalton, & Schneider, 2001; Ziegler, Jacobs, & Krueppel, 2001), generate more semantic categorization errors (e.g., Van Orden, 1987), and are more effective primes compared with carefully matched orthographic controls (e.g., Ferrand & Grainger, 1994; Frost, Ahissar, Gotesman, & Tayeb, 2003; Lukatela & Turvey, 1994; Perfetti & Bell, 1991; Ziegler, Ferrand, Jacobs, Rey, & Grainger, 2000; see Rastle & Brysbaert, 2006, for review).

Our multiple-route model makes one key prediction with respect to effects of pseudo-homophone primes. These effects should be eradicated by a transposed-letter manipulation, since precise letter order information is required along the fine-grained processing route that generates a sublexical phonological code. That is, according to our approach, precise letter order information is required in order to generate a pseudo-homophone priming effect because phoneme representations are activated via the fine-grained orthographic code. Now, one key empirical phenomenon provided the principal motivation for the theoretical shift from overly precise letter position coding schemes, such as the slot-coding scheme of the Interactive-Activation model (McClelland & Rumelhart, 1981), to more flexible coding schemes such as open-bigram coding. That was the effects of letter transpositions observed in masked priming experiments (e.g., Perea & Lupker, 2004; Schoonbaert & Grainger, 2004) or in nonword decision latencies or errors in unprimed lexical decision (e.g., Andrews, 1996; Chambers, 1979; O’Connor & Forster, 1981; Perea, Rosa, & Gomez, 2005). Rigid slot-coding schemes cannot account for transposed-letter effects (see Grainger, 2008, for review). Therefore, given the hypothesis that only fine-grained orthographic coding, and not coarse-grained coding, can generate an accurate sublexical phonological representation of a string of letters, and under the assumption that pseudo-homophone priming effects are subtended by sublexical phonology (see Rastle & Brysbaert, 2006, for review), then we predict that introducing a letter transposition in a pseudo-homophone stimulus should eradicate its ability to prime the corresponding baseword. Acha and Perea (2010) compared priming effects for transposed-letter words (e.g., *caniso-CASINO* vs. *caviro-CASINO*) and transposed-letter

pseudo-homophones (e.g., *kaniso-CASINO* vs. *kaviro-CASINO*). They found the standard TL priming effect for TL word primes, but no priming effect for TL pseudo-homophone primes, as predicted by our dual-route approach.

Another line of evidence in favor of this dual-route approach comes from experiments manipulating the orthographic regularity and pronounceability of nonword stimuli created by a letter transposition (Frankish & Turner, 2007; Frankish & Barnes, 2008). Transposing two letters of the baseword “storm”, for example, can generate either an orthotactically and phonotactically regular nonword such as “strom” or an orthotactically and phonotactically irregular nonword such as “sotrm”. Frankish and colleagues found that the illegal nonwords were more readily misperceived as the corresponding baseword in a perceptual identification task, took longer to reject as a nonword in the lexical decision task (Frankish & Turner, 2007), and were more effective primes for the corresponding baseword targets in a masked priming experiment (Frankish & Barnes, 2008). We agree with Frankish and colleagues, that the reduced TL effects seen with legal nonwords most likely reflect the more efficient computation of a phonological code with this kind of nonword compared with the illegal nonwords. This phonological code, generated via the fine-grained processing route in our model, would interfere with the output of the coarse-grained processing route, therefore reducing the evidence in favor of the presence of the baseword (evidence provided by the coarse-grained processing route).

Furthermore, in line with the architecture of the BIAM and its extension in the form of the multiple-route model shown in Figure 4, we know that phonological influences on visual word recognition are fast acting (Braun et al., 2009), but nevertheless require more processing than that required to obtain purely orthographic effects. This differential time-course of effects has been revealed on several occasions in studies of masked orthographic and phonological priming (Ferrand & Grainger, 1994; Ziegler et al., 2000). It has also been revealed in research combining masked priming and the recording of even-related potentials (ERPs: Grainger, Kiyonaga, & Holcomb, 2006). Most important is that the latter research compared priming from transposed-letter primes (e.g., *barin-BRAIN* vs. *bolin-BRAIN*) with priming from pseudo-homophone primes (e.g., *brane-BRAIN* vs. *brans-BRAIN*). Effects of transposed-letter primes were seen in the ERP waveforms about 50 ms before the emergence of pseudo-homophone priming effects.

Finally, future work will need to explore exactly how sublexical orthographic chunking, hypothesized to operate along the fine-grained orthographic processing route, could be coupled with a graphemic parser for grapheme-to-phoneme conversion, such as implemented in CDP+ (Perry et al., 2007) and the BIAM (Diependaele et al., 2010). One possibility, within a “both ends” coding scheme for fine-grained orthographic processing described above, is to add a sequential beginning-to-end graphemic parser that operates on the letter representations coded with respect to word beginnings. This phonologically constrained parser would interact with the orthographically driven chunking mechanism, reinforcing representations that are both orthographically and phonologically constrained, such as frequently occurring complex graphemes. Here we imply that purely orthographic constraints need to be supplemented with phonological constraints in order to optimize the sublexical conversion of print-to-sound. These phonological constraints could operate at the level of individual phonemes and phonological syllables, as in CDP++ (Perry et al., 2010).

Orthography and morphology

A large number of the words we read every day are morphologically complex (approximately 75% in French and 85% in English). These include prefixed and suffixed derivations (e.g., *rework*, *worker*), compounds (e.g., *work-place*), and inflected forms (e.g., *works*, *working*, *workers*). There is growing evidence that part of the process of reading morphologically complex words involves the sublexical segmentation of the word into its constituent morphemes (e.g., *work* + *er*). A large number of masked priming studies have shown that derived suffixed primes facilitate the recognition of stem targets (*worker-work*) relative to unrelated primes (e.g., Grainger, Colé, & Segui, 1991). Critically, morphological priming is larger than both form priming (*scandal-scan*) and semantic priming (*giraffe-horse*) (see Diependaele, Grainger & Sandra, 2010 for review). However, perhaps the key evidence in favor of sublexical morphological segmentation has been obtained from experiments using semantically opaque complex primes (*department-depart*) and pseudo-complex primes (*corner-corn*). Both these prime types show priming that is similar to that of semantically transparent complex primes (*worker-work*; see Rastle & Davis, 2008 for review), and significantly larger facilitation than primes consisting of the stem plus a non-suffix ending (*scandal-scan*). Furthermore, similar effects have also been found when comparing true derived word primes (*worker-work*) and complex pseudoword primes (*cornity-corn*), while

again no priming is found when the prime is formed by the stem plus a non-suffix ending (*cornal-corn*) (Longtin & Meunier, 2005, but see Morris, Porter, Grainger & Holcomb, 2010, for possible limits of such a finding).

