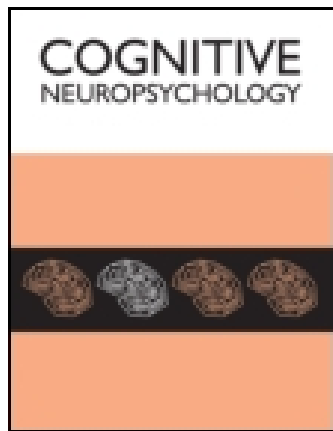


This article was downloaded by: [University of Waikato]

On: 11 July 2014, At: 06:25

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office:
Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Cognitive Neuropsychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pcgn20>

Representation of letter position in single-word reading: Evidence from acquired dyslexia

Michael McCloskey^a, Simon Fischer-Baum^b & Teresa Schubert^a

^a Department of Cognitive Science, Johns Hopkins University, Baltimore, MD, USA

^b Department of Psychology, Rice University, Houston, TX, USA

Published online: 10 Feb 2014.

To cite this article: Michael McCloskey, Simon Fischer-Baum & Teresa Schubert (2013) Representation of letter position in single-word reading: Evidence from acquired dyslexia, *Cognitive Neuropsychology*, 30:6, 396-428, DOI: [10.1080/02643294.2014.880675](https://doi.org/10.1080/02643294.2014.880675)

To link to this article: <http://dx.doi.org/10.1080/02643294.2014.880675>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Representation of letter position in single-word reading: Evidence from acquired dyslexia

Michael McCloskey¹, Simon Fischer-Baum², and Teresa Schubert¹

¹Department of Cognitive Science, Johns Hopkins University, Baltimore, MD, USA

²Department of Psychology, Rice University, Houston, TX, USA

Visual word recognition requires information about the positions as well as the identities of the letters in a word. This study addresses representation of letter position at prelexical levels of the word recognition process. We present evidence from an acquired dyslexic patient, L.H.D., who perseverates letters in single-word reading tasks: Far more often than expected by chance, L.H.D.'s reading responses include letters from preceding responses (e.g., SAILOR read as SAILOG immediately after FLAG was read correctly). Analyses carried out over two large data sets compared the positions of perseverated letters (e.g., the G in SAILOG) with the positions of the corresponding "source" letters (e.g., the G in FLAG). The analyses assessed the extent to which the perseverations preserved source position as defined by various theories of letter position representation. The results provided strong evidence for graded both-edges position representations, in which the position of each letter is encoded coarsely relative to both the beginning and the end of the word. Alternative position representation schemes, including letter-context and orthosyllabic schemes, were not supported.

Keywords: Alexia; Dyslexia; Reading; Serial order; Letter position; Word recognition.

Visual word recognition processes must consider not only the identities but also the positions of letters (e.g., CAT vs. ACT). A letter's position might be encoded on the basis of its distance from the beginning (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) or the midpoint (Caramazza & Hillis, 1990) of a word, its role in an orthographic syllable structure (e.g., Perry, Ziegler, & Zorzi, 2010), its context of surrounding letters (e.g., Grainger, Granier, Farioli, Van

Assche, & van Heuven, 2006; Whitney, 2001), or according to some other representational scheme. For example, the E in PERSON could be represented as the second letter from the beginning of the word ($B+2$), the letter in the first nucleus position (N1) of the first orthographic syllable, or the letter appearing after P and before R (P_R).

Letter position representation has received a great deal of attention in the reading literature over the

Correspondence should be addressed to Michael McCloskey, Department of Cognitive Science, Krieger Hall, Johns Hopkins University, Baltimore, MD 21218, USA (E-mail: michael.mccloskey@jhu.edu).

We are grateful to L.H.D. for her cooperation and patience over many sessions of testing.

The second author was supported by a Beckman Fellowship, and the third author was supported by a National Science Foundation Integrative Graduate Education and Research Traineeship grant to the Johns Hopkins Department of Cognitive Science.

past decade, and multiple computational theories have been proposed (e.g., Adelman, 2011; Cox, Kachergis, Recchia, & Jones, 2011; Davis, 2010b; Gomez, Ratcliff, & Perea, 2008; Grainger et al., 2006; Hannagan, Dupoux, & Christophe, 2011; Hannagan & Grainger, 2012; Norris & Kinoshita, 2012; Whitney, 2001). These theories differ on a number of dimensions, including their assumptions about how letter position is represented. A number of recent studies have attempted to adjudicate among the competing theories, typically on the basis of results from masked priming procedures (e.g., Davis, 2010a; Davis & Bowers, 2004, 2006; Gomez et al., 2008; Grainger et al., 2006; Kinoshita & Norris, 2009, 2013; Perea & Lupker, 2003; Whitney, Bertrand, & Grainger, 2012). However, much of the available masked priming evidence is compatible with multiple theories (see, e.g., Fischer-Baum, Charny, & McCloskey, 2011; Norris, Kinoshita, & van Casteren, 2010), and masked priming experiments with the potential to distinguish among competing theories have yielded inconsistent results (e.g., Davis, 2010b; Kinoshita & Norris, 2013; Whitney et al., 2012). These points suggest that progress in theory development and evaluation may require additional forms of empirical evidence.

In the domain of spelling Fischer-Baum, McCloskey, and Rapp (2010) used letter perseveration errors to probe representation of letter position. In spelling-to-dictation tasks, two acquired dysgraphic patients (C.M. and L.S.S.) perseverated letters from prior spelling responses into subsequent responses. For example, after writing FARM, C.M. intruded an A into his response on the next trial, writing BLUNT as BLANT. Significantly more often than expected by chance, perseverated letters occupied the same position in the prior and subsequent responses, when position was defined by a beginning- or end-based representational scheme. For example, the intruded A in BLANT appears in the same end-based position ($E-3$) as the A in the preceding response FARM. Perseverated letters that did not appear in exactly matching beginning- or end-based positions were more likely than expected by

chance to occupy approximately matching positions—that is, positions immediately adjacent to the exactly matching positions. Representational schemes other than the beginning- or end-based schemes made no contribution to accounting for the positions occupied by perseverated letters.

Fischer-Baum et al. (2010) took these results as evidence for both-edges representation of letter position in spelling. A both-edges position scheme encodes a letter's position relative to both the beginning and the end of the word, so that in PERSON, for example, the E has position representations $B+2$ and $E-5$. To explain the systematic tendency toward approximate as well as exact position matches, Fischer-Baum proposed that the position representations are graded, such that nearby positions (e.g., $B+2$ and $B+3$) have more similar representations than more distant positions (e.g., $B+2$ and $B+6$).

Applying similar logic, Fischer-Baum et al. (2011) used the illusory word paradigm (e.g., Davis & Bowers, 2004) to examine letter position representations in reading. Target-distractor word pairs (e.g., TUBE-TANGLE) were presented briefly, and participants reported the target word in each pair. Analyses focused on letter migration errors (e.g., TUBE reported as TUNE), in which a letter from the distractor (TANGLE) intruded into the response to the target. Fischer-Baum et al. found that letters systematically maintained their beginning- or end-based position when migrating from distractors to targets. For example, the distractors TANGLE and THRONE were both likely to elicit the TUBE \rightarrow TUNE migration error. In the case of TANGLE, the N migrates from position $B+3$ of the distractor to position $B+3$ of the target, and for THRONE the migration is from distractor position $E-2$ to target position $E-2$. Migrations that preserved approximate but not exact beginning- or end-based position (e.g., TUBE \rightarrow TUNE with the distractor TRANCE) were less frequent, but occurred at rates significantly above baseline. Fischer-Baum et al. (2011) interpreted these results as evidence for graded both-edges position representations in reading.

In the present study, we apply the methods from the Fischer-Baum et al. (2010) spelling study to an acquired dyslexic patient, L.H.D., who makes letter perseveration errors in single-word reading tasks. Extensive analyses of two very large data sets strongly support a graded both-edges scheme for representing letter position in early visual word recognition processes.

SINGLE-WORD READING

As illustrated in Figure 1, we assume that when a word is presented, prelexical reading processes first compute the identity and position of each letter (e.g., Caramazza & Hillis, 1990; Schubert & McCloskey, 2014). At the earliest level of processing, visual processes generate a visual feature map that represents low-level visual features (e.g., oriented bars). At the next level of representation, which we refer to as the character shape level, the stimulus is segmented into separate characters, and the shape of each character is represented as a configuration of visual features. The character shape level represents stimulus shapes that have not yet been recognized: No determination has been made that a shape corresponds to a particular letter, or even that it is a familiar form. The character shape representations activate stored representations at an allograph level that represents the various distinct visual forms of each letter—for example, a, α, A, and so on, for the letter A (e.g., Miozzo & Caramazza, 1998; Rapp & Caramazza, 1989; Schubert & McCloskey, 2014). Finally, the allograph representations activate representations at an abstract letter identity level. This level of representation abstracts away from visual properties of letters, representing for each position in the word only the letter identity.

The abstract representations of letter identity and position constitute the input to lexical and sublexical reading processes. In lexical processing, a representation is activated in an orthographic lexicon, leading in turn to activation of semantic and phonological representations. In sublexical processing, a plausible pronunciation is assembled

on the basis of grapheme-to-phoneme correspondence rules (e.g., the letter B usually corresponds to the phoneme /b/).

As we shall see, L.H.D.'s errors arise at the prelexical levels of the reading system, in the process of mapping allograph representations onto abstract letter identity representations. Hence, we argue that her errors speak to the representation of letter position at the allograph and abstract letter identity levels.

