Routlege
Taylor & Francis Group

Si Lexical is as lexical does: computational approaches to lexical representation

Anna M. Woollams*

Neuroscience and Aphasia Research Unit, School of Psychological Sciences, University of Manchester, Zochonis Building, Brunswick Street, Manchester M13 9PL, UK

In much of neuroimaging and neuropsychology, regions of the brain have been associated with 'lexical representation', with little consideration as to what this cognitive construct actually denotes. Within current computational models of word recognition, there are a number of different approaches to the representation of lexical knowledge. Structural lexical representations, found in original theories of word recognition, have been instantiated in modern localist models. However, such a representational scheme lacks neural plausibility in terms of economy and flexibility. Connectionist models have therefore adopted distributed representations of form and meaning. Semantic representations in connectionist models necessarily encode lexical knowledge. Yet when equipped with recurrent connections, connectionist models can also develop attractors for familiar forms that function as lexical representations. Current behavioural, neuropsychological and neuroimaging evidence shows a clear role for semantic information, but also suggests some modality- and task-specific lexical representations. A variety of connectionist architectures could implement these distributed functional representations, and further experimental and simulation work is required to discriminate between these alternatives. Future conceptualisations of lexical representations will therefore emerge from a synergy between modelling and neuroscience.

Keywords: lexical representation; word recognition; orthography; phonology; semantics; computational models

The term 'lexical representation' is commonly found in much recent work concerning the neural bases of normal and disordered language processing, and has a longer history in the context of experimental and theoretical psycholinguistics. Indeed, a search for this term on SCOPUS yields more than 3300 results. Yet, when this term is used, it is often without any formal definition: if one adds 'definition' to our search, this currently yields less than 100 results. Both the popularity of the term and the paucity of formal definition arise from the fact that most researchers in the field of language processing feel that we have an intuitive sense of what 'lexical representation' means. Broadly, it derives from our sense that there is some form of 'mental lexicon' or internal dictionary, in which the knowledge we have concerning the words we know is represented. Of course, our semantic system represents the meaning of known words, and there has been considerable debate concerning the extent to which this obviates the need for any other form of whole word representations. This paper surveys a range of possible implementations of lexical representations within current computational models of word recognition and production with particular attention to the role of semantics. In addition to semantic activation for familiar words, a proposal is put forward for distributed functional lexical representations, which incorporate a degree of modality and task specificity. Recent neuroimaging liter-ature concerning lexical representations is then considered that supports a proposal for multiple levels of distributed functional lexical representations. This highlights the areas in which further research is needed to understand the nature of lexical representations at the cognitive and neural levels.

Why do we need lexical representations?

The existence of some form of lexical representation is inferred when a behavioural processing advantage emerges for a familiar string of letters or phonemes (e.g., *DOG*) over a novel string (e.g., *POG*), which is termed the lexicality effect. The lexicality effect is pervasive across a variety of psycholinguistic tasks, with the key ones that have informed the development of models of written and spoken word recognition including letter/phoneme identi-fication, visual/auditory lexical decision and reading aloud/repetition. In surveying this literature, two issues emerge as important for theoretical interpretation of the lexicality effect. The first is the extent to which the advantage for words is seen over *comparable nonwords*, namely those with subword components that are similar to those seen in words (as measured by bigraph/biphone probabilities and/or neighbourhood/cohort sizes). If a lexicality effect is seen with nonwords that are distin-guishable in terms of their subword properties (e.g., *DOG* vs *ZQF*), this is not necessarily evidence for lexical

*Email: anna.woollams@manchester.ac.uk

representation, but rather of familiarity with subword components. The second is the extent to which lexicality effects seen with comparable nonwords may be driven by activation of *semantic representations*. If we see lexicality effects with closely matched nonwords that are always accompanied by significant effects of semantic variables such as imageability, then this suggests that in fact lexical representations could potentially be reduced to semantic information.

The classic version of the letter identification task involves brief presentation of a string of letters followed by two alternatives over a particular letter position (e.g., *DOG*, followed by a choice of G or T over the third position). In this task, there is a clear advantage for letters presented in a pseudoword relative to an unpronounceable letter string (i.e., better identification of G in *POG* vs *PZG*) (McClelland & Johnston, 1977; Paap, Newsome, McDonald, & Schvaneveldt, 1982). This pseudoword superiority effect indicates familiarity with subword components. More importantly, there is an additional advantage for letters presented in a real word relative to a pronounceable pseudoword (i.e., better identification of G in *DOG* vs *POG*) (Manelis, 1974; McClelland & Johnston, 1977). There is some evidence that this word superiority effect emerges from the familiarity of not only orthography, but also phonology, as participants can be misled when presented with pseudohomophonic non-words, although the activation of phonology may be subject to strategic influences (Hooper & Paap, 1997). While the word superiority effect equates to a lexicality effect, to date there has been no exploration of the influence of semantic variables upon letter identification performance. In the auditory domain, there is a clear influence of lexical knowledge on perception such that when required to judge whether 'p' or 'b' is heard, this is biased by lexicality, such that the shift occurs earlier when moving from 'beace' to 'peace', and later when moving from 'beef' to 'peef' (Ganong, 1980; Miller, Dexter, & Pickard, 1984). We are also more likely to fail to notice a missing phoneme in a heard word than nonword (Samuel, 1996). Whether this effect can be reduced to semantics is unclear, as while semantic context increases the likelihood of phoneme restoration (Liederman et al., 2011), the magnitude of such effects relative to those of lexicality have yet to be directly compared.

Turning to lexical decision, where the task is simply to judge if a string is a word, there is universally a clear advantage for words over nonwords. Lexical decision is a type of signal detection task where the nature of the nonword foils (the noise) relative to the word targets (the signal) critically determines the 'depth' of processing needed for effective discrimination. In visual lexical decision, the lexicality effect is negligible when the nonwords are less plausible (e.g., *KZT*), and increases with pronounceable foils (e.g., *KET*) and is largest with

pseudohomophones (e.g., *KAT*). At the same time, semantic effects, like imageability or semantic priming, increase as the nonwords foils become more word-like (Evans, Lambon Ralph, & Woollams, 2012; James, 1975; Joordens & Becker, 1997). Although lexicality and semantic effects in visual lexical decision emerge in parallel, the lexicality effects observed are larger than the semantic effects (Evans et al., 2012). This difference is, however, difficult to interpret, given that lexicality is inherently confounded with the response required. In auditory lexical decision, lexicality effects are also seen, and semantic effects are apparent for items from larger cohorts (Tyler, Voice, & Moss, 2000). Similar to the visual domain, it is harder to reject more word-like spoken nonwords (Vitevitch & Luce, 1999; Vitevitch, Luce, Pisoni, & Auer, 1999), although the influence of foil type upon performance for words has yet to be investigated in the auditory domain.

A clear lexicality effect is also seen when a spoken response is required, as in reading aloud (McCann & Besner, 1987) and repetition (Vitevitch & Luce, 1998). In repetition, there is an advantage for words over nonwords, and this diminishes the more wordlike the nonword (Vitevitch & Luce, 2005). There is also evidence of semantic effects in repetition (Tyler et al., 2000; Wurm, Vakoch, & Seaman, 2004), although these have not been directly compared to those of lexicality. In reading aloud, there is an issue around the consistency of the spelling to sound correspondences, both for words and nonwords. For words, performance is slower for items containing atypical correspondences, particularly when these items are low in frequency (Jared, 1997, 2002). For nonwords, the same effect can be seen, such that items containing inconsistent correspondences are slower (Andrews & Scarratt, 1998). Hence, the lexicality effect in this task is largest when comparing words with typical correspondences to non-words containing inconsistently pronounced elements. The issue of consistency is also relevant for the presence of semantic effects in this task, with these being strongest for words with atypical correspondences (Shibahara, Zorzi, Hill, Wydell, & Butterworth, 2003; Strain, Patterson, & Seidenberg, 1995, 2002; Woollams, 2005). However, there has yet to be a direct comparison between the size of the semantic effects seen in naming and the lexicality effect obtained with comparable nonwords. It is worth noting that the lexicality effect reflects familiarity of both form and meaning, whereas most semantic manipulations concern variations amongst meanings, and hence may be expected to be weaker.