These results all point to some form of sublexical morpho-orthographic processing that, when presented with a fully decomposable stimulus, segments the stem and affix thereby allowing activation of an orthographic representation of the stem. It is this boost in activation of the representation of the stem that generates facilitation during processing of the stem as the following target word. On the basis of these results, it is commonly agreed that morphology influences visual word recognition through fast and automatic morpho-orthographic segmentation (see Rastle & Davis, 2008, for a review of the findings). Here we hypothesize that such morpho-orthographic segmentation occurs along the fine-grained orthographic processing route of our model. More precisely, we propose that affix detection is the key mechanism underlying morpho-orthographic segmentation (see Andrews & Davis, 1999, for a concrete proposal). Therefore, since the detection of affixes, just like complex graphemes, requires precise letter position coding (e.g., to distinguish between the real suffix “ion” and the non-suffix ending “oin”), it can only be performed via fine-grained orthographic processing. Just as is the case with complex graphemes, coarse-grained coding would give rise to too many false affix detections (binding errors), such as detecting the suffix “er” in “their”.

Morpho-orthographic processing is, however, only part of the story of how morphology can influence reading. According to the account of morphological processing proposed by Diependaele et al. (2005; 2009; see also Morris et al., 2010), such morphological influences not only reflect morpho-orthographic processing, but also morpho-semantic processing. As shown in Figure 5, morpho-semantic representations are supralexical representations that impose a morphological organization on whole-word form representations (Grainger et al., 1991; Giraudo & Grainger, 2001). As such, they represent knowledge that certain words overlap in both form and meaning. Evidence in favor of this approach, and against purely morpho-orthographic accounts (Rastle & Davis, 2008), comes from masked priming studies showing an advantage for semantically transparent derived primes relative to semantically opaque and/or pseudo-derived primes (Diependaele et al. 2005; 2009; Morris, Frank, Grainger & Holcomb, 2007). Although these reports of significant effects of semantic transparency appear isolated relative to the numerous reports of non-

significant differences, recent meta-analyses by Feldman, O'Connor, and Moscoso del Prado Martin (2009) and Davis and Rastle (2010) show that despite the non-significance in the majority of individual studies, the effect of semantically transparent primes is indeed significantly larger than that of opaque primes when analyzed across studies.

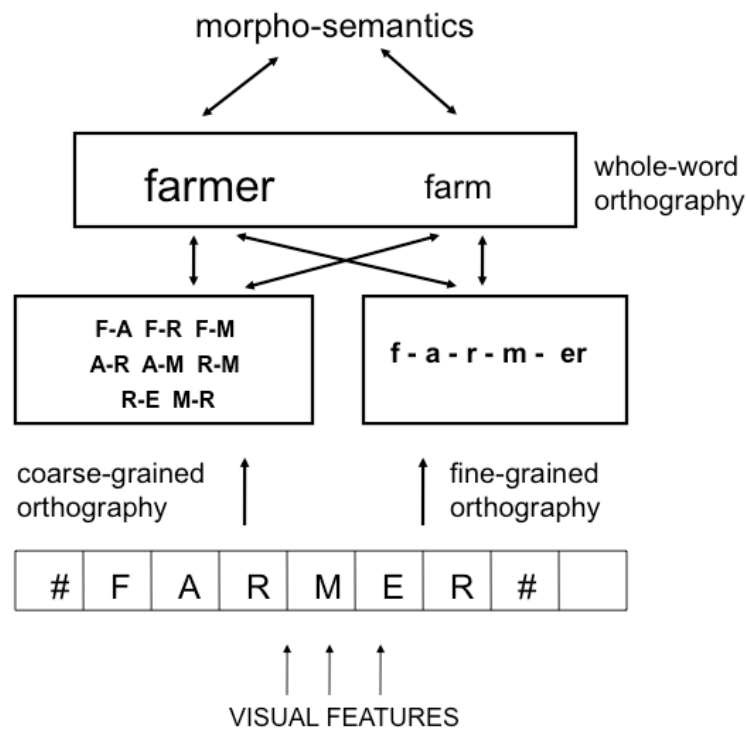


Figure 5. Morphological processing and the dual-route approach to orthographic processing. Fine-grained orthographic processing enables sublexical morpho-orthographic segmentation via the detection of affixes such as the suffix “er” in the stimulus “farmer”. Activation in these representations is fed-forward to whole-word orthographic representations, increasing the activation level of all compatible units (e.g., “farmer”, “farm”). Coarse-grained orthography activates compatible whole-word orthographic representations. Morpho-semantic representations provide bi-directional connectivity between whole-word representations belonging to the same morphological family.

Figure 5 shows how the account of morphological processing developed by Diependaele et al., fits within our dual-route approach to orthographic processing. Morpho-orthographic segmentation is achieved via fine-grained orthographic processing, which then feeds activation forward to compatible whole-word orthographic representations. At the same time, coarse-grained orthographic processing also feeds forward activation to whole-word orthographic representations, but independently of the morphological structure of the stimulus. Activation in whole-word form representations (both orthographic and

phonological, not shown in Figure 5) is further constrained by top-down feedback from morpho-semantic representations, such that activation of “farmer” leads to increased activation of the whole-word form “farm” via the shared morpho-semantic representation of “something to do with farms”. In sum, the key idea behind our dual-route proposal when applied to morphological processing, is that the mechanism that learns that the presence of the letter F somewhere before the letter R is a good indication that the whole-word being processed is “farmer” or “farm”, is not the same mechanism that learns that there is an E just before a final R, and that this combination of letters has a specific linguistic function.