REPRESENTING LETTER POSITION IN READING

We evaluate a broad range of potential schemes for representing letter position in reading, reflecting the spectrum of schemes that have been proposed in theories of visual word recognition. The schemes may be classified according to the stance they adopt on two principal issues. First, relative to what reference point(s) is the position of a letter specified? Second, are letter position representations discrete, such that each position has a representation that is maximally distinct from the representation for every other position; or are the position representations graded, such that nearby positions have more similar representations than distant positions?

Reference points

Assumptions about reference points for representing letter position fall into three broad classes. In *extrinsic-anchor* schemes, a letter's position is determined by distance (in number of letters) and direction from one or more anchor points that do not depend on the identities of the letters in the word. In *letter-context* schemes, a letter's position is defined relative to its specific surrounding letters. In *syllabic* schemes, a letter's position is defined relative to the syllable in which it appears and its role within that syllable.

Extrinsic-anchor schemes

We consider five extrinsic-anchor schemes: beginning-based, end-based, midpoint-based,

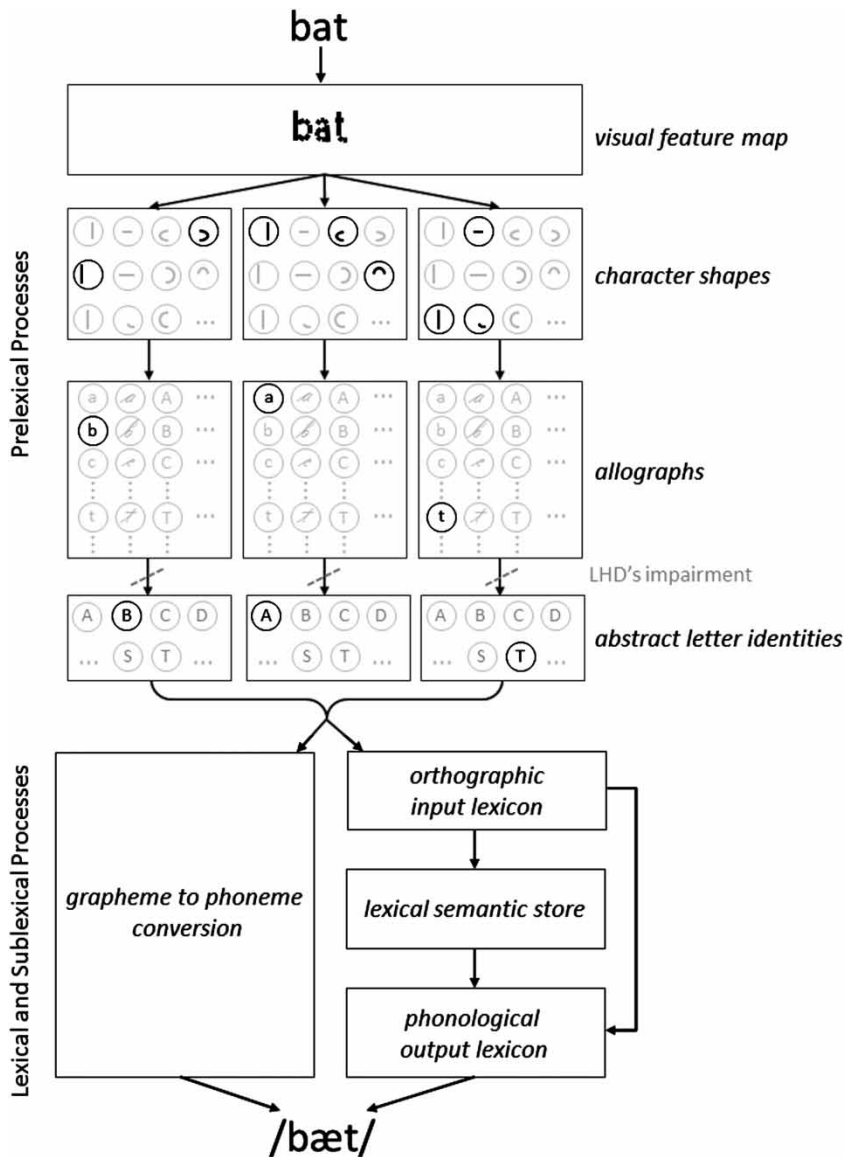


Figure 1. Schematic depiction of cognitive single-word reading system. Dashed line indicates the functional locus of L.H.D.'s reading deficit.

closer-edge, and both-edges. Figure 2 illustrates these schemes using a representational format introduced in McClelland and Rumelhart's (1981) interactive activation reading model. Each position in a word is represented by a distinct pool of 26 letter units. The six pools in each panel of Figure 2 can represent words of up to six letters.

Figure 2a illustrates a beginning-based representation for the word FACE. The F node is activated in the $B+1$ pool, the A node in the $B+2$ pool, and so forth. Beginning-based schemes have been adopted in several computational models of reading (e.g., Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart et al., 2001; Ellis, Flude, & Young, 1987; Seymour, 1979).

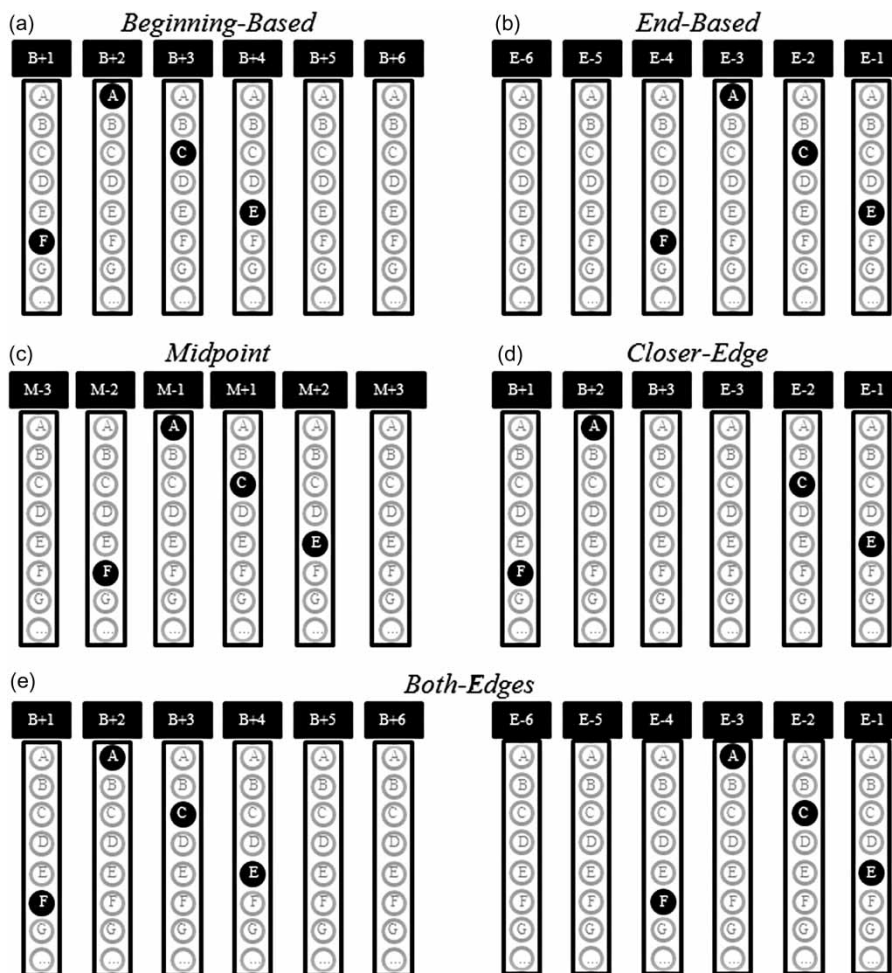


Figure 2. Representations of the letter sequence for the word *FACE*, according to several extrinsic-anchor position schemes.

The end-based scheme, depicted in Figure 2b, defines position relative to the end of the word. Although the end-based scheme has not been proposed as the sole basis for representing letter position in either reading or spelling, this scheme constitutes one component of the both-edges scheme.

The midpoint scheme, shown in Figure 2c, defines position relative to the centre of the word. In *FACE* the *F* occupies position $M-2$

(two positions prior to the midpoint), whereas the *E* is in position $M+2$.¹ Caramazza and Hillis (1990) proposed a midpoint scheme for representing letter position in reading and spelling.

The closer-edge scheme, proposed by Jacobs, Rey, Ziegler, and Grainger (1998) in a theory of visual word recognition, encodes letter position relative to the beginning or end of the word,

¹ Three versions of the midpoint scheme can be formulated, depending upon how a position is assigned to the central letter in words with an odd number of letters. Using the *A* in *STAMP* for purposes of illustration, this letter could be assigned to (a) position *M*, a position that does not occur for words of even length; (b) position $M+1$, or (c) position $M-1$. We considered all three of these variants in our analyses and report results for the most successful variant.

whichever is closer (see Figure 2d). For FACE, the F and A are represented relative to the beginning of the word and so have position representations $B+1$ and $B+2$, respectively. The C and E, however, are represented relative to the end of the word and so have position representations $E-2$ and $E-1$, respectively.²

Finally, the both-edges scheme assigns both beginning- and end-based position representations to each letter in a word. For example, in Figure 2e the F in FACE is represented by activating the F unit in both $B+1$ and $E-4$ pools. As we have seen, Fischer-Baum and colleagues have proposed graded both-edges position representations for reading (Fischer-Baum et al., 2011) and spelling (Fischer-Baum et al., 2010). Friedmann, Kerbel, and Shvimer (2010) have also proposed both-edges representations for reading on the basis of results from individuals with developmental dyslexia (although the evidence they report speaks only to end-based representations). In addition, Glasspool and colleagues posit both-edges representations in their competitive queuing models of spelling (Glasspool, 1998; Glasspool & Houghton, 2005; Houghton, Glasspool, & Shallice, 1994).