Evidence from psycholinguistics indicates that the lexicality *effect* is a basic phenomenon that all models of visual and auditory word recognition must accommodate. To the extent that this cannot be reduced to a semantic contribution, then the existence of some form of lexical *representation* is therefore necessary in any model. The

nature of such representations, as will become apparent, can vary from model to model across a variety of dimensions.

What do lexical representations look like?

Models of word recognition have adopted two broad approaches to the implementation of lexical representations: a localist structural view and a distributed functional view. Models that adopt localist representations have one or more sets of units within which there is a unit corresponding to each known word (see Figure 1a). These
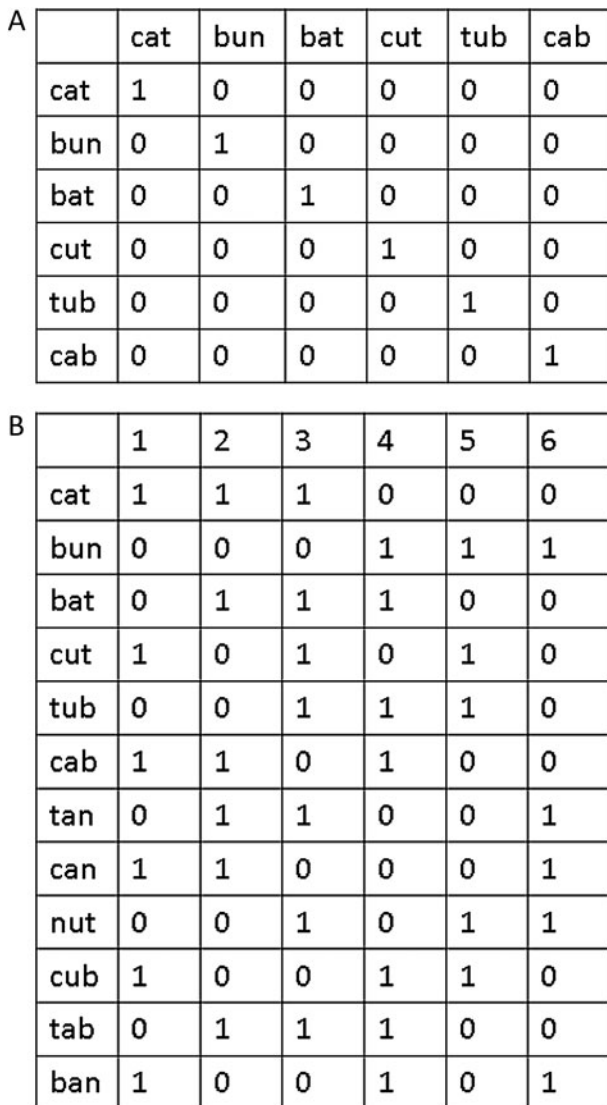

Figure 1. A schematic representation of activation of units encoding (a) localist representations and (b) distributed representations. In (a) representation of six words requires six units. In (b) representation of twelve words also requires six units. In (b) the representations are distributed at the word level but localist at the letter level for the purposes of exposition. A fully distributed scheme would have the capacity to represent many more words.

units are dedicated to the representation of words only, and cannot support accurate processing of novel strings. The lexical units are fixed, in the sense that they exist irrespective of the current stimulus or task requirements. As there is an identity relationship between lexical items and the representational units in the model, then words can be seen as existing as part of the structure of the model.

The localist structural approach has its origins in the earliest modern conceptions of lexical representations, outlined by Morton (1969) in his proposal of the logogen model. Within this verbal theory, the written form of each known word is represented by a unit, or logogen. Each logogen can hold activation for an entry generated by presentation of a letter string, with a higher level of activation the closer the match between the internal and external representation of that string. The resting level of a particular logogen is higher the greater our familiarity with that word, and recognition occurs when activation of a particular logogen exceeds threshold, which explains the advantage for high over low-frequency words (Forster & Chambers, 1973).

The logogen model formed the basis for one of the first implemented computational models of visual word recognition, the Interactive Activation and Competition (IAC) model of Rumelhart and McClelland (1982). In this model, letter features activate letter representations, and in turn these activate localist structural lexical representations. Crucially, this model included full bidirectional connectivity between the word and letter levels, as well as within level inhibitory connections. This allowed the model to account for the lexicality effect in letter perception. An analogous approach to auditory word recognition can be seen in the TRACE model (McClelland & Elman, 1986), which can explain the influence of lexical knowledge on perception of ambiguous phonemes.

The IAC model then provided a foundation for the implementation of the lexical pathway of the Dual Route Cascaded model of visual word recognition and reading aloud (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). This model went beyond previous implementations as it also aimed to describe the process of reading aloud in terms of co-operation between whole word lexical processing and subword rule–based grapheme-to-phoneme conversion. The model, therefore, included not only localist structural orthographic lexical representations in the form of the orthographic input lexicon, but also localist structural phonological representations in the form of the phonological output lexicon, both of which are independent from semantic representation. These two elements are also seen in other models of visual word recognition and reading aloud (e.g., MROM (Grainger & Jacobs, 1996); CDP+ and CDP++ (Perry, Ziegler, & Zorzi, 2007, 2010)). In the auditory domain, localist models have tended to focus more on either perception (e.g., MERGE (Norris,

McQueen, & Cutler, 2000)) or production (e.g., WEAVER (Levelt, Roelofs, & Meyer, 1999)).

The localist structural lexical representation approach is intuitively appealing and computationally relatively easy to implement, and models using this approach have accounted for a great deal of behavioural data concerning word recognition (Coltheart et al., 2001; Levelt et al., 1999; Perry et al., 2007). Yet, one of the major disadvantages of the localist approach of a dedicated structural unit of representation for every word is that it is by no means an efficient system of representation when scaled up to the vocabulary levels of a skilled adult (around 75,000 words (Sibley, Kello, Plaut, & Elman, 2008)). This obviously has consequences when thinking about how such a system may be implemented in the brain, which, although it has billions of neurons, nevertheless needs to accommodate many other functions apart from language processing.

In an attempt to address these concerns, Seidenberg and McClelland (1989) introduced a new approach to the modelling of visual word recognition and reading aloud. Within this model, orthographic and phonological representations consisted of subword units more akin to groups of neurons. These representations are distributed in the sense that the activation of multiple structural units contributes to the representation of a single written or spoken word, and these same units can also represent not only other words but also nonwords. Knowledge concerning words is not stored in the units of the model, but rather on the learnt weights on the connections between them. In this context, lexical representations for a given word are constructed each time they are presented, and hence are functional rather than structural – they do not have a direct relationship to the representational units of the model. This kind of system is more efficient in the sense that a given unit can participate in the representation of multiple different words (see Figure 1b) – more than 6000 units required to represent the written or spoken forms of all monosyllabic words in English can be reduced to 111 orthographic units or 200 phonetic feature units (Harm & Seidenberg, 2004). A similar shift to functional distributed representations of phonology and semantics can be seen in the domain of spoken word recognition (Gaskell & Marslen-Wilson, 1997) and production (Harm & Seidenberg, 2004).

The Seidenberg and McClelland (1989) model represented a major shift in thinking about reading, and its ability to perform lexical decision was scrutinised closely (Besner, Twilley, Seergobin, & McCann, 1990; Fera & Besner, 1992; Seidenberg & McClelland, 1990). Lexical decision is rapidly and accurately achieved by skilled readers and considered to be a basic function that any model of visual word recognition must be able to capture. As distributed functional representations involve units that can also represent nonwords, this was taken by some as problematic for lexical decision. While the initial model

was able to discriminate between words and nonwords on the basis of their subword orthographic and/or phonological familiarity when the nonwords were less familiar, form-based representations were not sufficient in the case where the nonword foils were closely matched to the words (Seidenberg & McClelland, 1990).