One prediction of this approach to morphological processing within the multiple-route framework, is that effects that are driven by morpho-orthographic processing should be selectively impaired by manipulations that are thought to principally affect fine-grained orthographic processing. This prediction has been the object of recent experimentation where we compared the effects of letter transpositions on priming from semantically transparent derivations and pseudo-derivations (Diependaele, Morris, & Grainger, 2011). In this study, the standard comparison of morphologically transparent primes (e.g., *farmer-farm*) with pseudo-morphologically related primes (e.g., *corner-corn*), was augmented with a TL manipulation involving the two letters across the morpheme (pseudo-morpheme) boundaries (e.g., *faremr-farm*; *corenr-corn*). These priming effects were measured relative to standard double substitution control primes (e.g., *farivr-farm*; *corivr-corn*).

As predicted by our theoretical approach, we found significant priming from intact derived primes (e.g., *farmer-farm*) and TL derived primes (e.g., *faremr-farm*), as well significant priming from intact pseudo-derived primes (e.g., *corner-corn*), but most important, no priming from TL pseudo-derivations (e.g., *corenr-corn*). According to our dual-route model, letter transpositions selectively interfere with fine-grained orthographic processing, and therefore selectively perturb sublexical morpho-orthographic segmentation. Since this is hypothesized to be the only source of priming for pseudo-derived relations (e.g., *corner-corn*), a TL manipulation eliminates priming in this condition. On the other hand, true morphological relations (e.g., *farmer-farm*) still benefit from morpho-semantic facilitation obtained via coarse-grained coding, such that the prime “faremr” strongly activates the whole-word orthographic representation “farmer” which connects with the whole-word orthographic representation for the word “farm” via shared morpho-semantic representations.

Finally, it is interesting to note the fact that standard TL priming effects are not found in Semitic languages, at least for the true Semitic words of these languages (Perea, Abu Mallouh, & Carreiras, 2010; Velan & Frost, 2009; 2011). Within the framework of our dual-route approach, this would imply that Semitic words are processed via fine-grained orthographic processing, which would enable the accurate detection of tri-consonantal roots. Letter transpositions would therefore disrupt the processing of these root morphemes, causing as much damage as letter substitutions because of the fine-grained nature of the coding. Furthermore, Velan and Frost (2011) have recently demonstrated that morphologically simple Hebrew words of non-Semitic origin show the standard effects of letter transpositions seen with words in languages like English and French. Within our dual-route approach, these non-Semitic Hebrew words would be processed just like English and French words, with robustness to TL manipulations arising from involvement of the coarse-grained processing route. Semitic root-derived words, on the other hand, would be predominantly processed via the fine-grained route. The key question here is therefore why these Semitic root-derived words would not also involve coarse-grained processing, given that we do not expect the two processing routes to be subject to any kind of on-line control mechanism that could disable one or the other processing route. Here we tentatively suggest that it is the efficiency in the mapping of form to meaning via root morphemes and verbal patterns in the fine-grained processing route (i.e., morphological decomposition) that prevents development of coarse-grained orthographic processing (i.e., whole-word access) for Semitic root-derived words during the course of learning to read.⁶

In terms of Frost's (2009) hypothesis that reading words in Semitic languages involves qualitatively different processes than those used to read words in Indo-European languages, within our dual-route framework this translates into prioritizing fine-grained processing as opposed to coarse-grained processing. As noted above, such prioritizing is not thought to occur on-line, but arises for each individual word as a function of the relative efficiency of the process of mapping orthographic information onto semantics along each of the two routes. This is therefore more a quantitative than a qualitative distinction, and in line with this reasoning, Velan and Frost (2011) reported the results of an interesting intermediate case of

⁶ This is just a re-statement of the issue of how whole-word processing might trade-off with morphological decomposition as a function of the morphological structure of a language (e.g., agglutinative or not), or for certain categories of words within a language (e.g., Semitic vs. non-Semitic Hebrew words). See Velan and Frost (2011) for arguments for why priority might be given to morphological decomposition when reading Semitic words.

Hebrew words that have an existing verbal pattern combined with a non-productive or pseudo-root. These words were found to produce a pattern of priming effects lying in between the root-derived words and the morphologically simple non-Semitic words. Within our dual-route approach, the likely reduced efficiency in morpho-orthographic processing of such words compared with root-derived words, would allow the development of coarse-grained processing, without actually reaching the same level of efficiency as achieved when processing non-Semitic words in the coarse-grained route.

A multiple-route account of learning to read words

In the final section of this work, we examine the implications of our dual-route approach to orthographic processing with respect to the process of learning to read words. This is an essential extension of the approach, given that the two types of orthographic coding postulated in our model are thought to emerge as the result of specific constraints operating during reading acquisition. It is therefore important to begin to understand how and when such constraints might come into play, and what factors might modulate their contribution to orthographic learning.

The main task of the beginning reader of a language that uses an alphabetic script is to associate letter identities with sounds in order to make contact with whole-word phonological representations of known words (phonological recoding). Initially, this will involve a serial letter-by-letter reading strategy, since the mechanism for parallel letter identification is not yet established. By shifts of the eyes and shifts of attention, the beginning reader identifies the different letters of the word one at a time, and learns what sounds they correspond to. This mechanism simply capitalizes on the two key sources of information that the beginning reader has available – knowledge of the alphabet and spoken vocabulary.

Apart from the initial acquisition of a small sight vocabulary (involving the most frequently occurring words), we agree with Share (1995) and others that phonological recoding is the essential first step in reading acquisition (e.g., Ehri, 1992). Inspired by the self-teaching hypothesis of Share (1995), we propose that it is during this relatively slow and painstaking process of phonological recoding that exposure to printed words enables the setting-up of the specialized system for parallel orthographic processing that is hypothesized

in our model of orthographic processing. According to Share, each successful decoding that is achieved via the laborious serial procedure provides the beginning reader with an opportunity to set up connections between the printed word and the decoded meaning. Within the present theoretical framework, this specifically involves the development of parallel, independent letter processing. It is the development of parallel letter processing, thought to involve some form of location-specific letter code, that then leads to the development of the two types of location-invariant, sublexical, orthographic codes that form the basis of skilled silent reading according to our dual-route approach. This developmental progression is illustrated in Figure 6.