Letter-context schemes

In letter-context schemes, a letter's position is represented with respect to the identities of other letters in the word. Various theories of word recognition (Dehaene, Cohen, Sigman, & Vinckier, 2005; Grainger & Van Heuven, 2003; Mozer, 1987; Seidenberg & McClelland, 1989; Whitney, 2001) posit some form of letter-context coding. Figure 3a illustrates a simple letter-context scheme, in which the position of a letter is defined by the immediately preceding letter. Each pool of letter units corresponds to a preceding letter context. For example, the A_ pool is used to represent any letter that is preceded by an A. The C in FACE is therefore represented by activating the C unit in this pool. The #_ pool represents the position preceded by the left word

boundary, and hence the F in FACE is represented by a unit in this pool.

This preceding-letter scheme is a bigram representation, because the word is represented as a set of bigrams (e.g., #F, FA, AC, CE, E# for FACE). Alternatives to this simple bigram scheme have been proposed. In trigram schemes (Brown & Loosemore, 1994; Seidenberg & McClelland, 1989) the position of a letter is defined jointly by the immediately preceding and immediately following letters, so that the C in FACE, for example, could be represented by a unit in an A_E pool that represents letters preceded by A and followed by E. In open bigram schemes (Dehaene et al., 2005; Grainger & Van Heuven, 2003; Whitney, 2001) the position of a letter is defined not only by immediately preceding and/or following letters, but also by more distant letters. For example, representing the C in FACE might involve activating the C unit in an F_ pool as well as in an A_ pool, to indicate that the C was preceded by both an F and an A.

Orthosyllabic schemes

Many researchers have argued that letter-sequence representations implicated in reading or spelling include information about orthographic syllable structure (e.g., Badecker, 1996; Caramazza & Miceli, 1990; Olson & Caramazza, 2004; Prinzmetal, Treiman, & Rho, 1986; Rapp, 1992; Spoehr & Smith, 1973; Taft, 1979). This claim does not necessarily imply that syllabic information plays a role in representing letter position. Nevertheless, letter position could be represented through a syllabic scheme, as assumed in several simulation models of reading aloud (Harm & Seidenberg, 1999; Plaut, McClelland, Seidenberg, & Patterson, 1996; Zorzi, Houghton, & Butterworth, 1998).

Figure 3b shows a simplified orthosyllabic position representation for the word FACE, which we assume to be bisyllabic orthographically (FA-CE) although monosyllabic phonologically. The F is

² Two variants of the closer-edge scheme may be defined, differing in the position assigned to the central letter in words with an odd number of letters. In these variants the A in STAMP is assigned either to position $B+3$ or to position $E-3$. Both variants were considered in our analyses; reported results are for the more successful variant.

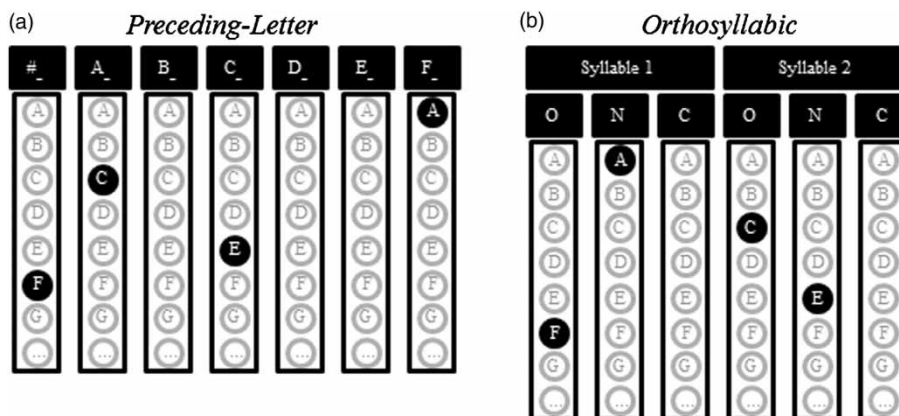


Figure 3. Examples of letter context and orthosyllabic representations for the word FACE.

represented by a unit in a Syllable 1 onset pool, and the A by a unit in a Syllable 1 nucleus pool. Similarly, the C and E are represented by units in Syllable 2 onset and nucleus pools, respectively. Orthographic onsets, nuclei, and codas may contain multiple letters (as in BREATH); as a consequence, multiple onset, nucleus, and coda pools are needed to fully specify letter position orthosyllabically (e.g., Onset 1, Onset 2, Onset 3).

Discrete versus graded position representations

Schemes for representing letter position vary not only in assumptions about reference points, but also in whether the position representations are assumed to be discrete or graded. In a discrete scheme, two position representations are either the same or different; representations of nearby positions (e.g., $B+1$ and $B+2$) are no more similar than representations of distant positions (e.g., $B+1$ and $B+5$). In a graded scheme, position representations have a similarity structure, such that the closer two positions are, the more similar their representations (e.g., Gomez et al., 2008). Figure 4 illustrates one possible instantiation of a graded beginning-based scheme. Each letter in a word is represented by activation in more than one pool of units, and each pool contributes to representing more than one of the word's letters. For example, the A in FACE is represented by

activating the A unit strongly in the $B+2$ pool, somewhat less strongly in the $B+1$ and $B+3$ pools, and so forth. In this scheme, the representation of each letter is distributed across several pools of units, and each pool has a gradient of activation across letters. Thus, in Figure 4 the $B+1$ pool shows a gradient of activation in which the F unit is activated mostly strongly, the A unit less strongly, and so forth (e.g., Houghton et al., 1994). As a consequence, the F and A have more similar position representations than the F and the E. Discrete representations do not have this property (see, e.g., Figure 2a).

The distinction between discrete and graded position representations applies to orthosyllabic and letter-context schemes as well as to extrinsic-anchor schemes. Some syllabic schemes (e.g., Harm & Seidenberg, 1999; Plaut et al., 1996) assume graded representations in which positions within a syllabic role (e.g., first vs. second onset position) have more similar representations than positions for different syllabic roles (e.g., first onset vs. first coda position). Other syllabic schemes (e.g., Zorzi et al., 1998) assume discrete position representations with no similarity structure. For letter-context schemes, simple bigram and trigram schemes are discrete, with no similarity structure among letter positions. In contrast, at least some open-bigram schemes are graded in the sense that nearby positions in a word share some preceding- and/or following-letter

Graded Beginning-Based

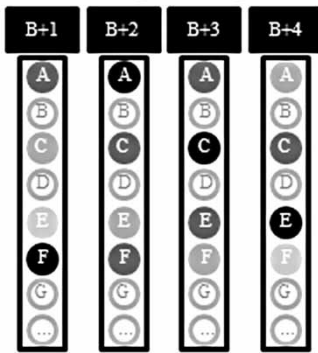


Figure 4. Graded beginning-based representation for the word *FACE*.

representations, whereas distant positions do not. For example, given an open-bigram scheme that represents a letter's position relative to the three preceding and following letters (e.g., Grainger & Van Heuven, 2003), the E and C in *SECTION* share the preceding-letter representation S_{-} (indicating that both are preceded by S) and the following-letter representations $_{-}T$ and $_{-}I$ (indicating that both are followed by T and I). Distant letters such as the E and N, however, do not share any position representations.

CASE HISTORY

L.H.D. is a 72-year-old woman with a Master's degree who worked as an executive in the banking industry before her retirement. In February 2007 she suffered a ruptured left posterior cerebral artery aneurysm and underwent surgery to clip the aneurysm. Magnetic resonance imaging (MRI) demonstrated left-hemisphere damage extending from the posterior occipital lobe to the medial anterior temporal lobe, affecting striate and extrastriate visual areas, the fusiform and parahippocampal gyri, and the hippocampus. Ophthalmological examination revealed a right homonymous hemianopia without macular sparing. (For additional case history information, MRI images, and neuropsychological assessment results, see Schubert & McCloskey, 2014)

Table 1. Examples of L.H.D.'s test-session reading errors for word and nonword stimuli

Target	Response
carpet	CARPE <u>F</u>
person	PERCO <u>B</u>
magazine	MAGAZI <u>E</u>
carriage	CARRI <u>E</u> Q <u>F</u>
bishop	DISHOP
dust	DUS <u>K</u>
cest	CES <u>Y</u>
raradil	RARA <u>F</u> IL

Immediately after her surgery, L.H.D. was severely aphasic, but many of her language abilities recovered substantially. She was left with moderately severe word-finding difficulty. However, her principal complaint was severe difficulty in reading. A highly intelligent woman, she had been a voracious reader, but now struggled to identify nearly every word. Table 1 presents examples of L.H.D.'s reading errors in test sessions during the present study. In these and subsequent examples, we use lower case for stimuli (e.g., *carpet*) and upper case for written transcriptions of L.H.D.'s responses (e.g., CARPEF), regardless of the case in which the stimulus was presented. As the examples illustrate, the vast majority of L.H.D.'s errors took the form of nonword responses that differed from the corresponding targets by the substitution of one or (occasionally) more letters (e.g., *carpet* → CARPEF, *person* → PERCOB).