In models using distributed functional representations, when the task is to discriminate words from very word-like nonwords, recourse has to be made to semantic information. For example, Plaut's (1997) model included an opaque set of semantic representations (in that each unit did not correspond to an underlying feature). Harm and Seidenberg's (2004) model included a transparent set of semantic representations (where each unit did correspond to an underlying feature), as did Chang, Lambon Ralph, Furber, and Welbourne's (2013) model. All of these models were able to discriminate between words (e.g., *BRAIN*) and closely matched nonwords that are homophonic with real words (i.e., pseudohomophones like *BRANE*) effectively, albeit via different metrics. In Plaut's (1997) model, discrimination was based on semantic stress, which reflects how easily a pattern of activation settles. In Harm and Seidenberg's (2004) model, the discrimination was based on how closely the internally generated orthography matched the stimulus. In Chang et al.'s (2013) model, the decision was made on the basis of pooled activation over orthographic, phonological and semantic units. This latter model has also simulated the larger semantic effects are seen in lexical decision performance for words when presented in a difficult relative to an easy foil context. In these models, therefore, semantic information clearly contributes to the representation of lexical items.

Lexical versus semantic representations

The emergence of an alternative to localist structural representations in the form of distributed functional representations led to considerable debate in the literature concerning the existence of structural lexicons in the mind and brain (e.g., Coltheart, 2004; Elman, 2004). As described earlier, while there is evidence of the involvement of semantic information in tasks demonstrating lexicality effects, it is only neuropsychological evidence that speaks to the necessity of such information in visual and auditory word recognition. The localist structural view invokes one or more form-based lexica that are dedicated to the representation of words, irrespective of semantics. In contrast, the distributed functional view invokes activation of semantic information to support word recognition in the absence of a structural form–based lexicon. This would seem parsimonious given that the ultimate goal of language processing is of course to communicate meaning.

Localist versus distributed approaches offer contrasting perspectives on lexical representation that make rather different predictions concerning the consequences of semantic damage for word recognition. According to the localist structural account, accurate lexical decision should still be possible in the face of severe semantic deficits, as lexical and semantic representations are independent. Evidence for this view is provided by preserved performance in visual and auditory lexical decision tasks in some cases of neuropsychological patients with impaired access to word meaning (Blazely, Coltheart, & Casey, 2005; Bormann & Weiller, 2012). Cases of co-occurring deficits in word recognition and semantic knowledge are considered to be due to the contiguity of semantic and lexical processing areas in the brain (Noble, Glosser, & Grossman, 2000).

The distributed functional approach predicts that semantic impairments will undermine lexical decision performance, but only if the foils are equally familiar to the words in terms of their orthographic and phonological components. Indeed, patients with semantic dementia show very impaired recognition of words when presented with close nonword foils, but accurate performance with distant nonword foils (Patterson et al., 2006; Rogers, Lambon Ralph, Hodges, & Patterson, 2004). This view proposes that the cases of semantic impairment with preserved lexical decision performance arise from the use of distractors that can be discriminated from words on the basis of their orthographic or phonological properties (Woollams, Ralph, Plaut, & Patterson, 2007).

It is worth noting that the implementation of semantic representations does vary over models employing distributed functional representations. Most models dealing with word recognition are trained to instantiate target semantic patterns, and in this sense, meaning functions as an output layer (Chang et al., 2013; Gaskell & Marslen-Wilson, 1997; Harm & Seidenberg, 2004; Plaut, 1997). In other models considering multiple modalities beyond word recognition, then the hidden units that mediate between the external input/output layers are allowed to develop their own representational structure (Dilkina, McClelland, & Plaut, 2008, 2010; Rogers, Lambon Ralph, Garrard, et al., 2004). In either case, amodal semantic representations can function as lexical representations. However, this is not the only possible location for distributed representations that can index word knowledge within connectionist models.

Lexical representations as attractors

Connectionist models of word recognition have usually considered inputs and outputs as some form of orthographic or phonological representations (Dilkina et al., 2010; Gaskell & Marslen-Wilson, 1997; Harm & Seidenberg, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996). As such, these do not encode the visual or auditory signal per se, but our judgement as to a representational system that captures the salient aspects of the domain (e.g., letter or phonetic features). When semantic and phonological representations are trained in the presence of recurrent connections, attractors form that represent known patterns (Plaut et al., 1996; Plaut & Shallice, 1993). Recurrence involves feedback connections between levels of representation, and/or connections between units within a layer (this latter can also be accomplished via connections to a smaller set of clean-up units, e.g. Plaut & Shallice, 1993). The point of recurrence is that it allows the activation of a unit to be affected by other units in a continuous fashion over time. This means that there is a variety of initial activation values that will come to converge on the same final pattern. The initial range of activation values forms the attractor basin for a particular item.

The utility of attractor networks in the simulation of normal and impaired reading has been demonstrated at the level of semantics (Plaut & Shallice, 1993) and at the level of phonology (Plaut et al., 1996, Simulation 3). This latter simulation demonstrated that although semantics seems to naturally encode whole word knowledge in connectionist networks, functional lexical representations can develop in networks with a direct mapping between orthographic and phonological forms. The formation of such attractors does not imply that novel patterns cannot be represented effectively, but rather that words, as familiar patterns, will have their distributed functional representation instantiated more easily over time, as initial approximate patterns of activation can fall into the attractor, whereas the initial patterns for novel strings must avoid doing so. Attractor networks therefore naturally reproduce a processing advantage for words over nonwords.

It is possible that such attractors could be harnessed to inform word recognition and support performance in tasks like lexical decision independently of semantic activation, although this possibility has been under-explored to date. This is because recurrence at the input level risks drowning out the original signal – if pre-existing internal knowledge is allowed to influence activation very early via top–down connections, then perception will be biased towards familiar items and may become inaccurate as processing progresses (e.g., novel strings come to be seen as familiar words). As most connectionist models start with input layers of orthographic or phonological representation, then recurrent connections are not included and attractors will not form. This is not, however, because there is an assumption that this is in fact the raw input to the word recognition process, but rather because the challenges associated with modelling perception of raw inputs have been so great that it was more tractable to start with higher-level representations.

More recently, models using distributed functional representations have begun to consider mapping from raw visual inputs to meaning (Chang et al., 2013; Plaut & Behrmann, 2011), or from the acoustic signal to meaning to articulatory features (Kello & Plaut, 2004; Plaut & Kello, 1999). In these kinds of models, what we think of as orthographic or phonological lexical representations are more likely to be found in the activations of the hidden units that map between the physical inputs to outputs, either directly or via meaning (a similar idea has been eloquently articulated by Elman (2004)). The intermediate location of the orthographic/phonological representations in such models allows for recurrent connections and the formation of attractors as these are no longer direct input layers. A depiction of this kind of model is given in Figure 2. In addition, the structure of such representations will be fully distributed, in that the units no longer need to correspond to letters or phonetic features, which is even more economical. Although it has been suggested that the notion of lexical representation as hidden unit activations lacks explanatory capacity (Green, 1998), the structure of such units can be revealed using techniques such as cluster analysis (e.g., Chang, Furber, & Welbourne, 2012; Rogers, Lambon Ralph, Garrard, et al., 2004).

Specificity of lexical representations

The concept of a lexical representation is an internal one, in that it implies at some level a degree of abstraction from physical inputs and outputs. The idea that connectionist models could contain a form of 'lexical' representation in the hidden units that map between different input and output domains has the advantage that it allows for a degree of specificity in terms of modality (written/spoken) and task (recognition/production), as shown in Figure 2. The need for such specificity has been suggested by the study of neuropsychological patients. With respect to modality there have been a number of reports of brain-damaged patients with problems with reading yet intact speech comprehension associated with damage to regions of the left ventral occipito-temporal cortex (vOTC) (Roberts, Lambon Ralph, & Woollams, 2010; Roberts et al., 2013; Woollams, Hoffman, Roberts, Lambon Ralph, & Patterson, 2014). Conversely, there are reports of patients with clear deficits in speech comprehension as a result of perisylvian lesions who nevertheless demonstrate relatively proficient reading comprehension (Ellis, Miller, & Sin, 1983; Lytton & Brust, 1989).