The nature of the two types of sublexical location-invariant orthographic codes that are hypothesized in our approach, is thought to be determined by the constraints imposed by the general goal of optimizing the mapping of letters onto meaning while learning to read. On the one hand, these constraints involve optimization of the mapping of letters onto whole-word orthographic representations, and from there, onto the associated semantic representations. This is the coarse-grained processing route that provides direct access to semantics via orthographic information alone. Here, optimization is thought to involve the development of letter combination detectors that best constrain word identity (the diagnosticity constraint), in the same way that parts of objects act as clues to object identity in certain theoretical approaches to visual object recognition (e.g., Ullman, Vidal-Naquet, & Sali, 2002). For printed words, we hypothesize that this involves selecting letter combinations that maximize the visibility of the constituent letters and maximize the amount of information they carry with respect to word identity (Dandurand et al., 2011; Grainger & Dufau, 2011).

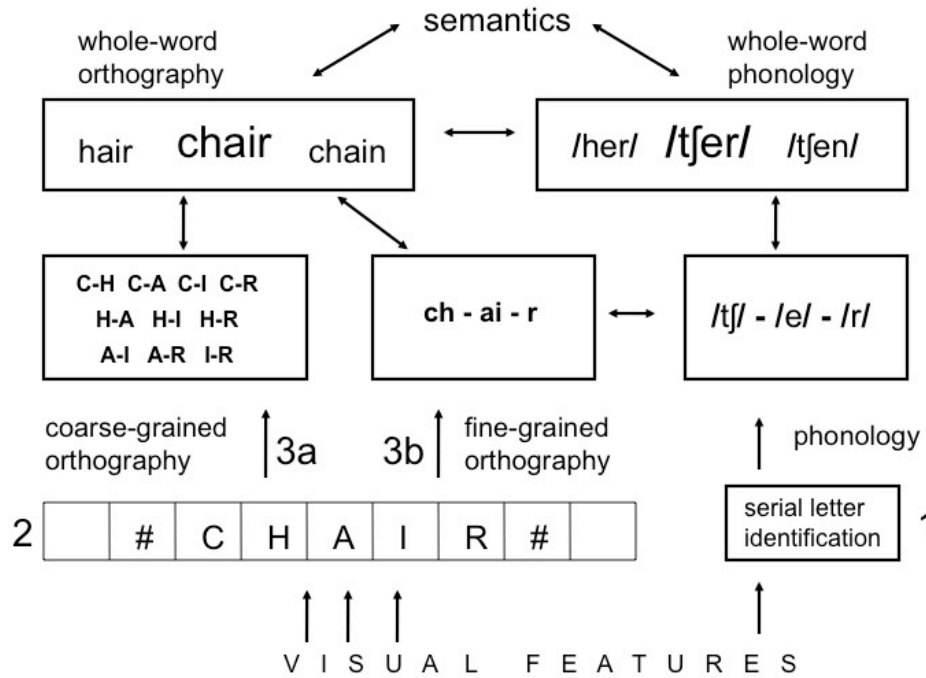


Figure 6. The major steps involved in learning to read words described within the framework of a multiple-route model of silent reading. 1) Orthographic input is initially processed letter-by-letter, and the corresponding sounds are derived from letters and letter combinations (phonological recoding). 2) Development of parallel independent letter processing in the form of a bank of location-specific letter detectors. 3) Development of two types of location-invariant sublexical representation: a) coarse-grained representations for fast access to semantics from orthography, and b) fine-grained representations involving a modification of the process used to translate print-to-sound (grapheme representations) and the development of morpho-orthographic representations (affixes).

Constraints during reading acquisition also operate to optimize the mapping of letters onto meaning by connecting letters with the pathway that is already used to map speech onto meaning during spoken language comprehension. This is the fine-grained processing route that provides access to semantics via phonological and morphological representations. Here, optimization is hypothesized to involve the development of orthographic representations that facilitate the mapping of letter representations onto pre-existing sublexical representations involved in spoken word comprehension.⁷ Given the nature of these pre-existing representations, this optimization is thought to involve detection of frequently co-occurring letter combinations (the chunking constraint). Here, constraints operate not to maximize information with respect to word identity, but to facilitate the transformation of the orthographic code into a different type of linguistic code that has already been optimized for

⁷ We acknowledge here that such pre-existing representations might in turn be modified by the process of reading acquisition.

mapping onto meaning. Frequently co-occurring groups of letters often represent the orthographic equivalent of phonemes and morphemes.

As noted above, the process of phonological recoding is thought to initially involve a letter-by-letter reading strategy, where order information is provided by the sequence of encoding events. The development of parallel letter identification is therefore hypothesized to cause a shift from a strictly sequential letter encoding (that outputs an ordered set of phonemes) to a more parallel mapping of letters onto higher level orthographic representations such as graphemes and affixes, that retains the same level of precision as the strictly sequential mechanism (see Alario, De Cara, & Ziegler, 2007). Whether or not the mapping of graphemes onto phonemes also becomes more parallel is another issue still open to debate (see Carreiras et al., 2005, for evidence that this process is sequential at the level of syllable representations). Here it is hypothesized that a precise word-centered letter-position coding scheme enables the chunking of frequently co-occurring contiguous letter combinations, such as multi-letter graphemes and affixes, independently of eye fixation location in the word. The development of a parallel version of this fine-grained orthographic coding is particularly important for the extraction of suffixes or rhymes, for which the positions of the component letters are defined with respect to word endings (e.g., Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995).

Unsupervised learning algorithms, such as implemented in Self-Organizing Maps (e.g., Dufau, Lété, Touzet, Glotin, Ziegler, & Grainger, 2010; Kohonen, 1982) or Adaptive Resonance Theory (e.g., Glotin et al., 2010; Grossberg, 1987), provide specific mechanisms for performing orthographic chunking on the basis of a parallel letter input. In these approaches, frequency of co-occurrence determines the formation of higher-order categories. Finally, we follow Zorzi, Houghton, & Butterworth (1998) and Perry et al. (2007) in proposing that the associations established between sublexical orthographic and phonological representations (such as graphemes and phonemes) arise via supervised learning. However, this supervision can be both externally driven (teacher) and internally driven, by using the output of the letter-by-letter strategy to modify the mapping of the parallel orthographic code onto phonology.