L.H.D.'s reading accuracy was not affected by the length of the stimulus, but her reading latencies increased with length. Serial position functions showed that L.H.D.'s error rate was low for letters near the beginning of a stimulus, with higher error rates toward the end.

FUNCTIONAL LOCUS OF IMPAIRMENT

Schubert and McCloskey (2014) report extensive testing pointing to the conclusion that L.H.D.'s

reading errors arise from an impairment in visual letter identification. More specifically, L.H.D.'s letter processing is intact at visual feature map, character shape, and allograph levels, as indicated by her normal performance in making same-different judgements for within-case letter strings (e.g., DFA–DFA), in copying pseudoletters, and in discriminating real letters from pseudoletters. However, she is impaired in mapping allographs onto abstract letter identities, as evidenced by impaired performance in cross-case same-different judgements (e.g., RMH–rmh), in transcoding between upper- and lower-case letters, and in naming letters presented individually or in strings. According to this account, L.H.D. misread *bishop* as DISHOP despite activating the correct allograph representation for the stimulus letter *b* (i.e., the lower-case print *b* allograph), because her impairment in mapping allographs to abstract letter identities led to abnormally low activation of the correct letter identity *B*, with the result that the incorrect letter *D* was selected instead.

Many of L.H.D.'s letter identification errors take the form of letter perseverations, in which a letter from a prior response is substituted for the target letter. For example, L.H.D.'s misreading of *bishop* as DISHOP occurred immediately after she had read *dust* correctly. The letter perseverations presumably stem from residual activation of letter representations. Thus, when the *B* abstract letter identity representation was poorly activated during the processing of *bishop*, *D* may have been selected instead because the *D* representation retained some activation from prior processing of *dust*. Note that the residual activation may be entirely normal, giving rise to errors only when the correct letter identity representation has abnormally low activation.

EXPERIMENTAL STUDY

Testing for the present study began in October 2010. Over the next 11 months L.H.D. showed some gradual improvement in single-word reading, although the qualitative pattern (error

types, effects of variables) remained stable. Then, in September 2011 she was injured in a car accident. Although she did not strike her head, she suffered a left-hemisphere subdural bleed and underwent surgery to evacuate the resulting hematoma. No new brain damage was evident on MRI. When our testing resumed in November 2011, L.H.D.'s reading performance was qualitatively the same as before the accident, but her overall accuracy was approximately 10–15 percentage points lower. Some improvement was evident over the next several months, and by the end of the study her reading accuracy had recovered to nearly the pre-accident level. Roughly half of the data for the present study were collected prior to L.H.D.'s car accident, and approximately half after the accident. Analyses performed separately on pre- and post-accident data yielded the same pattern of results, and hence we report analyses on data pooled across the pre- and post-accident periods.

Test session data

Across 60 testing sessions carried out between November 2010 and July 2012, L.H.D. read aloud 5915 words and 629 nonwords. Stimuli were presented individually on a computer monitor in 36-point Courier New or Arial font. The majority of the stimuli (4005/6544) were presented in lower case. Most of the remaining items (2160) were presented in upper case, and a small number (379) were presented in mixed case (e.g., BrANcH). Because of L.H.D.'s right homonymous hemianopia, stimuli were displayed to the left of a central fixation point. The tasks were untimed, and each stimulus remained in view until L.H.D. had completed her reading response. She was encouraged to fixate the central point prior to stimulus presentation, but was not required to maintain fixation while attempting to read the item. In transcribing L.H.D.'s spoken reading responses, we gave her the benefit of the doubt by assuming the minimal error consistent with the response. For example, the response /bleɪn/ to the stimulus *brain* was transcribed as BLAIN rather than BLANE, on the assumption

that the error reflected a single letter misidentification ($r \rightarrow l$), rather than a complex combination of letter misidentification and transposition. On trials in which one or more letter identifications could not plausibly be determined from L.H.D.'s response, she was asked to report the letters in question. For example, in her misreading of *tiny* as /tɪnz/ the final letter might have been misidentified as either S or Z, and so L.H.D. was asked to report that letter.

L.H.D.'s whole-word reading accuracy was 62% for word stimuli (3658/5915) and 44% for nonwords (277/629), $\chi^2(1, N = 6544) = 75.2, p < .001$. Accuracy did not differ between lower-case (61%) and upper-case (59%) stimuli, $\chi^2(1, N = 6165) = 2.62, p > .10$.

At-home practice data

L.H.D. also completed at-home practice sessions from July 2011 through July 2012. In these sessions, words were presented individually on a computer monitor as in the test sessions conducted in our lab. However, stimulus presentation time was limited, with exposure durations ranging across sessions from 250 to 2000 ms. L.H.D. read the stimulus silently and then, after stimulus offset, wrote her response on an answer sheet. Across the 13-month period she provided written responses to a remarkably large number of stimuli: 21,971 words and 2447 nonwords. Nearly all of the stimuli (23,763/24,418) were presented in lower case.

L.H.D.'s accuracy in the home-practice task was 69% (15,153/21,971) for words and 45% for nonwords (1106/2447), $\chi^2(1, N = 24,418) = 558.9, p < .001$. Table 2 presents examples of her errors. The home-practice task might be viewed as a delayed-copy task rather than a reading task. Note, however, that L.H.D. could not have performed the task by copying the visual forms of the letters without identifying them, because she wrote her responses in cursive script. Furthermore, by any measure L.H.D.'s home-practice performance was virtually indistinguishable from her performance in the reading-aloud tasks carried out during the testing sessions

Table 2. Examples of L.H.D.'s home-practice errors for word and nonword stimuli

Target	Response
group	GROUF
borrow	DORREW
family	FAMITY
weapon	WEATOR
alone	ALOVE
dust	DUSK
dac	DAS
pesh	PESN

in our lab. Accordingly, we assume that the early stages of the reading process, including letter identification and (in the case of words) activation of a lexical representation, were shared by the spoken-response task used in the lab testing sessions and the written-response task from the at-home practice sessions. In subsequent analyses we use both test-session and home-practice data to infer letter position representations.

LETTER PERSEVERATION ANALYSIS

In both the test sessions and the at-home practice sessions the vast majority of L.H.D.'s reading errors contained intruded letters—that is, letters not present in the target word. Nearly all of the intrusions took the form of letter substitutions, as in *sailor* read as SAILOG, which contains an intruded G. Occasionally L.H.D. produced intrusions in the form of letter insertions, as in *steady* → STREADY.

Very often the intruded letters were present in one or more of the preceding responses, suggesting that at least some of the intrusions were perseverations from prior responses. Table 3 presents an illustrative sequence of trials: the trial on which L.H.D. read *sailor* as SAILOG, labelled I (for Intrusion trial), and the five immediately preceding trials (I-1, I-2, etc.). The intruded G in SAILOG was present in the responses for

Table 3. *A sequence of test-session reading trials*

<i>Trial</i>	<i>Target</i>	<i>Response</i>
I-5	vine	VINE
I-4	goat	GOAN
I-3	rib	RIB
I-2	college	COLLEGE
I-1	flag	FLAG
I	sailor	SAILOG

trials I-1 (FLAG), I-2 (COLLEGE), and I-4 (GOAN). Conceivably, then, the G in SAILOG was a perseveration from one or more of these preceding responses. However, an intruded letter could be present in a prior response simply by chance. To determine whether L.H.D.'s letter intrusions included true perseverations from prior responses, we evaluated whether intruded letters were present in prior responses more often than expected by chance.

Using methods developed in our earlier work on perseveration errors in spelling (Fischer-Baum et al., 2010; McCloskey, Macaruso, & Rapp, 2006; see also Cohen & Dehaene, 1998, for similar methods), we conducted two letter perseveration analyses, one for the 2599 intruded letters in the test-session data, and one for the 8480 intrusions in the home-practice data. The analysis for each data set evaluated whether an intruded letter was more likely than expected by chance to appear in each of the five trials immediately preceding the intrusion error (trials I-1 through I-5). For each intruded letter a computer program first tabulated whether the letter was present in the response on trial I-1. For example, the intruded G in SAILOG was present in the I-1 response FLAG. From these results, the program calculated the proportion of intruded letters that were present in the corresponding I-1 responses. For example, the analysis for the test-session data revealed that 1376 of the 2599 intruded letters (a proportion of .53) appeared in the immediately preceding response.

Next the program computed the proportion of intruded letters that did not appear in the I-1 response but did appear in the response on trial I-2. For

example, in L.H.D.'s reading error *gun* → GUK, the intruded K was not present in the I-1 response WATER, but did appear in the I-2 response ANKLE. The program then computed the proportion of intruded letters that did not appear in the I-1 or I-2 response, but did appear in the response on trial I-3, and so forth through trial I-5.

The results are shown in the solid curves of Figure 5a (test-session data) and 5b (home-practice data). The .57 value plotted for trial I-1 in Figure 5b reflects the fact that 4831 of L.H.D.'s 8480 intruded letters in the home-practice data (57%) were present in the response immediately prior to the error. The .33 value for trial I-2 indicates that 33% of the intruded letters that did not appear in the I-1 response were present in the response on trial I-2.