There is also some evidence of specialisation according to task requirements for comprehension or production. One of the most striking demonstrations is seen in pure alexia, where patients with profound reading difficulties are nevertheless able to write proficiently (Turkeltaub et al., 2014). This contrasts with cases of pure dysgraphia, where spelling is compromised but reading ability is retained (Miceli, Silveri, & Caramazza, 1985). Turning to spoken language processing, rare patients with word
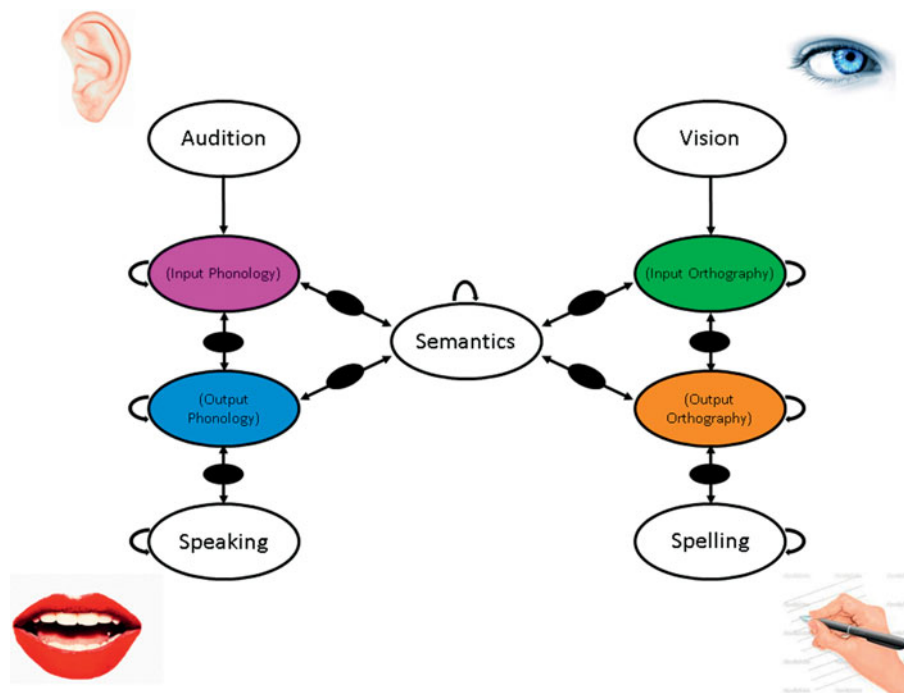

Figure 2. A schematic diagram of a model of visual and auditory word recognition and production showing the location of hidden unit layers that could house distributed functional lexical representations in the form of attractors. Note bidirectional connections in all cases bar those from input. Additional hidden layers are shown in black. Within level connections are shown with U-shaped arrows.

deafness due to bilateral posterior perisylvian damage are unable to understand single words but manage to speak fluently (Best & Howard, 1994). Conversely, many nonfluent patients with profound speech difficulties due to damage to Broca's area can still show good speech comprehension at the single word level (Hickok, Costanzo, Capasso, & Miceli, 2011). It should be noted, however, that it is not yet clear to what extent these modality-specific problems are due to disruption of internal orthographic/phonological representations or rather to deficits in more basic visual and auditory processing.

Modality and task specificity of hidden unit representations can be handled in a number of different architectures in connectionist models, with the most widely used to date being the separation of hidden units onto different processing pathways (e.g., Harm & Seidenberg, 2004; Chang et al., 2013), as shown in Figure 2. A different solution in which modality specificity was captured across a set of hidden units linked their inputs or outputs was adopted by Plaut (2002), and this framework was extended to word recognition by Dilkina et al. (2010). Plaut's (2002) model set out to simulate performance in optic aphasia, where a patient is unable to name an item when presented visually, but can do so on the basis of other inputs, such as touch. Within this model, the proximity of the hidden units in two-dimensional space to a particular set of inputs or outputs captured their degree of specialisation in a graded rather than categorical way. The graded specialisation emerges because the weights on the connections of input to hidden units are biased to become stronger for closer units. When the hidden units close to visual inputs and phonological outputs were lesioned, the modality-specific naming deficit seen in optic aphasia emerged. A version of this approach applied to visual and auditory word recognition is presented in Figure 3. An interesting aspect of this approach is that as lesions to the model move towards the middle of the hidden unit space, the more domain general (i.e., semantic) their effects become. Although the spatial location of the hidden units in such models is not intended to map directly on to neural structure (Plaut, 2002), these models are interesting in that the graded specialisation according to modality and task seems to map well on to recent neuroimaging evidence concerning visual and auditory word recognition (e.g., Binney, Parker, & Lambon Ralph, 2012; Vinckier et al., 2007).

Localisation of lexical representations

Although connectionist approaches with distributed representations were initially motivated by a desire to seek a model that is more neurally plausible than those described previously, it is only in more recent times that neural data has been used explicitly as an additional constraint in model construction and evaluation (Dell, Schwartz, Nozari, Faseyitan, & Branch Coslett, 2013; Laszlo & Plaut, 2012; Ueno, Saito, Rogers, & Lambon Ralph, 2011). Not all agree that information concerning neural structure and function are informative in terms of understanding the cognitive mechanisms that permit written and spoken word recognition and production (Coltheart, 2006, 2013). If, however, 'the fact of the matter lies in the brain' (Davis, 2004), then neuroimaging data provides a valuable source of information with which to limit the space of potential computational implementations.

The notion of distributed functional representations that are to some extent modality and task specific predicts that there will be more than one area of the brain that will show sensitivity to lexicality. Recent meta-analyses of both written (Cattinelli, Borghese, Gallucci, & Paulesu, 2013; Taylor, Rastle, & Davis, 2013) and spoken word recognition and production (Gow, 2012; Vigneau et al., 2006) have converged on a framework in which direct mappings between form representations (i.e., reading aloud/repetition) are attributed to a dorsal pathway while semantically mediated mappings used for word recognition are associated with a largely ventral pathway. In all of these reviews, lexicality effects are observed in multiple areas along *both* pathways. The subword direct mappings between input and outputs are flagged by greater activation for nonwords than words. In contrast, the whole word semantically mediated mappings are indicated by greater activation for words than nonwords.