Furthermore, although learning of coarse-grained and fine-grained representations may well both largely involve implicit, unsupervised learning algorithms, attention might play

a different role in these two cases. Recent research suggests that attention might be a critical factor in learning dependencies among elements (e.g., Pacton & Perruchet, 2008; see Le Pelley, 2010, for a review of the evidence). According to Pacton and Perruchet (2008), the reason why non-adjacent contingencies are typically harder to learn than adjacent contingencies is because it is easier to focus attention on contiguous events than non-contiguous events. In line with this reasoning, examples of successful learning of non-adjacent dependencies involve situations where it is arguably easier to attend to the non-contiguous elements (e.g., Gomez, 2002; Pena, Bonatti, Nespor, & Mehler, 2002). In the specific case of learning to read words, two factors could help focus attention on contiguous dependencies between letters. One would be an initial phase of supervised learning, whereby the beginning reader is told that a given complex grapheme corresponds to a particular phoneme. The other is the presence of pre-existing phonological and/or morphological representations that correspond to a contiguous sequence of letters. So far, so good for contiguous letter combinations, but we still need to know how chunks of non-contiguous letters could be formed, such as the root morphemes of Semitic languages. One possibility here is that the consonant-vowel distinction (roots are formed exclusively of consonants) helps focus attention on these non-contiguous elements during the learning of spoken language, and that it is this learned association that enables attention to be brought to focus on the orthographic equivalent during reading acquisition.

Our general account of learning to read words, illustrated in Figure 6, predicts that the initial dominance of serial phonological recoding should rapidly be replaced by parallel orthographic processing. According to this account, the development of parallel orthographic processing will enable 1) faster access to semantic representations via the development of a coarse-grained orthographic code, 2) greater efficiency in the sublexical translation of orthography to phonology via the development of a fine-grained orthographic code, and 3) the emergence of both morpho-semantic and morpho-orthographic representations via the combination of coarse-grained and fine-grained orthographic processing. Among the empirical consequences of the development of parallel orthographic processing are: 1) a reduction in the effects of word length (e.g., Acha & Perea, 2008; Aghababian & Nazir, 2000; Bijeljac-Babic, Millogo, Farioli, & Grainger, 2004), 2) a reduction in phonological effects that arise principally via phonological recoding (e.g., Grainger et al., 2011), 3) an increased sensitivity to orthographic priming (e.g., Acha & Perea, 2008; Castles, Davis, Cavalot, & Forster, 2007; Grainger et al., 2011), and 4) an increased sensitivity to morphological

structure (e.g., Colé, Bouton, Leuwers, Casalis, & Sprenger-Charolles, 2010). Future research should provide more fine-grained analyses of these different developmental patterns, as well as specific tests of the implications of our dual-route model of orthographic processing for reading development.

Dual-routes for reading in the brain?

It is tempting to link our functional dual-route approach to orthographic processing with the oft-made distinction between ventral and dorsal neuro-anatomical pathways for reading. This proposition builds on an analogy with the well-established distinction between ventral (what) and dorsal (where) pathways for visual object processing. Several authors have mapped this classic ventral-dorsal pathway distinction onto processes involved in reading words, but this has been done in various ways. Here we briefly summarize prior accounts of this mapping, and discuss how they could be applied in order to reveal the neural underpinnings of the component processes of our dual-route model.

One approach, pitched within the framework of standard dual-route theory (Coltheart et al., 2001), is to associate the ventral pathway with the direct mapping of orthography onto semantics, and the dorsal pathway with the mapping of orthography onto phonology (e.g., Borowsky, Cummine, Owen, Friesen, Shih, & Sarty, 2006; Pugh et al., 2000; Whitney & Cornelissen, 2005). The first sentence of the abstract in Borowsky et al. (2006) provides a nice summary of this approach and attests to its popularity: “Most current models of the neurophysiology of basic reading processes agree on a system involving two cortical streams: a ventral Stream (occipital-temporal) used when accessing familiar words encoded in lexical memory, and a dorsal Stream (occipital-parietal-frontal) used when phonetically decoding words (i.e., mapping sublexical spelling onto sounds).” This account fits with the evidence that brain structures located on the dorsal pathway, such as the supramarginal gyrus and the angular gyrus, are thought to be involved in the mapping of orthography to phonology (e.g., Booth, Burman, Meyer, Gitelman, Parrish, & Mesulam, 2003; Wilson, Tregellas, Slason, Pasko, & Rojas, 2011), whereas orthographic processing would be performed by brain structures in ventral occipital-temporal (VOT) regions (e.g., Cohen, Lehericy, Chochon, Lemer, Rivaud, & Dehaene, 2002).

According to certain authors, however (e.g., Cohen & Dehaene, 2009; Rosazza, Cai, Minati, Paulignan, & Nazir, 2009), the dorsal pathway would only play a limited role in normal reading, mostly for unfamiliar words or familiar words presented in an unfamiliar format. Cohen, Dehaene, Vinckier, Jobert, and Montavont (2008) associated the dorsal route with “abnormal” attention-driven, letter-by-letter reading (see Cohen & Dehaene, 2009, for a review of the evidence). According to Cohen et al., all “expert” processes involved in reading printed words are performed by neural structures in the left VOT, and more specifically by a region of the left fusiform gyrus referred to as the “visual word form area” (VWFA). Processing by this brain region would then connect with semantic and phonological representations in middle and superior temporal gyri and inferior frontal cortex. The dorsal route in Cohen et al.’s account involves brain structures in intraparietal sulcus thought to be responsible for the type of attentional control necessary for letter-by-letter reading. Such reading strategies can be induced in expert readers by changing the visual format of words (Cohen et al., 2008; Rosazza et al., 2009).

Although Cohen and Dehaene (2009) acknowledge the possible role of sublexical phonology in skilled visual word recognition, the phonological reading route in their schema of the reading system is more about connections between whole-word orthography and phonology, with neurons in anterior fusiform gyrus connecting with neurons in supramarginal gyrus, superior temporal gyrus, and pars opercularis of the inferior frontal gyrus. On the other hand, the estimated timing of sublexical orthographic and phonological influences compared with lexical influences during visual word recognition (Braun, Hutzler, Ziegler, Dambacher, & Jacobs, 2009; Grainger et al., 2006; see Grainger & Holcomb, 2009, for a review of the evidence from ERP studies), would suggest that phonological effects are coming into play earlier than in Cohen and Dehaene’s proposal (see Wheat, Cornelissen, Frost, & Hansen, 2010, for evidence for very early effects of phonology using MEG). This could be achieved by having, as suggested by Cohen and Dehaene (2009), multiple outlets from the ventral stream, such that neurons coding for orthographic structures of different grain sizes could make contact with neurons coding for phonological representations of different grain sizes (e.g., Ziegler & Goswami, 2005).