The perseveration analysis program also estimated the likelihood of an intruded letter being present by chance in the I-1 through I-5 responses, using methods described in McCloskey et al. (2006) and Fischer-Baum et al. (2010). These methods also provide a basis for evaluating the statistical significance of differences between observed and chance proportions. The dashed curves in Figure 5 present the mean estimated chance proportions for trials I-1 through I-5 in each data set. For both the test-session and home-practice data the observed proportion of preceding responses containing the intruded letter was reliably higher than the proportion expected by chance for trials I-1 through I-3 ($p < .0001$), but not for trials I-4 and I-5 ($p > .05$). These results demonstrate for both data sets that at least some of L.H.D.'s letter intrusions are true perseverations and further indicate that perseverations may originate from responses as far back as three trials prior to the intrusion trial. We refer to this range of preceding trials as the perseveration window.

LETTER POSITION ANALYSES

We next conducted analyses using L.H.D.'s letter perseverations as a basis for inferences about the representation of letter position at the prelexical levels of the reading system where the errors

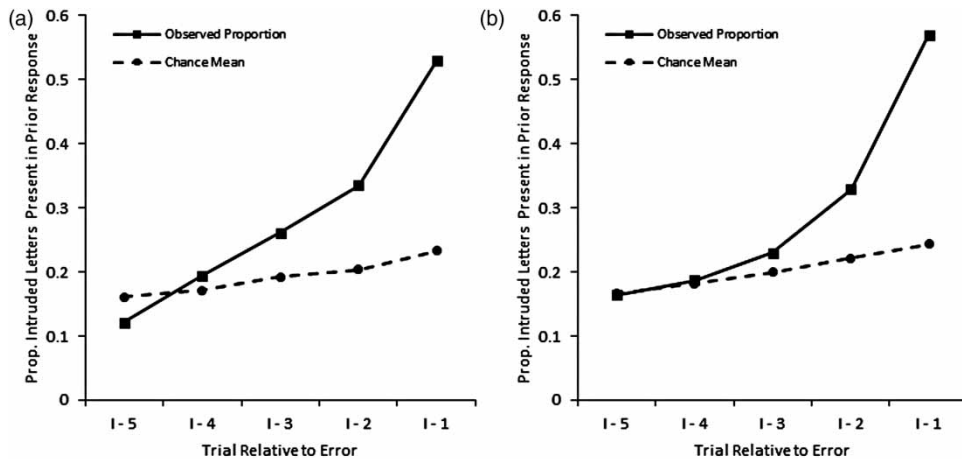


Figure 5. Perseveration analysis results for (a) the test-session and (b) the home-practice data sets. The solid line in each panel shows the proportion of intruded letters present in the response on the five immediately preceding trials. The dashed line is the mean proportion expected by chance.

arise. Schubert and McCloskey (2014) demonstrated that residual activation at both allograph and abstract letter identity levels play a role in the perseveration errors, and they also presented evidence that letter position is represented in the same manner at both of these levels. Accordingly, our letter position analyses speak to representation of letter position at allograph and abstract letter identity levels of the reading system.

In describing these analyses, we refer to a response with a perseverated letter as a *perseveration response* and the prior response from which the letter was perseverated as the *source response*. For example, assuming that the G in the error *sailor* → SAILOG was a perseveration from the immediately preceding response FLAG, SAILOG is a perseveration response, and FLAG is the corresponding source response. Similarly, we refer to the G in FLAG as the source letter for the perseverated G in SAILOG.

The rationale for the position analyses was as follows: If the abstract letter representations activated in reading a word are associated with particular positions, then perseverated letters may nonaccidentally maintain source position—that is, the perseveration position may match the source position for reasons other than chance.

Consider the beginning-based representation for FLAG shown in Figure 6a. Suppose that after FLAG is read, the G unit retains some activation, and the letter in position $B+4$ of the next stimulus word—for example, the L in SAILOR—is activated only weakly. Under these circumstances, the G from FLAG may perseverate into the response for SAILOR and more specifically may perseverate into position $B+4$, yielding the error SAIGOR. In this example, the G maintains the same beginning-based position when perseverating from the source response FLAG to the perseveration response SAIGOR.

Suppose, however, that position is defined by an end-anchored scheme, as illustrated in Figure 6b. Under this scheme, the G in FLAG is in position $E-1$. Consequently, if the G perseverated into the response for SAILOR, we might expect the perseveration to occur at position $E-1$, yielding SAILOG. As these examples illustrate, the various position schemes will sometimes differ with respect to the position in a source response that matches the position of a perseverated letter in a perseveration response. These differences between schemes provide a basis for inferring which scheme(s) underlie representation of letter position at the level where L.H.D.'s perseverations arise.

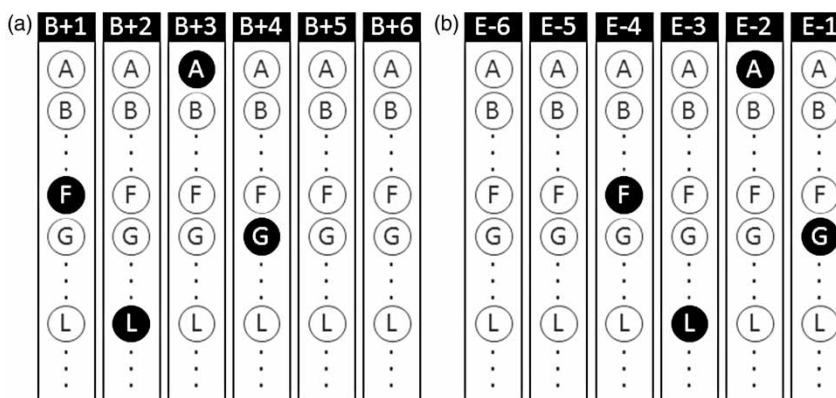


Figure 6. Beginning- and end-based representations for the word FLAG.

The position analyses were carried out over *potential perseveration–source pairs*. A potential perseveration is a letter intrusion error in which the intruded letter appears in one or more of the preceding responses within the perseveration window of three prior responses. The preceding responses that contained the intruded letter are potential source responses. In the case of L.H.D.’s error *sailor* → SAILOG, the intruded letter G is present in two of the three immediately preceding responses (FLAG and COLLEGE, but not RIB; see Table 3). SAILOG is therefore a potential perseveration response, and SAILOG–FLAG and SAILOG–COLLEGE are potential perseveration–source pairs. The total number of pairs was 2900 for the test-session corpus and 9070 for the home-practice corpus.

The set of potential perseveration–source pairs for each corpus unavoidably includes not only *true* perseveration–source pairs (i.e., pairs associating a true perseveration with its true source), but also what may be called *pseudo* perseveration–source pairs. In a pseudo pair the intruded letter either is not a perseveration from a prior response or is not paired with its true source. Suppose, for example, that L.H.D. intruded a G into the response SAILOG for reasons having nothing to do with the presence of Gs in prior responses. In this case both SAILOG–FLAG and SAILOG–COLLEGE would be pseudo perseveration–source pairs. Suppose, alternatively, that the G in

SAILOG was a true perseveration, and that the only source for the perseveration was the immediately preceding response FLAG. The pair SAILOG–FLAG would then be a true perseveration–source pair, but SAILOG–COLLEGE would be a pseudo pair. Because we had no way of knowing which particular pairs were true perseveration–source pairs and which were pseudo pairs, we included all potential perseveration–source pairs in our analyses.

Fortunately, we can determine from this mixture of true and pseudo pairs whether true perseverations maintain the position of the perseverated letter in the true source more often than expected by chance. For pseudo pairs, the letter intrusion in the potential perseveration response has nothing to do with the presence of the letter in the potential source, and therefore the intruded letter should maintain source position no more often than expected by chance. Consequently, if we find that position is maintained at an above-chance rate in the full set of potential perseveration–source pairs, we can conclude that true perseverations maintain true source position more often than expected by chance.

We can also estimate the number of true perseveration–source pairs from each data set, despite our inability to classify any individual pair. Applying the method described in Appendix A (see also Fischer-Baum et al., 2010), we obtain an estimate of 1586 true perseveration–source

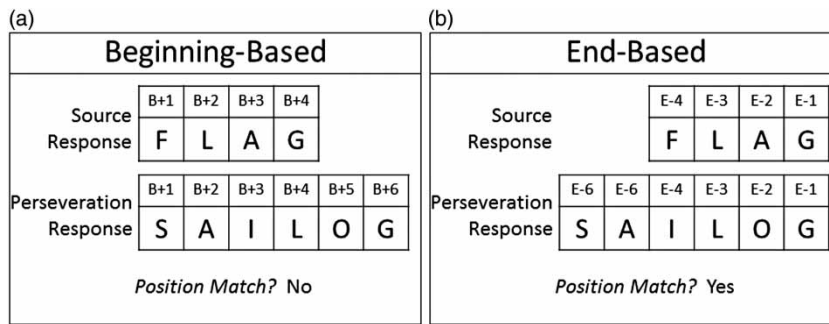


Figure 7. Example illustrating a perseveration–source position match for the end-based but not the beginning-based position scheme.

pairs for the test-session data (approximately 55% of the 2900 potential perseveration–source pairs). For the home-practice data the estimated number of true perseveration–source pairs is 5042 (56% of the 9070 potential pairs). These results indicate that we have very substantial numbers of true perseveration–source pairs from each data set, providing a solid basis for position analyses.