It is in the areas more responsive to words than nonwords that some form of lexical representation would seem indicated. Interestingly, posterior parietal regions including the angular gyrus that show lexicality effects overlap with areas implicated in both semantic and phonological processing, and it has been suggested that this area acts as a 'phonological lexicon' (Davis & Gaskell, 2009; Gow, 2012; Taylor et al., 2013). Similarly, the inferior temporal regions including the fusiform gyrus that show lexicality effects have been implicated in both semantic and orthographic processing, suggesting that it functions as an 'orthographic lexicon' (Cattinelli et al., 2013; Davis & Gaskell, 2009; Gow, 2012; Taylor et al., 2013). These lexicality effects hold even when only studies that have carefully matched their nonwords to the words are included (Taylor et al., 2013). The interpretation of the effects in these areas is not straightforward, however. They may indicate form-based lexical representations independent of semantic activation, or the intersection of form-based and semantic processing. Both of these are consistent with the proposal of lexical representations as attractors (Chang et al., 2013; Plaut, 1996), although they vary in the units across which these would form (form-based representation vs. hidden units closer to semantics). But there is another possibility, which is that these areas are only showing higher activation for words
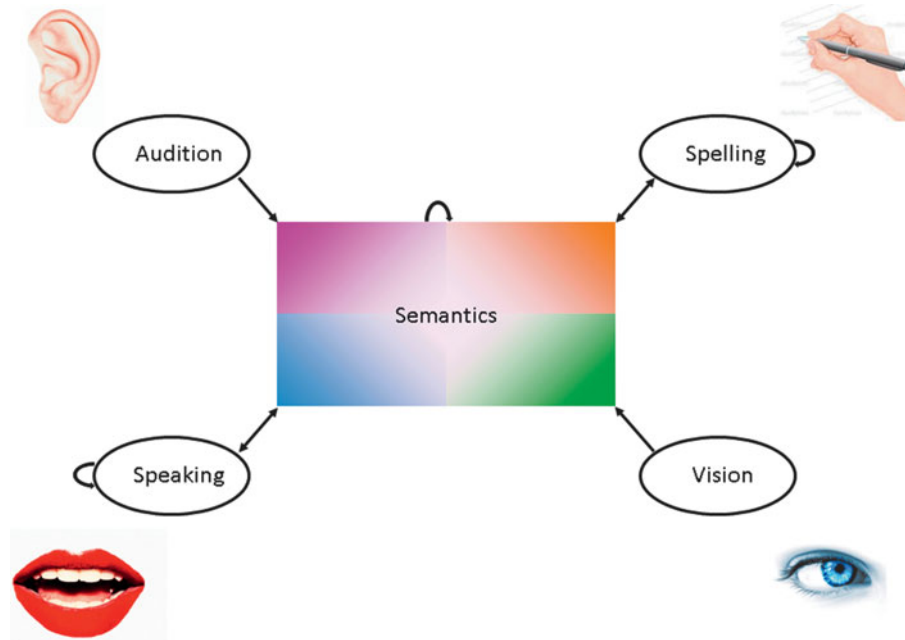
Figure 3.   A version of the previous model of visual and auditory word recognition and production containing one large set of hidden units. Learning in the network occurs under a topographic bias that favours short connections. This allows graded modality specificity to emerge in the network, such that units close to a particular input or output participate more in tasks involving them, while units close to the centre are increasingly amodal.

due to feedback from higher-level semantic representations, which is consistent with the idea that lexical representations equate to semantic representations.

The use of more parametric or graded approaches to the design of neuroimaging studies on written and spoken word recognition can be informative in teasing apart alternative interpretations of the lexicality effect. An elegant study by Vinckier et al. (2007) revealed that with the progression of activation from raw visual input along the ventral pathway to the most anterior portions of the temporal lobe that contain amodal semantic representations, there is increasing sensitivity to larger orthographic clusters from single letters, to bigrams, to quadrigrams. Similarly, work by Hauk, Davies et al. (2006) has used an item-based regression approach to identify areas associated with different dimensions of form through to meaning during visual lexical decision (see also Graves, Desai, Humphries, Seidenberg, & Binder, 2010 for reading aloud). The results revealed sensitivity to more form-based variables like length and typicality for both words and nonwords in posterior brain areas, moving through to higher-level intermediate form-based representations that indexed lexicality and frequency, with more anterior areas reflecting an influence of semantic coherence (the degree to which meaning is consistent across morphologically related words). These imaging studies indicate a gradient of representation from more posterior form based to more anterior meaning based along the temporal lobe in visual word recognition. Sensitivity to

lexicality in mid to posterior regions does not seem to reduce to semantic feedback as these areas are not sensitive to semantic dimensions.

As proposed earlier, in connectionist models, what we think of as orthographic or phonological lexical representations may well be found in the activations of the hidden units that map between physical inputs and outputs. Within models that incorporate a single set of hidden units and encode their modality and task specificity in terms of proximity between inputs and outputs (e.g., Dilkina et al., 2008; Plaut, 2002), the gradation from modality-specific form-based representation at the edges of the space to progressively more amodal semantic representations at the centre of the space agrees with recent conceptions of the ventral pathway in visual and auditory word recognition (see Figure 4 from Binney et al., 2012). Linking the previous proposal concerning lexical representations in hidden units to neuroimaging, along the ventral pathway, 'lexical representations' would be located at the intersection of form and meaning-based information. They can be localised by looking for activation that permits reliable discrimination between words and nonwords in a particular task, with newer approaches such as searchlight analysis of patterns of activation over multiple voxels seeming particularly promising in this regard (e.g., Nestor, Behrmann, & Plaut, 2013). The more similar the nonwords to words in any given task, the closer activation should move to areas involved in representation of meaning.

Not only would the precise location of lexical representation depend on stimulus properties, it would also be affected by task requirements (e.g., Binder, Medler, Westbury, Liebenthal, & Buchanan, 2006; Chen, Davis, Pulvermüller, & Hauk, 2013; Gan, Büchel, & Isel, 2013; Twomey, Kawabata Duncan, Price, & Devlin, 2011). For example, the direction of the lexicality effect observed in vOTC depends upon whether the task involves passive viewing or active discrimination (e.g., Vinckier et al., 2007; cf. Woollams, Silani, Okada, Patterson, & Price, 2011), and its presence appears to depend on the opportunity for phonological recoding (Mano et al., 2013). Perhaps even more strikingly, differing patterns of lexicality effects are observed across multiple brain regions in matching versus reading tasks for exactly the same stimuli (Vogel, Petersen, & Schlaggar, 2013). Moreover, while we see higher activation for words than nonwords in areas that process orthographic and phonological input, there is also a network of left frontal regions, which show less activation for easy-to-pronounce words than hard-to-pronounce words and nonwords (Cattinelli et al., 2013; Taylor et al., 2013). In this case, the lexicality effect occurring in speech production regions shows less effort associated with familiar word stimuli. The mediating effects of stimulus and task on the location of the brain areas specifically sensitive to the lexicality of orthographic or phonological strings emphasises the fluidity of lexical representations, as expected according to a distributed functional view.

Activation of lexical representations

Connectionist models of word recognition are highly interactive, and feedback from higher to lower levels is integral in terms of the formation and operation of these models. As noted earlier, recurrent connections are essential for the formation of the attractors that this paper nominates as a potential candidate for functional lexical representations. From IAC and TRACE onwards, the majority of models of word recognition permit some degree of feedback activation from whole to subword information (although this is not carried as far as the input layer). A commitment to feedback is also reflected in current neural models of language processing (e.g., Price & Devlin, 2011; Ueno et al., 2011). However, others have strongly challenged the necessity of feedback connections in cognitive models of word recognition (e.g., McQueen, Cutler, & Norris, 2003; Norris et al., 2000) and also neural models (Davis, Ford, Kherif, & Johnsrude, 2011). It is very difficult to conclusively establish a role for feedback in word recognition in behavioural studies, as the response represents a summation of activation over the course of processing, and is therefore subject to 'post-lexical' effects. The same problem arises in studies using PET or fMRI, as again, the brain activation observed is a summation of processes over time. This leads to interpretative difficulties when dealing with lexicality effects in areas associated with form-based processing, as it is unclear to what extent these may be reduced to feedback from higher-level semantic representation.

However, neuroimaging modalities that offer good temporal resolution, such as EEG or MEG, can offer insight into the flow of brain activation during word recognition and production (for reviews see: Carreiras, Armstrong, Perea, & Frost, 2014; Mattys, Davis, Bradlow, & Scott, 2012). This can allow an understanding of the time course of the component processes involved in a behaviour, such as picture naming (Indefrey, 2011), and can also speak to the interactivity between these different levels of representation. There is now an emerging imaging literature that strongly supports the very rapid influence of higher-level semantic and phonological knowledge upon visual and auditory word recognition. For example, in an ERP study of visual lexical decision (Hauk, Patterson, et al., 2006), initial activation in the region of left vOTC at 100 ms was driven by bigram frequency in a visual lexical decision task, but after strong activation in the left anterior temporal lobe at around 150 ms for words of low bigram frequency, the next surge of left vOTC activation patterned with lexicality at 200 ms, which was well before the behavioural response (see also Hauk, Coutout, Holden, & Chen, 2012). Turning to speech, Sohoglu, Peelle, Carlyon and Davis (2012) manipulated prior knowledge of the content of degraded speech using previously presented text, and found that this modulated EEG/MEG activity in the left inferior frontal gyrus as early as 90–130 ms, which was well before that seen in the superior temporal gyrus after 270 ms.