Within the framework of our dual-route approach to orthographic processing, we would argue that both coarse-grained and fine-grained orthographic processing is performed by neural structures in the ventral occipital-temporal junction, and specifically in the left

fusiform gyrus (the VWFA, Cohen et al., 2002). The evidence at present suggests that orthographic processing proceeds in a posterior-anterior sweep through left fusiform gyrus, with an increase in location-invariance (Dehaene, Jobert, Naccache, Ciuciu, Poline, & Le Bihan, 2004) accompanied by increased sensitivity to orthographic structure (Vinckier, Dehaene, Jobert, Dubus, Sigman, & Cohen, 2007). The increased sensitivity to orthographic structure could reflect both an increase in the size of coarse-grained orthographic representations (e.g., open n-grams, with increasing n), and the increase in size of fine-grained orthographic representations (e.g., graphemes and morphemes of increasing length). With respect to the role of fusiform gyrus in fine-grained orthographic processing as defined in our account, it is important to note that there is evidence suggesting that sublexical morpho-orthographic processing is performed by neural structures in this region (Bick, Goelman, & Frost, in press; Gold & Rastle, 2007; Solomyak & Marantz, 2009).

As argued by several authors, this ventral orthographic processing pathway would be one component of a triangular reading network involving ventral, dorsal, and frontal regions of the left hemisphere (e.g., Mechelli et al., 2005; Frost et al., 2007; Jobert, Crivello, & Tzourio-Mazoyer, 2003; Sandak, Mencl, Frost, & Pugh, 2004). According to Cohen and Dehaene (2009), processing in fusiform gyrus would bifurcate out onto more anterior neural structures in middle temporal gyrus, and inferior frontal gyrus, as well as onto more dorsal neural structures such as supramarginal gyrus and superior temporal gyrus. Given that phonological processing is associated with both projections from fusiform to frontal regions (pars opercularis) and projections from fusiform to dorsal regions, the utility of a ventral – dorsal distinction in describing these projections is called into question. As argued by Cohen and Dehaene (2009), the ventral – dorsal distinction might be more useful when contrasting the type of serial, letter-by-letter reading processes thought to be under attentional control in beginning readers (dorsal pathway), and the processes involved in expert reading (ventral stream plus projections to frontal and dorsal regions). In line with this proposition, and as argued in our account of reading development (see above), the mechanisms involved in sublexical conversion of print to sound in skilled reading are thought not to be the same as the very early serial mechanisms used by beginning readers.

Conclusions

We have described a dual-route approach to orthographic processing that posits the existence of two fundamentally different types of sublexical, word-centered, orthographic representations. We have shown how this distinction is easily integrated within a generic bi-modal interactive-activation model (BIAM) of word recognition, and we have discussed the implications of this integration for accounting for phonological and morphological influences on visual word recognition in skilled readers. Then we described our dual-route approach from a developmental perspective, in the form of a multiple-route model of learning to read words. In this way, phenomena that have typically been examined independently of each other, now find their place within a comprehensive account of the process of visual word recognition and its development during reading acquisition. Finally, we provided a tentative link between the component processes of our model with underlying neural structures. Most important, however, is that the overarching theoretical framework generates testable predictions, which have been the object of recent behavioral research, with promising results so far.

REFERENCES

- Acha, J. & Perea, M. (2010). Does *kaniso* activate *CASINO*? Input coding schemes and phonology in visual-word recognition. *Experimental Psychology*, in press.
- Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers, *British Journal of Psychology*, 99, 245–264.
- Aghababian, V., & Nazir, T. (2000). Developing normal reading skills : Aspects of the visual processes underlying word recognition. *Journal of Experimental Child Psychology*, 76, 123-150.
- Alario, F. X., De Cara, B., & Ziegler, J. C. (2007). Automatic activation of phonology in silent reading is parallel: Evidence from beginning and skilled readers. *J Exp Child Psychol*, 97, 205-219.
- Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed letter confusability. *Journal of Memory and Language*, 35, 775-800.
- Andrews, S., & Davis, C. J. (1999). Interactive activation accounts of morphological decomposition: Finding the trap in Mousetrap? *Brain and Language*, 68, 355-361.
- Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmidt, A.M., Dale, A.M., Hamalainen, M.S., Marinkovic, K., Schacter, D.L., Rosen, B.R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences*, 103, 449–454.
- Bick, A.S., Goelman, G., & Frost, R. (2011). Hebrew brain vs. English brain: Language modulates the way it is processed. *Journal of Cognitive Neuroscience*, in press.
- Bijeljac-Babic, R., Millogo, V., Farioli, F., & Grainger, J. (2004). A developmental investigation of word length effects in reading using a new on-line word identification paradigm. *Reading and Writing*, 17, 411-431.
- Bonatti, L. L., Pena, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, 16, 451–459.
- Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2003). Relation between brain activation and lexical performance. *Human Brain Mapping*, 19, 155-169.