Extrinsic-anchor schemes

In extrinsic-anchor position schemes a letter's position is represented relative to anchor points that do not depend on the identities of the letters in the word. We evaluated five discrete extrinsic-anchor schemes: beginning-based, end-based, midpoint, closer-edge, and both-edges schemes. For each data set a computer program calculated for each scheme the proportion of potential perseveration–source pairs in which the intruded letter was in the same position in the perseveration and source responses. Figure 7 illustrates the process for the pair SAILOG–FLAG, with the beginning- and end-based schemes. When position is defined by the beginning-based scheme (Figure 7a), the intruded G in SAILOG occupies a different position than the source G in FLAG ($B+6$ vs. $B+4$). According to the end-based scheme, however, the G occupies the same position ($E-1$) in the perseveration and source responses (Figure 7b). Hence, for SAILOG–FLAG the analysis program tallied a position match for the end-based scheme but not for the beginning-based scheme.

The program also estimated for each scheme the proportion of position matches expected by chance. The chance likelihood of an intruded letter having the same position in the perseveration and source responses can be estimated by examining source control responses. A source control response was defined as a response that had the same number of letters as the source response and contained the intruded letter, but came from a trial outside the vicinity of the perseveration response (more specifically, a trial outside the range from five trials preceding through five trials following the perseveration response). For example, source control responses for the pair SAILOG–FLAG were four-letter responses that contained a G (e.g., SIGH, GATE, LONG, BAGE, HUGE), from trials not in the vicinity of the SAILOG error. For each perseveration–source pair, the analysis program generated a pool of source control responses. A Monte Carlo analysis used the source control responses to estimate for each representational scheme the proportion of position matches expected by chance. For each potential perseveration–source pair, the program randomly sampled a response from the source control pool and tabulated for each position scheme whether the perseverated letter was in the same position in that response as in the perseveration response. For example, on one run of the analysis, SIGH was chosen as the control for the source response FLAG in the SAILOG–FLAG pair. According to both beginning-based and end-based schemes, the G in the perseveration response (SAILOG) and the G in the source

control response (SIGH) are in different positions ($B+6$ vs. $B+3$, respectively, for the beginning-based scheme, and $E-1$ vs. $E-2$ for the end-based scheme). The result of each run of the Monte Carlo analysis is a single estimate of the proportion of perseverated letters expected by chance to match position in the perseveration and source responses. The entire process was carried out 10,000 times, yielding a distribution of chance proportions for each scheme.

For both the test-session and home-practice data sets, the observed proportion of position matches was greater than each of the 10,000 chance estimates for all five extrinsic-anchor schemes ($p < .0001$ for both data sets for all schemes). For example, for the beginning-based scheme in the test-session analysis, the observed proportion of position matches was .33, whereas the mean proportion expected by chance was only .22. These results demonstrate that letters were more likely than expected by chance to maintain the same extrinsic-anchor position when perseverating from a source response to a perseveration response.

The results do not necessarily mean, however, that all of the extrinsic-anchor schemes contribute to representation of letter position in reading. The alternative schemes agree with one another regarding which position in a source response matches the position of an intruded letter in a perseveration response for any pair in which the perseveration and source responses have the same number of letters. Consider the pair TOUK–LACK, in which the potential perseveration response TOUK has an intruded K. The position of the K in TOUK differs by scheme (e.g., $B+4$, $M+2$, $E-1$), but all of the extrinsic-anchor schemes agree that the K occupies the same position in TOUK (potential perseveration response) and LACK (potential source response). Hence, if letter position were encoded according to any of the extrinsic-anchor schemes, one would expect all of the schemes to show above-chance proportions of positional matches. Accordingly, we need methods for comparing schemes and assessing their contributions to representation of letter position in reading.

Comparing schemes

The observed proportions of position matches were calculated over all potential perseveration–source pairs. As previously noted, however, the pool of potential perseveration–source pairs includes not only true perseveration–source pairs (i.e., pairs consisting of a true perseveration and its true source), but also pseudo perseveration–source pairs (i.e., pairs in which the intruded letter is present in the potential source by chance). Because the likelihood of a position match for a pseudo perseveration–source pair may vary across position schemes, the pseudo pairs may make a greater or lesser contribution to the observed proportion of position matches, depending upon the scheme. Consequently, we cannot contrast schemes simply by comparing their observed proportions of position matches.

We can, however, estimate a scheme's proportion of position matches for the true perseveration–source pairs alone ($pPM_{forTruePS}$). The method is described in Appendix B (see also Fischer-Baum et al., 2010). For each position scheme, $pPM_{forTruePS}$ provides a measure of how well the scheme performs in using the position of a perseverated letter in a perseveration response to predict the position of the corresponding source letter in the true source response. For example, a $pPM_{forTruePS}$ of .60 would indicate that in 60% of the true perseveration–source pairs the scheme in question correctly predicted the position of the source letter from the position of the perseverated letter.

Table 4 presents, for each data set, the estimated $pPM_{forTruePS}$ value for each of the discrete extrinsic-anchor schemes. For both data sets, $pPM_{forTruePS}$ was highest for the both-edges scheme. In nearly two thirds of the true perseveration–source pairs, this scheme was successful in using the position of the perseverated letter in the perseveration response to predict the position of the source letter in the source response.

Results from the Monte Carlo procedure provide a basis for assessing the statistical reliability of differences among schemes in $pPM_{forTruePS}$. For each position scheme, each Monte Carlo run provided a different estimate of

Table 4. *Estimated proportions of position matches (pPMforTruePS) for true perseveration–source pairs for discrete extrinsic-anchor position schemes*

Position scheme	Data set	
	Test session	Home practice
Beginning-based	.42*	.38*
End-based	.50*	.54*
Midpoint	.44*	.44*
Closer-edge	.54*	.55*
Both-edges	.64	.67

*Significantly lower than both-edges ($p < .0001$).

the chance probability of a position match. Using these chance estimates, we estimated pPMforTruePS for each scheme on each of the 10,000 Monte Carlo runs. For both data sets the estimated pPMforTruePS was highest for the both-edges scheme on all 10,000 runs; hence, for each corpus the difference between the both-edges scheme and each of the other schemes was significant at $p < .0001$.

Controlling for number of predicted positions. Although the pPMforTruePS value for a scheme is a useful measure of a scheme's success, these values do not alone provide an adequate basis for comparing schemes. The beginning-based, end-based, midpoint, and closer-edge schemes are simple schemes, assigning only one position representation to each letter in a word. In contrast, the both-edges scheme is a compound scheme; that is, the scheme is a combination of beginning- and end-based schemes and so assigns both beginning- and end-based position representations to each letter (e.g., $B+4$ and $E-1$ for the G in FLAG). In evaluating compound schemes, one must ask whether each component scheme (e.g., beginning- and end-based components for the both-edges scheme) plays a role in the success of the overall scheme.

Comparisons of pPMforTruePS values cannot answer this question. For each potential perseveration–source pair, a simple scheme designates at most one position in the source response as the position where the intruded letter should be

found. Consider the pair SOLAT–PLENTY. According to the end-based scheme, the intruded T in SOLAT (a misreading of SOLAR) occupies position $E-1$, and so this scheme predicts that the T in the source response should be found at position $E-1$. However, the T in the source response PLENTY is in position $E-2$, and the end-based scheme fails to score a position match for the SOLAT–PLENTY pair.

In contrast to simple schemes, compound schemes often designate multiple source positions as predicted positions of the intruded letter. For example, the both-edges scheme assigns position representations $B+5$ and $E-2$ to the T in SOLAT. The scheme therefore designates positions $B+5$ and $E-2$ as predicted source positions and scores a position match if the source T is found at either of these positions. In the source PLENTY, the T is at position $B+5$, and so the both-edges scheme scores a position match.

As this example illustrates, the both-edges scheme often designates more predicted source positions than either the beginning- or end-based scheme, and so has more opportunities for scoring position matches. This point has implications for comparisons of pPMforTruePS values across schemes. Consider the finding that estimated pPMforTruePS was significantly higher for the both-edges scheme than for the end-based scheme. For all potential perseveration–source pairs, the both-edges scheme predicts the same source position as the end-based scheme (e.g., $E-1$ for the SOLAT–PLENTY pair). However, for many pairs the both-edges scheme also predicts an additional position via beginning-based coding (e.g., the $B+5$ position in PLENTY). Now suppose that the beginning-based scheme plays no role in representing letter position during visual word recognition and that as a consequence the additional predicted beginning-based positions are unrelated to the positions predicted by the scheme that actually encodes letter position. In other words, suppose that the additional positions predicted by the beginning-based component of the both-edges scheme are random relative to the positions predicted by the actual scheme for letter position representation.

Even a random guess will occasionally be correct, and so we expect the additional beginning-based predictions to yield occasional position matches for true perseveration–source pairs (as long as the end-based component of the both-edges scheme does not always predict source position correctly for these pairs). As a consequence, we expect the $p_{PMforTruePS}$ value to be higher for the both-edges scheme than for the end-based scheme, even if the beginning-based component of the both-edges scheme makes no systematic (i.e., nonrandom) contribution to predicting the source positions of perseverated letters. Hence, the finding that $p_{PMforTruePS}$ was higher for the both-edges scheme than for the end-based scheme does not necessarily mean that the beginning-based component of the both-edges scheme made a systematic contribution over and above that of the end-based scheme in predicting the source positions of perseverated letters.