These results are very interesting because they indicate a role for feedback during word recognition in a way not permitted by consideration of behaviour alone. There is clearly very early activation of regions involved in semantic processing. The fact that these areas show sensitivity to lexicality/familiarity before more posterior form-based regions could be interpreted as supporting a view in which lexical representations reduce to top–down activation from semantic representations. However, it may also be that these form-based regions follow a different time course of activation, with initial sensitivity to subword properties shifting to a lexicality effect. This latter possibility could be accommodated by the distributed functional view where attractors act as lexical representations. It is very difficult to determine the extent to which early activation in one brain region causes a pattern of activation in another region, and consideration of the time course of unit activations in connectionist models could provide concrete predictions to test in future neuroimaging experiments.

Development of lexical representations

Connectionist models, by definition, derive their representations via adjustments on initially random weights between processing units through a process of learning that involves exposure and usually error correction. As such, issues around learning of lexical representations may seem to map fairly directly onto the issue of distributed versus localist representations; however, this is not necessarily so. There are in fact models that include localist structural representations where the weights on connections between them are learnt (e.g., Dandurand, Hannagan, & Grainger, 2013; Mirman, McClelland, & Holt, 2006; Perry et al., 2007, 2010). When this is adopted as a theoretical standpoint, it can be considered 'localist connectionism' (Grainger & Jacobs, 1998); however, localist structural representations have also been adopted in the connectionist domain to render large-scale models tractable in terms of computational processing demands (Kello, 2006; Sibley et al., 2008).

Allowing the connections between localist representations to be learnt does not speak directly to their initial formation, yet understanding of this process is crucial if we are to accurately capture language development. A recent simulation allows for the formation of a localist lexical orthographic representation once initial translation of a letter string via subword orthography to phonology mappings makes contact with a pre-existing phonological lexical representation (Ziegler, Perry, & Zorzi, 2014). This process is a computational implementation of Share's (1995) self-teaching hypothesis, but it merely defers the problem of formation of a lexical entry from the orthographic to the phonological level – it does not tell us how the phonological lexical representations are formed. In contrast, in models that map directly from raw visual or auditory inputs, distributed lexical representations will emerge over the course of learning in the hidden units that mediate access to semantic and articulatory outputs via recurrent connections within and between layers (Chang et al., 2013; Plaut & Kello, 1999). The formation of distributed functional lexical representations therefore falls naturally from models that focus on mapping raw inputs to meaning and speech.

While connectionist models learn their representations, this does not mean that they necessarily do so in precisely the same way as children. Some connectionist models of reading development have captured this process well with a standard training regime (e.g., Harm & Seidenberg, 1999; Karaminis & Thomas, 2010; Ueno et al., 2011). Other models have used pre-training and more psychologically plausible training vocabularies and techniques, which have improved the fit of the models to children's data considerably (Hutzler, Ziegler, Perry, Wimmer, & Zorzi, 2004; Powell, Plaut, & Funnell, 2006). Clearly, developmental trajectories, like functional neuroanatomy,

reflect a potential source of constraint upon model design that can be utilised in future. More generally, the fact that connectionist models can learn their representations allows them to speak to issues concerning not only normal and disordered language development (Harm, McCandliss, & Seidenberg, 2003; Harm & Seidenberg, 1999; Joanisse & Seidenberg, 2003), but also language function in progressive disorders (Rogers, Lambon Ralph, Garrard, et al., 2004; Woollams et al., 2007) and the possibility for relearning and rehabilitation after brain damage (Abel, Grande, Huber, Willmes, & Dell, 2005; Plaut, 1996; Welbourne & Lambon Ralph, 2005, 2007).

Conclusions and future directions

This paper has considered a number of different computational implementations of lexical representations that seem warranted to account for the lexicality effects seen in visual and auditory language-processing tasks. Distributed functional representations were favoured by virtue of their efficiency and flexibility. Within connectionist models using such representations, one possibility is that lexical knowledge reduces to semantic activation. An alternative proposal outlined here is that whole word knowledge is also captured by a system of attractors at levels intermediate between form and meaning, formed during learning via recurrent connections between units within and across levels. Within this view, the nature of the stimuli and the demands of the task will determine where and when these functional lexical representations emerge, hence they are fluid. This latter proposal has the virtue of allowing a degree of modality and task specificity suggested by selective impairments in some neuropsychological patients and variability in the timing and location of lexicality effects in neuroimaging studies.

Further research is needed to tease apart the contributions of semantic versus form-based lexical knowledge to word recognition. This will require direct comparisons between lexicality effects and semantic effects, and also more parametric manipulations sensitive to graded specialisation. Comparisons over modalities and tasks are needed to determine the degree of specificity of lexical representations. Such research will preferably use neuroimaging techniques with sufficient temporal resolution to understand the flow of activation within the system prior to behavioural response. Simulation of these data will require models that start with approximations of raw visual and auditory inputs and incorporate at least some degree of recurrence between and within levels. Within such attractor networks, it is necessary to explore how the various possible performance metrics correspond to both behavioural and neuroimaging data.

In summary, this paper has highlighted that there are a variety of implementation options for lexical representations in connectionist models, both in terms of the nature

of the semantic system and also for distributed functional representations as hidden unit attractors, ranging from discrete sets of hidden units for each modality/task to graded specialisation within a single integrative layer. Neuroimaging evidence can be used to constrain the range of possible implementations, but most current analysis techniques yield discrete cortical areas and are not sensitive to graded influences of form- and meaning-based variables. Neuroimaging evidence has its own limitations in terms of the cognitive interpretation of activation differences, and consideration of hidden unit dynamics in connectionist models with different architectures may shed light on the neural data. Ultimately, a full understanding of the source of lexicality effects in behaviour will require a dialogue between modelling and neuroscience through which they will converge on the precise form of lexical representations in the mind and brain.

Disclosure statement

Funding

References

Abel, S., Grande, M., Huber, W., Willmes, K., & Dell, G. S. (2005). Using a connectionist model in aphasia therapy for naming disorders. *Brain and Language*, *95*(1 SPEC. ISS.), 102–104.

Andrews, S., & Scarratt, D. R. (1998). Rule and analogy mechanisms in reading non words: Hough Dou Peapel Rede Gnew Wirds? *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1052–1086. doi:10.1037/0096-1523.24.4.1052

Besner, D., Twilley, L., Seergobin, K., & McCann, R. S. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review*, *97*, 432–446. doi:10.1037/0033-295X.97.3.432

Best, W., & Howard, D. (1994). Word sound deafness resolved? *Aphasiology*, *8*, 223–256. doi:10.1080/02687039408248655

Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., & Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *NeuroImage*, *33*, 739–748. doi:10.1016/j.neuroimage.2006.06.053

Binney, R. J., Parker, G. J. M., & Lambon Ralph, M. A. (2012). Convergent connectivity and graded specialization in the rostral human temporal Lobe as revealed by diffusion-weighted imaging probabilistic tractography. *Journal of Cognitive Neuroscience*, *24*, 1998–2014. doi:10.1073/pnas.0607061104

Blazely, A. M., Coltheart, M., & Casey, B. J. (2005). Semantic impairment with and without surface dyslexia: Implications for models of reading. *Cognitive Neuropsychology*, *22*, 695–717. doi:10.1080/02643290442000257

Bormann, T., & Weiller, C. (2012). 'Are there lexicons?' A study of lexical and semantic processing in word-meaning deafness suggests 'yes'. *Cortex*, *48*, 294–307.

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, *18*(2), 90–98. doi:10.1016/j.tics.2013.11.005

Cattinelli, I., Borghese, N. A., Gallucci, M., & Paulesu, E. (2013). Reading the reading brain: A new meta-analysis of functional imaging data on reading. *Journal of Neurolinguistics*, *26*, 214–238. doi:10.1016/j.jneuroling.2012.08.001

Chang, Y.-N., Furber, S., & Welbourne, S. (2012). Modelling normal and impaired letter recognition: Implications for understanding pure alexic reading. *Neuropsychologia*, *50*, 2773–2788. doi:10.1016/j.neuropsychologia.2012.07.031

Chang, Y.-N., Lambon Ralph, M. A., Furber, S., & Welbourne, S. (2013). *Modelling graded semantic effects in lexical decision*. Proceedings of the 35th Annual Meeting of the Cognitive Science Society, July 31-August 3, Berlin, Germany, 310–315.