- Borowsky, R., Cummine, J., Owen, W.J., Friesen, C.K., Shih, & Sarty, G.E. (2006). FMRI of ventral and dorsal processing streams in basic reading processes: Insular sensitivity to phonology. *Brain Topography*, 18, 233-239.
- Braun, M., Hutzler, F., Ziegler, J. C., Dambacher, M., & Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping*, 30, 1977-1989.
- Caramazza, A., & Hillis, A. E. (1990). Spatial representation of words in the brain implied by studies of a unilateral neglect patient. *Nature*, 346, 267-269.
- Carreiras, M., Ferrand, L., Grainger, J., & Perea, M. (2005). Sequential effects of masked phonological priming. *Psychological Science*, 16, 585-589.
- Castles, A., Davis, C., Cavalot, P., & Forster, K. I. (2007). Tracking the acquisition of orthographic skills in developing readers: masked form priming and transposed-letter effects. *Journal of Experimental Child Psychology*, 97, 165-182.
- Chambers, S.M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Behavior*, 18, 225-241.
- Cohen, L. & Dehaene, S. (2009). Ventral and dorsal contributions to word reading. In M.S. Gazzaniga (Ed.), *Cognitive Neuroscience*, 4th edition. Cambridge, MA: MIT Press.
- Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *Neuroimage*, 40, 353-366.
- Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, 125, 1054-1069.
- Colé, P., Bouton, S., Leuwers, C., Casalis, S., & Sprenger-Charolles, L. (2010). Stem and derivational suffix processing during reading by French second and third graders. *Applied Psycholinguistics*, in press.
- Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud - Dual-route and parallel-distributed-processing approaches. *Psychological Review*, 100, 589-608.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J.C. (2001). DRC: A dual-route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204-256.
- Dandurand, F., Grainger, J., & Dunabeitia, J.A. (2011). *On coding non-contiguous letter combinations*. Manuscript submitted for publication.

- Davis, C.J. (2010). The spatial coding model of visual word identification. *Psychological Review*, 117, 713-758.
- Davis, M. H. & Rastle, K. (2010). Form and meaning in early morphological processing: Comment on Feldman, O'Connor and Moscoso del Prado Martin. *Psychonomic Bulletin & Review*, 17.
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences*, 9, 335-341.
- Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., et al. (2004). Letter binding and invariant recognition of masked words. *Psychological Science*, 15, 307-313.
- Diependaele, K., Morris, J., & Grainger, J. (2010). *Letter transpositions affect morpho-orthographic not morpho-semantic processing*. Manuscript submitted for publication.
- Diependaele, K., Grainger, J., & Sandra, D. (2010). Derivational morphology and skilled reading: An empirical overview. In M. Spivey, M. Joanisse & K. McRae (Eds.), *The Cambridge Handbook of Psycholinguistics*, Cambridge UK: Cambridge University Press.
- Diependaele, K., Sandra, D., & Grainger, J. (2005). Masked cross-modal morphological priming: Unraveling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes*, 20, 75-114.
- Diependaele, K., Sandra, D., & Grainger, J. (2009). Semantic Transparency and Masked Morphological Priming: The Case of Prefixed Words. *Memory & Cognition*, 37, 895-908.
- Diependaele, K., Ziegler, J., & Grainger, J. (2010). Fast phonology and the bi-modal interactive activation model. *European Journal of Cognitive Psychology*, 22, 764-778.
- Dufau, S., Lété, B., Touzet, C., Glotin, H., Ziegler, J., & Grainger, J. (2010). A developmental perspective on visual word recognition: New evidence and a self-organizing model. *European Journal of Cognitive Psychology*, 22, 669-694.
- Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. B. Gough, L. E. Ehri & R. Treiman (Eds.), *Reading Acquisition* (pp. 105-143). Hillsdale, NJ: LEA.
- Ellis, A. W., & Young, A. (1988). *Human cognitive neuropsychology*. London: Erlbaum.
- Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A

- violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review*, 16, 684-691.
- Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 47A, 365-382.
- Fischer-Baum, S., McCloskey, M., & Rapp, B. (2010). Representation of letter position in spelling: Evidence from acquired dysgraphia. *Cognition*, in press.
- Frankish, C. & Barnes, L. (2008). Lexical and sublexical processes in the perception of transposed-letter anagrams. *The Quarterly Journal of Experimental Psychology*, 61, 381-391.
- Frankish, C. & Turner, E. (2007). SIHGT and SUNOD: The role of orthography and phonology in the perception of transposed letter anagrams. *Journal of Memory and Language*, 56, 189-211.
- Frost, R. (2009). Reading in Hebrew vs. reading in English: Is there a qualitative difference? In K. Pugh & P. McCrindle (Eds.), *How children learn to read: Current issues and new directions in the integration of cognition, neurobiology and genetics of reading and dyslexia research and practice* (pp. 235-254). London: Psychology Press.
- Frost, R., Ahissar, M., Gotesman, R., & Tayeb, S. (2003). Are phonological effects fragile? The effect of luminance and exposure duration on form priming and phonological priming. *Journal of Memory & Language*, 48, 346-378.
- Giraudo, H. & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review*, 8, 127-131.
- Glotin, H., Warnier, P., Dandurand, F., Dufau, S., Lété, B., Touzet, C., Ziegler, J., & Grainger, J. (2010). An adaptive resonance theory account of the implicit learning of orthographic word forms. *Journal of Physiology (Paris)*, 104, 19-26.
- Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431-436.
- Gomez, P., R. Ratcliff, and M. Perea. 2008. The overlap model: A model of letter position coding. *Psychological Review*, 115, 577-601
- Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2001). Pseudohomophone effects and phonological recoding procedures in reading development in English and German. *Journal of Memory and Language*, 45, 648-664.