To assess whether the beginning-based component of the both-edges scheme makes a systematic contribution beyond that of the end-based component, we can conduct an additional analysis that includes only the potential perseveration–source pairs for which the end-based scheme failed to predict the source position of the intruded letter (e.g., the SOLAT-PLENTY pair). For these *residual* pairs the analysis assesses whether the observed proportion of position matches for the beginning-based scheme is significantly higher than the proportion expected by chance. In other words, we ask whether the beginning-based scheme succeeds at above-chance rates in those cases where the end-based scheme fails. This residuals analysis is conceptually similar to analyses evaluating the ability of a variable to account for variance remaining after the variance explained by one or more other variables has been partialled out.

The residuals analysis revealed that the observed proportion of position matches for the beginning-based scheme was significantly higher than the proportion expected by chance in both the test-session and home-practice data, $ps < .0001$. For example, in the analysis of home-practice data, the beginning-based scheme correctly

predicted the source position for 920 of the residual pairs (18%), whereas the number expected by chance was only 639 (12%). These results demonstrate that the beginning-based component of the both-edges scheme made a significant contribution to the scheme's success, over and above the contribution of the end-based component. In complementary residuals analyses, we applied the end-based scheme to the pairs for which the beginning-based scheme failed to predict source position. These analyses revealed that the end-based component of the both-edges scheme made a systematic contribution beyond that of the beginning-based component ($ps < .0001$).

The both-edges scheme can also be viewed as a combination of closer-edge and farther-edge schemes. As we have seen, the closer-edge scheme codes letters in the first half of the word relative to the beginning and letters in the second half relative to the end. A farther-edge scheme, on the other hand, uses end-based coding for the first half of the word and beginning-based coding for the second half. A compound closer-edge-plus-farther-edge scheme therefore codes each letter relative to both the closer edge and the farther edge—that is, relative to both the beginning and end of the word—and so is equivalent to the both-edges scheme. Accordingly, we can evaluate whether the both-edges scheme outperforms the closer-edge scheme by asking whether the farther-edge component of the both-edges scheme makes a systematic contribution beyond that of the closer-edge component to the success of the both-edges scheme. Residuals analyses conducted over the potential perseveration–source pairs for which the closer-edge scheme failed in predicting the source position of the intruded letter found that the observed proportion of position matches for the farther-edge scheme reliably exceeded the proportion expected by chance in both test-session and home-practice data sets, $p < .0001$.

These analyses demonstrate that the success of the compound both-edges scheme relative to its component beginning-based, end-based, closer-edge, and farther-edge components cannot be explained by assuming that one of the components

(e.g., the closer-edge component) does all of the systematic work in predicting source positions of perseverated letters, with the complementary component (e.g., the farther-edge component) contributing only in the sense of making essentially random source position predictions that occasionally happen to yield position matches for true perseverations. Rather, each component (beginning- and end-based components given one decomposition of the scheme, closer- and farther-edge components given an alternative decomposition) plays a significant role in the success of the both-edges scheme.

Discrete versus graded position representations

The preceding analyses demonstrated that in the majority of L.H.D.'s true perseverations (64% in the test-session data and 67% in the home-practice data) the perseverated letter appeared at a source location exactly matching the beginning-based or end-based position of the letter in the perseveration response. However, for the remaining true perseverations the discrete both-edges scheme did not correctly predict the source location of the perseverated letter. Consider the perseveration-source pair PARONE-BRONZE, and assume that the intruded N in PARONE (a misreading of *parole*) is a true perseveration from the source response BRONZE. In PARONE the N occupies positions $B+5$ and $E-2$; hence, the discrete both-edges scheme predicts that an N will be found at position $B+5$ or $E-2$ of the source response BRONZE. In fact, though, the N in BRONZE occupies positions $B+4$ and $E-3$.

One potential explanation for the predictive failures of the discrete both-edges scheme is that position representations are graded. Considering first the beginning-based component of the both-edges scheme, the N in BRONZE might be represented by activating the N unit strongly in the $B+4$ pool, moderately in the $B+3$ and $B+5$ pools, and perhaps more weakly in pools for more distant positions. Given this pattern of activation, the N would be most likely to perseverate into position $B+4$ of a subsequent response, but might also perseverate into an adjacent position ($B+3$ or $B+5$), or, rarely, into a more

distant position. The same point applies to the end-based component of the both-edges scheme: If position representations are graded, we should expect the N in BRONZE to perseverate most often into position $E-3$ of a subsequent response, but also sometimes into position $E-2$ or $E-4$, or perhaps even a more distant position.

The same logic can be applied to analyses in which the position of a perseverated letter in a perseveration response is used to predict the position of the letter in the source response. If position representations are graded, we should expect sometimes to find the source letter not at an exactly matching position, but instead at a nearby position. For example, a graded both-edges scheme predicts that the N in PARONE should most often be found in source position $B+4$ or $E-3$ (the exactly matching positions), but might also be found at one of the nearby source positions (e.g., $B+4$, $B+6$, $E-2$, $E-4$).

We first evaluated a narrowly graded both-edges scheme that defined the predicted source positions to be the exactly matching positions plus the immediately adjacent positions. For PARONE-BRONZE, the scheme predicted that we would find the source N at position $B+4$, $B+5$, $B+6$, $E-2$, $E-3$, or $E-4$. In this example the beginning- and end-based positions coincide, and consequently the graded both-edges scheme identifies three predicted source positions (those occupied by the O, N, and Z). Because the N appears at one of these positions, the graded scheme is credited with a position match for this pair.

Estimated pPMforTruePS values for the narrowly graded both-edges scheme were remarkably high. Whereas the values for the discrete both-edges scheme were .64 and .67 in the test-session and home-practice data sets, respectively, the corresponding values for the graded scheme were .91 and .94. These values indicate that for over 90% of L.H.D.'s true perseverations, the source letter occupied either an exactly matching both-edges position in the source response, or an immediately adjacent position.

To determine whether the graded both-edges scheme significantly outperformed the discrete

scheme, we treated the graded scheme as a combination of the discrete both-edges scheme and what we call an adjacent both-edges scheme, which predicts that the source letter will be found at a source position immediately adjacent to an exact-match position. We then carried out a residuals analysis on the potential perseveration–source pairs for which the discrete both-edges scheme failed to predict the source position of the intruded letter, and we found that the adjacent scheme performed significantly better than chance at predicting source position, $p < .0001$ for each data set. This finding demonstrates that the adjacent component of the graded both-edges scheme contributed significantly to the success of the scheme, and hence that positing graded rather than discrete position representations significantly improves our ability to predict the source positions of perseverated letters.

We next considered a more broadly graded both-edges scheme that extended the predicted source positions to those with a distance of two from an exactly matching position. The broader grading increased the estimated pPMforTruePS values from .91 to .96 for the test-session data and from .94 to .97 for the home-practice data. Residuals analyses revealed that for both data sets the broadly graded scheme performed significantly better than chance at predicting the source position of intruded letters for which the narrowly graded scheme failed to make a correct prediction, $ps < .001$. This finding suggests that letter position representations are broadly graded at the level(s) giving rise to L.H.D.'s letter perseverations.

As illustrated in Figure 4, we assume that in graded position representations activation of a letter unit is strongest in the pools for exactly matching positions and progressively weaker in pools for progressively more distant positions. We therefore expect a perseveration's source letter to be found most often at an exactly matching both-edges position, less often at an immediately adjacent position, and still less often at a greater distance. The results for the both-edges scheme show just this pattern. The discrete both-edges scheme, which treats only exactly-

matching source positions as predicted positions, accounted for 64% and 67% of the true perseverations in the test-session and home-practice data, respectively. The narrowly graded scheme, which extended the predicted positions to those immediately adjacent to exact-match positions, accounted for an additional 27% of the true perseverations in both data sets; and the broadly graded scheme, which extended the predictions to a distance of two from an exact-match position, accounted for a further 5% (test-session data) or 3% (home-practice data).

One caveat should be noted: The graded both-edges schemes treat large numbers of source positions as predicted positions. The narrowly graded scheme designates as predicted positions a mean of 2.9 positions per source word in the test-session data set and 3.0 letters per word in the home-practice data, constituting 56% and 58% of the source letters, respectively. For the broadly graded scheme, 71% of the source positions are designated as predicted positions for both data sets. However, the success of the graded schemes is nevertheless impressive: The actual source position for more than 90% of the true perseverations fell within the 56–58% of source positions picked out by the narrowly graded both-edges scheme, and virtually all of the actual source positions fell within the 71% of source positions predicted by the broadly graded scheme.

The midpoint scheme

The final set of extrinsic-anchor analyses considered the midpoint scheme, which is not a sub-component of the both-edges scheme. We have already seen that estimated pPMforTruePS values were much lower for the discrete midpoint scheme than for the discrete both-edges scheme (see Table 4). A similar analysis contrasting broadly graded midpoint and both-edges schemes yielded the same result: The pPMforTruePS value for the graded midpoint scheme was .86 in both data sets, significantly lower than the .96 and .97 values for the graded both-edges scheme in the test-session and home-practice data, respectively ($ps < .0001$). Taken together with the evidence that each component

of the both-edges scheme contributes systematically to its success, these results indicate that the midpoint scheme is not a viable alternative to the both-edges scheme. Nevertheless, the midpoint scheme might conceivably play a role *in addition to* that of the both-edges scheme in coding letter position for reading.

To address this question we conducted residuals analyses on the potential perseveration–source pairs for which the graded both-edges scheme failed to predict the source position of the intruded letter. For both the test-session and home-practice data sets the graded midpoint scheme made no contribution beyond that of the graded both-edges scheme in predicting source position ($ps > .1$).