Chen, Y., Davis, M. H., Pulvermüller, F., & Hauk, O. (2013). Task modulation of brain responses in visual word recognition as studied using EEG/MEG and fMRI. *Frontiers in Human Neuroscience (JUN)*, *7*, 376.

Coltheart, M. (2004). Are there lexicons? *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *57*, 1153–1171. doi:10.1080/02724980443000007

Coltheart, M. (2006). What has functional neuroimaging told us about the mind (so far)? *Cortex*, *42*, 323–331. doi:10.1016/S0010-9452(08)70358-7

Coltheart, M. (2013). How can functional neuroimaging inform cognitive theories? *Perspectives on Psychological Science*, *8*(1), 98–103. doi:10.1177/1745691612469208

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256. doi:10.1037/0033-295X.108.1.204

Dandurand, F., Hannagan, T., & Grainger, J. (2013). Computational models of location-invariant orthographic processing. *Connection Science*, *25*(1), 1–26. doi:10.1080/09540091.2013.801934

Davis, M. H. (2004). Units of representation in visual word recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 14687–14688. doi:10.1073/pnas.0405788101

Davis, M. H., Ford, M. A., Kherif, F., & Johnsrude, I. S. (2011). Does semantic context benefit speech understanding through 'top-down' processes? Evidence from time-resolved sparse fMRI. *Journal of Cognitive Neuroscience*, *23*, 3914–3932. doi:10.1016/j.neuroimage.2006.04.199

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 3773–3800.

Dell, G. S., Schwartz, M. F., Nozari, N., Faseyitan, O., & Branch Coslett, H. (2013). Voxel-based lesion-parameter mapping: Identifying the neural correlates of a computational model of word production. *Cognition*, *128*, 380–396. doi:10.1016/j.cognition.2013.05.007

Dilkina, K., McClelland, J. L., & Plaut, D. C. (2008). A single-system account of semantic and lexical deficits in five semantic dementia patients. *Cognitive Neuropsychology*, *25*(2), 136–164. doi:10.1080/02643290701723948

Dilkina, K., McClelland, J. L., & Plaut, D. C. (2010). Are there mental lexicons? The role of semantics in lexical decision.

Brain Research, *1365*, 66–81. doi:10.1016/j.brainres.2010.09.057

Ellis, A. W., Miller, D., & Sin, G. (1983). Wernicke's aphasia and normal language processing: A case study in cognitive neuropsychology. *Cognition*, *15*(1–3), 111–144. doi:10.1016/0010-0277(83)90036-7

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, *8*, 301–306. doi:10.1016/j.tics.2004.05.003

Evans, G. A. L., Lambon Ralph, M. A., & Woollams, A. M. (2012). What's in a word? A parametric study of semantic influences on visual word recognition. *Psychonomic Bulletin and Review*, *19*, 325–331. doi:10.3758/s13423-011-0213-7

Fera, P., & Besner, D. (1992). The process of lexical decision: More words about a parallel distributed processing model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 749–764. doi:10.1037/0278-7393.18.4.749

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*, 627–635. doi:10.1016/S0022-5371(73)80042-8

Gan, G., Büchel, C., & Isel, F. (2013). Effect of language task demands on the neural response during lexical access: A functional magnetic resonance imaging study. *Brain and Behavior*, *3*, 402–416. doi:10.1002/brb3.133

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 110–125. doi:10.1037/0096-1523.6.1.110

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*, 613–656. doi:10.1080/016909697386646

Gow, D. W. (2012). The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. *Brain and Language*, *121*, 273–288. doi:10.1016/j.bandl.2012.03.005

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*, 518–565. doi:10.1037/0033-295X.103.3.518

Grainger, J., & Jacobs, A. M. (1998). Localist connectionism fits the bill. *Psycoloquy*, *9*, 7.

Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2010). Neural systems for reading aloud: A multiparametric approach. *Cerebral Cortex*, *20*, 1799–1815. doi:10.1093/cercor/bhp245

Green, C. D. (1998). Are connectionist models theories of cognition? *Psycoloquy*, *9*, 8.

Harm, M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, *7*, 155–182. doi:10.1207/S1532799XSSR0702_3

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*, 491–528. doi:10.1037/0033-295X.106.3.491

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662–720. doi:10.1037/0033-295X.111.3.662

Hauk, O., Coutout, C., Holden, A., & Chen, Y. (2012). The time-course of single-word reading: Evidence from fast behavioral and brain responses. *NeuroImage*, *60*, 1462–1477. doi:10.1016/j.neuroimage.2012.01.061

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*, *30*, 1383–1400. doi:10.1016/j.neuroimage.2005.11.048

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermüller, F., & Rogers, T. T. (2006). [Q:] When would you prefer a sossage to a sausage? [A:] At about 100 msec. ERP correlates of orthographic typicality and lexicality in written word recognition. *Journal of Cognitive Neuroscience*, *18*, 818–832.

Hickok, G., Costanzo, M., Capasso, R., & Miceli, G. (2011). The role of Broca's area in speech perception: Evidence from aphasia revisited. *Brain and Language*, *119*, 214–220. doi:10.1016/j.bandl.2011.08.001

Hooper, D. A., & Paap, K. R. (1997). The use of assembled phonology during performance of a letter recognition task and its dependence on the presence and proportion of word stimuli. *Journal of Memory and Language*, *37*, 167–189. doi:10.1006/jmla.1997.2520

Hutzler, F., Ziegler, J. C., Perry, C., Wimmer, H., & Zorzi, M. (2004). Do current connectionist learning models account for reading development in different languages? *Cognition*, *91*, 273–296. doi:10.1016/j.cognition.2003.09.006

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 255.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(2), 130–136. doi:10.1037/0096-1523.1.2.130

Jared, D. (1997). Spelling-sound consistency affect the naming of high-frequency words. *Journal of Memory and Language*, *36*, 505–529. doi:10.1006/jmla.1997.2496

Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language*, *46*, 723–750. doi:10.1006/jmla.2001.2827

Joanisse, M. F., & Seidenberg, M. S. (2003). Phonology and syntax in specific language impairment: Evidence from a connectionist model. *Brain and Language*, *86*(1), 40–56. doi:10.1016/S0093-934X(02)00533-3

Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning Memory and Cognition*, *23*, 1083–1105. doi:10.1037/0278-7393.23.5.1083

Karaminis, T., & Thomas, M. (2010). A cross-linguistic model of the acquisition of inflectional morphology in English and Modern Greek. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of 32nd Annual Conference of the Cognitive Science Society* (pp. 730–735), August 11–14, 2010, Portland.

Kello, C. T. (2006). Considering the junction model of lexical processing. In S. A. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 50–75). Hove: Psychology Press.

Kello, C. T., & Plaut, D. C. (2004). A neural network model of the articulatory-acoustic forward mapping trained on recordings of articulatory parameters. *Journal of the Acoustical Society of America*, *116*, 2354–2364. doi:10.1121/1.1715112

Laszlo, S., & Plaut, D. C. (2012). A neurally plausible parallel distributed processing model of event-related potential word reading data. *Brain and Language*, *120*, 271–281.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–75.

Liederman, J., Gilbert, K., Fisher, J. M., Mathews, G., Frye, R. E., & Joshi, P. (2011). Are women more influenced than men by top-down semantic information when listening to disrupted speech? *Language and Speech*, *54*(1), 33–48.

Lytton, W. W., & Brust, J. C. M. (1989). Direct dyslexia. Preserved oral reading or real words in Wernicke's aphasia. *Brain*, *112*, 583–594.