- Goswami, U., & Ziegler, J. C. (2006). A developmental perspective on the neural code for written words. *Trends in Cognitive Sciences*, 10, 142-143.
- Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes*, 23, 1-35.
- Grainger, J., Colé, P., & Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of Memory and Language*, 30, 370-384.
- Grainger, J., Diependaele, K., Spinelli, E., Ferrand, L., & Farioli, F. (2003). Masked Repetition and Phonological Priming Within and Across Modalities, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29, 1256-1269.
- Grainger, J. & Dufau, S. (2011). 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., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language*, 33, 218-233.
- 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). Watching the Word Go by: On the Time-course of Component Processes in Visual Word Recognition. *Language and Linguistics Compass*, 3, 128-156.
- Grainger, J. & Jacobs, A.M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, 103, 518-565.
- Grainger, J., Kiyonaga, K., & Holcomb, P.J. (2006). The time-course of orthographic and phonological code activation. *Psychological Science*, 17, 1021-1026.
- Grainger, J., Lété, B., Bertrand, D., Dufau, S., & Ziegler, J.C. (2010). *Evidence for multiple routes in learning to read*. Manuscript submitted for publication.
- 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 human mind read words as a whole? *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.
- Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance, *Cognitive Science*, 11, 23-63.
- Jacobs, A.M. & Grainger, J. (1994). Models of visual word recognition - Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1311-1334.
- Jacobs, A.M., Rey, A., Ziegler, J.C, & Grainger, J. (1998). MROM-P: An interactive activation, multiple read-out model of orthographic and phonological processes in visual word recognition. In J. Grainger & A.M. Jacobs (Eds.), *Localist connectionist approaches to human cognition*. Mahwah, NJ.: Erlbaum.
- Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *Neuroimage*, 20, 693–712.
- Kohonen, T. (1982). Self-organizing formation of topologically correct feature maps. *Biological Cybernetics*, 43, 5969.
- Le Pelley, M. E. (2010). Attention and human associative learning. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Associative Learning: From Brain to Behaviour*. Oxford: Oxford University Press.
- Longtin, C.-M. & Meunier, F. (2005). Morphological decomposition in early visual word processing. *Journal of Memory and Language*, 53, 26-41.
- Lukatela, G., & Turvey, M.T. (1994). Visual access is initially phonological: 2. Evidence from phonological priming by homophones, and pseudohomophones. *Journal of Experimental Psychology: General*, 123, 331-353.
- 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.
- Mechelli, A., Crinion, J. T., Long, S., Friston, K. J., Lambon Ralph, M. A., Patterson, K., et al. (2005). Dissociating reading processes on the basis of neuronal interactions. *Journal of Cognitive Neuroscience*, 17, 1753-1765.
- Mel, B.W. & Fiser, J. (2000). Minimizing binding errors using learned conjunctive features. *Neural Computation*, 12, 731-762.

- Morris, J., Frank, T., Grainger, J., & Holcomb, P.J. (2007). Semantic Transparency and Masked Morphological Priming: An ERP Investigation. *Psychophysiology*, 44, 506-521.
- Morris, J., Porter, J. H., Grainger, J. & Holcomb, P. J. (2011). Effects of lexical status and morphological complexity in masked priming: An ERP study. *Language and Cognitive Processes*, in press.
- Moser, M.C. (1991). *The perception of multiple objects: A connectionist approach*. Cambridge, Mass.: The MIT Press.
- O'Connor, R. E. & Forster, K. I. (1981). Criterion bias and search sequence bias in word recognition. *Memory & Cognition*, 9, 78-92.
- Pacton, S. & Perruchet, P. (2008). An attention-based associative account of adjacent and non-adjacent dependency learning. *Journal of Experiment Psychology: Learning, Memory, and Cognition*, 34, 80-96.
- Perea, M., Abu Mallouh, R., & Carreiras, M. (2010). The search of an input coding scheme: Transposed-letter priming in Arabic. *Psychonomic Bulletin and Review*, 17, 375–380.
- Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory & Language*, 51, 231-246.
- Perea, M., Rosa, E., & Gomez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception and Psychophysics*, 67, 301-314.
- Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception and Psychophysics*, 61, 691–706.
- Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory & Language*, 30, 473-485.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. *Psychological Review*, 114, 273-315.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond Single Syllables: Large-Scale Modeling of Reading Aloud with the Connectionist Dual Process (CDP++) Model. *Cognitive Psychology*, in press.

- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56–115.
- Pugh, K.R., Mencl, W.E., Jenner, A.R., Katz, L., Frost, S.J., Lee, J.R., Shaywitz, S.A. & Shaywitz, B.A. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Reviews*, 6, 207–213.
- Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*, 53, 97-145.
- Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, 23, 942-971.
- Rosazza, C., Cai, Q., Minati, L., Paulignan, Y. & Nazir, T.A. (2009). Early involvement of dorsal and ventral pathways in visual word recognition: An ERP study. *Brain Research*, 1272, 32-44.
- Sandak, R., Mencl, W.E., Frost, S.J. & Pugh, K.R. (2004). The neurobiological basis of skilled and impaired reading: Recent findings and new directions. *Scientific Studies of Reading*, 8, 273–292.
- 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, developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55, 151–218.
- Solomyak, O. & Marantz, A. (2009). Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience*, 22, 2042-2057.
- Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, 124, 107-136.
- Ullman, S., Vidal-Naquet, M., & Sali, E. (2002). Visual features of intermediate complexity and their use in classification. *Nature Neuroscience*, 5, 682-687.

- Van Assche, E., & Grainger, J. (2006). A study of relative-position priming with superset primes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 399–415.
- Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, 15, 181-198.
- Velan, H., & Frost, R. (2009). Letter-transposition effects are not universal: The impact of transposing letters in Hebrew. *Journal of Memory and Language*, 61, 285–302.
- Velan, H., & Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic and morphological processing? *Cognition*, 118, 141-156.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron*, 55(1), 143–156.
- von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron*, 24, 95-104
- Welvaert, M., Farioli, F., & Grainger, J. (2008). Graded effects of number of inserted letters in superset priming. *Experimental Psychology*, 55, 54-63.
- Wheat, K.L., Cornelissen, P.L., Frost, S.J., & Hansen, P.C. (2010). During visual word recognition, phonology is accessed within 100 ms and may be mediated by a speed production code: Evidence from Magnetoencephalography. *Journal of Neuroscience*, 30, 5229-5233.
- 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. (2005). Letter-position encoding and dyslexia. *Journal of Research in Reading*, 28, 274-301.
- Whitney, C. & Cornelissen, P. (2008). SERIOL reading. *Language and Cognitive Processes*, 23, 143-164.
- Wilson, L.B., Tregallas, J.R., Slason, E., Pasko, B.E., & Rojas, D.C. (2011). Implicit phonological priming during visual word recognition. *NeuroImage*, 55, 724-731.
- Ziegler, J. C., Ferrand, L., Jacobs, A. M., Rey, A., & Grainger, J. (2000). Visual and phonological codes in letter and word recognition: Evidence from incremental priming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 53A, 671-692.
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological*

Bulletin, 131, 3-29.

- Ziegler, J. C., Jacobs, A. M., & Klueppel, D. (2001). Pseudohomophone effects in lexical decision: Still a challenge for current word recognition models. *Journal of Experimental Psychology: Human Perception & Performance*, 27, 547-559.
- Zorzi, M. (2010). The connectionist dual process (CDP) approach to modelling reading aloud. *European Journal of Cognitive Psychology*, 22, 836-860.
- Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception & Performance*, 24, 1131-1161.