Manelis, L. (1974). The effect of meaningfulness in tachistoscopic word perception. *Perception and Psychophysics*, *16*, 182–192. doi:10.3758/BF03203272

Mano, Q. R., Humphries, C., Desai, R. H., Seidenberg, M. S., Osmon, D. C., Stengel, B. C., & Binder, J. R. (2013). The role of left occipitotemporal cortex in reading: Reconciling stimulus, task, and lexicality effects. *Cerebral Cortex*, *23*, 988–1001.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*, 953–978.

McCann, R. S., & Besner, D. (1987). Reading Pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 14–24.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. doi:10.1016/0010-0285(86)90015-0

McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception and Psychophysics*, *22*, 249–261. doi:10.3758/BF03199687

McQueen, J. M., Cutler, A., & Norris, D. (2003). Flow of information in the spoken word recognition system. *Speech Communication*, *41*, 257–270. doi:10.1016/S0167-6393(02)00108-5

Miceli, G., Silveri, M. C., & Caramazza, A. (1985). Cognitive analysis of a case of pure dysgraphia. *Brain and Language*, *25*, 187–212. doi:10.1016/0093-934X(85)90080-X

Miller, J. L., Dexter, E. R., & Pickard, K. A. (1984). Influence of speaking rate and lexical status on word identification. *Journal of the Acoustical Society of America*, *76*, 589.

Mirman, D., McClelland, J. L., & Holt, L. L. (2006). An interactive Hebbian account of lexically guided tuning of speech perception. *Psychonomic Bulletin and Review*, *13*, 958–965. doi:10.3758/BF03213909

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*, 165–178.

Nestor, A., Behrmann, M., & Plaut, D. C. (2013). The neural basis of visual word form processing: A multivariate investigation. *Cerebral Cortex*, *23*, 1673–1684. doi:10.1093/cercor/bhs158

Noble, K., Glosser, G., & Grossman, M. (2000). Oral reading in dementia. *Brain and Language*, *74*(1), 48–69.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*, 299–325, 363–370. doi:10.1017/S0140525X00003241

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word-superiority effect. *Psychological Review*, *89*, 573–594.

Patterson, K., Ralph, M. A. L., Jefferies, E., Woollams, A., Jones, R., Hodges, J. R., & Rogers, T. T. (2006). 'Presemantic' cognition in semantic dementia: Six deficits in search of an explanation. *Journal of Cognitive Neuroscience*, *18*, 169–183.

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*, *61*(2), 106–151. doi:10.1016/j.cogpsych.2010.04.001

Plaut, D. C. (1996). Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language*, *52*(1), 25–82.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, *12*, 765–805.

Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, *19*, 603–639.

Plaut, D. C., & Behrmann, M. (2011). Complementary neural representations for faces and words: A computational exploration. *Cognitive Neuropsychology*, *28*, 251–275.

Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language* (pp. 381–415). Mahwah, NJ: Erlbaum.

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. doi:10.1037/0033-295X.103.1.56

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*, 377–500.

Powell, D., Plaut, D., & Funnell, E. (2006). Does the PMSP connectionist model of single word reading learn to read in the same way as a child? *Journal of Research in Reading*, *29*, 229–250.

Price, C. J., & Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*, 246–253.

Roberts, D. J., Lambon Ralph, M. A., & Woollams, A. M. (2010). When does less yield more? The impact of severity upon implicit recognition in pure alexia. *Neuropsychologia*, *48*, 2437–2446.

Roberts, D. J., Woollams, A. M., Kim, E., Beeson, P. M., Rapcsak, S. Z., & Lambon Ralph, M. A. (2013). Efficient visual object and word recognition relies on high spatial frequency coding in the left posterior fusiform gyrus: Evidence from a case-series of patients with ventral occipito-temporal cortex damage. *Cerebral Cortex*, *23*, 2568–2580.

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*, 205–235.

Rogers, T. T., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2004). Natural selection: The impact of semantic impairment on lexical and object decision. *Cognitive Neuropsychology*, *21*, 331–352.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, *89*(1), 60–94.

Samuel, A. G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General*, *125*(1), 28–51.

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

Seidenberg, M. S., & McClelland, J. L. (1990). More words but still no lexicon: Reply to besner et al. (1990). *Psychological Review*, *97*, 447–452.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*, 151–218.

Shibahara, N., Zorzi, M., Hill, M. P., Wydell, T., & Butterworth, B. (2003). Semantic effects in word naming: Evidence from English and Japanese Kanji. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *56A*, 263–286. doi:10.1080/02724980244000369

Sibley, D. E., Kello, C. T., Plaut, D. C., & Elman, J. L. (2008). Large-scale modeling of wordform learning and representation. *Cognitive Science: A Multidisciplinary Journal*, *32*, 741–754. doi:10.1080/03640210802066964

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience*, *32*, 8443–8453.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1140–1154.

Strain, E., Patterson, K., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling-sound consistency. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*, 207–214.

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin*, *139*, 766–791.

Turkeltaub, P. E., Goldberg, E. M., Postman-Caucheteux, W. A., Palovcak, M., Quinn, C., Cantor, C., & Coslett, H. B. (2014). Alexia due to ischemic stroke of the visual word form area. *Neurocase*, *20*, 230–235.

Twomey, T., Kawabata Duncan, K. J., Price, C. J., & Devlin, J. T. (2011). Top-down modulation of ventral occipito-temporal responses during visual word recognition. *NeuroImage*, *55*, 1242–1251.

Tyler, L. K., Voice, J. K., & Moss, H. E. (2000). The interaction of meaning and sound in spoken word recognition. *Psychonomic Bulletin and Review*, *7*, 320–326.

Ueno, T., Saito, S., Rogers, T., & Lambon Ralph, M. (2011). Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*, *72*, 385–396. doi:10.1016/j.neuron.2011.09.013

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., … Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, *30*, 1414–1432. doi:10.1016/j.neuroimage.2005.11.002

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. doi:10.1016/j.neuron.2007.05.031

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*, 325–329.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*, 374–408. doi:10.1006/jmla.1998.2618

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, *52*, 193–204.

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *68*, 306–311. doi:10.1006/brln.1999.2116

Vogel, A. C., Petersen, S. E., & Schlaggar, B. L. (2013). Matching is not naming: A direct comparison of lexical manipulations in explicit and implicit reading tasks. *Human Brain Mapping*, *34*, 2425–2438.

Welbourne, S. R., & Lambon Ralph, M. A. (2005). Using computational, parallel distributed processing networks to model rehabilitation in patients with acquired dyslexia: An initial investigation. *Aphasiology*, *19*, 789–806. doi:10.1080/026870305002688110

Welbourne, S. R., & Lambon Ralph, M. A. (2007). Using parallel distributed processing models to simulate phonological dyslexia: The key role of plasticity-related recovery. *Journal of Cognitive Neuroscience*, *19*, 1125–1139. doi:10.1080/02687030500268811

Woollams, A. M. (2005). Imageability and ambiguity effects in speeded naming: Convergence and divergence. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*, 878–890. doi:10.1037/0278-7393.31.5.878

Woollams, A. M., Hoffman, P., Roberts, D. J., Lambon Ralph, M. A., & Patterson, K. (2014). What lies beneath: A comparison of reading aloud in pure alexia and semantic dementia. *Cognitive Neuropsychology*, *31*, 461–481.

Woollams, A. M., Ralph, M. A. L., Plaut, D. C., & Patterson, K. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review*, *114*, 316–339. doi:10.1037/0033-295X.114.2.316

Woollams, A. M., Silani, G., Okada, K., Patterson, K., & Price, C. J. (2011). Word or word-like? Dissociating orthographic typicality from lexicality in the left occipito-temporal cortex. *Journal of Cognitive Neuroscience*, *23*, 992–1002. doi:10.1016/j.neuroimage.2006.07.029

Wurm, L. H., Vakoch, D. A., & Seaman, S. R. (2004). Recognition of spoken words: Semantic effects in lexical access. *Language and Speech*, *47*, 175–204. doi:10.1177/00238309040470020401

Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*, 1634.