Letter-context schemes

Letter-context schemes represent a letter's position with respect to the identities of other letters in the word. We first evaluated adjacent-letter context schemes, in which context letters were limited to letters immediately preceding and/or following the to-be-represented position. The *preceding-letter scheme* is a bigram scheme in which the position of each letter is defined by the immediately preceding letter. According to this scheme, illustrated in Figure 3a, the A in FACE occupies the F_ position (i.e., the immediately preceded by F position). The *following-letter scheme* similarly represents each letter relative to the immediately following letter. In this scheme, the A in FACE occupies the _C position. The compound *preceding-letter + following-letter scheme* assigns each letter both a preceding- and a following-letter position representation (e.g., F_ and _C for the A in FACE). In the *trigram scheme* each letter is assigned a single position representation defined jointly by the immediately preceding and immediately following letter, such that, for example, the A in FACE occupies the F_C position. The adjacent-letter context schemes included word-boundary representations. For example, the preceding-letter scheme represented the position of the F in FACE as #_ (immediately following the word-initial boundary), and the following-letter

scheme represented the E's position as _# (immediately preceding the word-final boundary).

Finally, we evaluated an *open-bigram scheme* in which context letters were not limited to those immediately adjacent to the to-be-represented position. The scheme closely resembles the constrained open-bigram model proposed by Grainger and van Heuven (2003; see also Grainger et al., 2006). Each letter was assigned preceding- and following-letter position representations relative to the letters in the three preceding and following positions. For example, the P in HOSPITAL was assigned preceding-letter representations H_, O_, and S_, as well as following-letter representations _I, _T, and _A. The open-bigram scheme is a graded letter-context scheme, in that nearby letters in a word have similar position representations. For example, the P and I in HOSPITAL are each assigned six position codes, and four of these (O_, S_, _T, _A) are the same for the two letters. In accord with the assumptions of the constrained open-bigram model, word boundary representations were not included. For example, the position representations for the H in HOSPITAL were limited to _O, _S, and _P.

As in the analyses of the extrinsic-anchor schemes, a position match was counted for a letter-context scheme whenever the intruded and source letters shared one or more position representations. Consider the potential perseveration–source pair SOLAT–PLENTY. According to the open-bigram scheme, the intruded T in SOLAT has position representations O_, L_, and A_, and the source T in PLENTY has representations L_, E_, N_, and _Y. Because the intruded and source Ts share the position representation L_, the open-bigram scheme scores a position match for this pair.

In the position analyses for the letter-context schemes, we excluded intruded letters from the set of potential perseverations if a context letter in the perseveration response was also an intruded letter. For example, in the error *lamb* → LAWF, we excluded the intruded F from analyses of the preceding-letter scheme, because the preceding-context letter W was also an intrusion. As

Table 5. *Estimated proportions of position matches for true perseveration–source pairs (pPMforTruePS) for letter–context position schemes*

Position scheme	Data set	
	Test session	Home practice
Preceding–letter	.25*	.24*
Following–letter	.40*	.51*
Preceding–letter + Following–letter	.53*	.61*
Trigram	.09*	.11*
Open–bigram	.59*	.55*

*Significantly lower than graded both–edges ($p < .0001$).

discussed in Fischer–Baum et al. (2010, p. 482), this procedure is necessary to avoid over- or underestimating the pPMforTruePS values for the letter–context schemes.

As shown in the first four rows of Table 5, the pPMforTruePS values for the adjacent–letter context schemes ranged from .09 to .61. In contrast, when the broadly graded both–edges scheme was applied to the set of potential–perseveration source pairs used in each of the adjacent–letter context scheme analyses, the pPMforTruePS values were significantly higher, ranging from .96 to .97 ($ps < .0001$). Clearly, none of the adjacent–letter context schemes are viable alternatives to the graded both–edges scheme as candidates for representing letter position at the level of representation giving rise to L.H.D.’s reading errors.

To assess whether any of these schemes might nevertheless play a role in addition to the both–edges scheme, we conducted residuals analyses in which each adjacent–letter context scheme was applied to the potential perseveration–source pairs for which the broadly graded both–edges scheme failed to predict the source position of the intruded letter. For both the test–session and home–practice data, none of the letter–context schemes predicted source position at above–chance rates ($ps > .25$). These results strongly suggest that adjacent–letter bigram and trigram schemes do not play a role in representing letter position at the level of representation giving rise to L.H.D.’s reading errors.

It is worth noting that the adjacent–letter bigram schemes achieved even modest pPMforTruePS values only because word–boundary representations were included in these schemes. However, coding of word–initial position as #_ is equivalent to beginning–based extrinsic–anchor coding of that position as $B+1$, and the _# coding of word–final position is equivalent to end–based coding of that position as $E-1$. If intrusions at positions represented by word–boundary codes are excluded from the position analyses, the maximum pPMforTruePS value plummets from .53 to .22 in the test–session data, and from .61 to .27 in the home–practice data. This finding indicates that the adjacent–letter context schemes perform poorly in predicting source position on the basis of contextual representations in which the context element is a letter (e.g., F_ or _E). In contrast, the pPMforTruePS values for the broadly graded both–edges scheme remain very high (.95–.97) when these intrusion positions are excluded.

Results for the open–bigram scheme are presented in the bottom row of Table 5. This broadly graded scheme, which represents each letter relative to the three preceding and three following letters, achieved pPMforTruePS values of only .59 in the test–session data and .55 in the home–practice data. Applied to the same sets of potential perseveration–source pairs, the broadly graded both–edges scheme showed pPMforTruePS values of .97 and .98, respectively ($ps < .0001$). Furthermore, residuals analyses revealed that the open–bigram scheme made no contribution beyond that of the both–edges scheme in predicting source positions of intruded letters ($ps > .5$). These findings suggest that open–bigram schemes (e.g., Dehaene et al., 2005; Grainger et al., 2006; Grainger & van Heuven, 2003; Whitney, 2001) are not implicated in letter position representation at the level of the reading system from which L.H.D.’s errors originate.

Orthosyllabic schemes

We turn now to orthosyllabic position schemes, in which a letter’s position is represented in terms of

its role in an orthographic syllable structure. A computer program syllabified L.H.D.'s responses, assigning each letter a position defined jointly by the position of the syllable in which the letter appeared and the role of the letter within that syllable. The position of a syllable within a response was defined by a both-edges scheme. For example, the syllable PER in PERSON is the first syllable from the beginning of the word and the second syllable from the end and so was assigned syllable position representations $B+1$ and $E-2$. To avoid confusion between position representations for syllables and those for letters, we denote syllable positions with σ (for syllable) subscripted with a beginning- or end-based position, as in σ_{B+1} and σ_{E-2} . In referring to a particular syllable in a word, we include both the beginning- and end-based syllable position representations; thus, the syllable PER in PERSON is in position $\sigma_{B+1/E-2}$.

Within each syllable, the position of a letter was defined by its role as an onset, nucleus, or coda letter. For example, in STRING, the S, T, and R occupy onset positions, the I a nucleus position, and the N and G coda positions. Letters were further differentiated by their positions within an onset, nucleus, or coda. In STRING the S, T, and R occupy the first, second, and third onset positions (O1, O2, and O3), respectively.

Two alternative syllabification algorithms were applied to the responses. The onset maximization (OMAX) algorithm assigned as many letters as possible to syllable onsets, whereas the coda maximization (CMAX) algorithm assigned letters preferentially to codas. For example, the OMAX algorithm syllabified the response REGENT as RE-GENT, assigning the G to the onset of the second syllable, whereas the CMAX algorithm placed the G in the first-syllable coda, yielding REG-ENT. Onset maximization is frequently assumed in phonological research (e.g., Kahn, 1976), whereas coda maximization is a fundamental feature of the basic orthographic syllable structure (BOSS) proposed by Taft (e.g., Taft, 1979, 1987, 1992, 2001, 2002; Taft & Krebs-Lazendic, 2013) for representing orthographic syllable structure in the mental lexicon. All analyses were

carried out with both OMAX and CMAX syllabifications.

Position analysis method

The position analysis method that we used to evaluate extrinsic-anchor and letter-context schemes is not entirely appropriate for syllabic-scheme analyses, for reasons described in Appendix C. The appendix also explains the modified analysis method that we developed to circumvent the problem. Rather than using the location of the intruded letter in the perseveration response to predict the location of the source letter in the source response, the modified method does the opposite, using the position of the source letter to predict the position of the perseverated letter. We refer to this modified analysis method as the prospective method because it looks forward from the source to the perseveration response, in contrast to the original retrospective method, which looks backward from the perseveration response to the source response.

We used the prospective method to evaluate a discrete orthosyllabic position scheme, as well as two graded schemes. For each scheme we generated predicted perseveration responses for each perseveration-source pair and tallied a correct prediction for a scheme if any of the predicted responses matched the actual perseveration response. Appendix C describes the procedure in detail; here we provide a brief overview.

For the discrete orthosyllabic scheme, we assumed that the source letter could perseverate only into an exactly matching orthosyllabic position in the perseveration response. Consider the perseveration-source pair SAVOL-SLANO. The perseveration response SLANO is a misreading of *slant* and so contains an intruded O. The source response SAVOL contains an O in the N1 position of syllable $\sigma_{E-1/B+2}$. Therefore, the discrete orthosyllabic scheme predicts that if the O from SAVOL perseverates into the response for *slant*, the perseverated letter will appear at position N1 of syllable σ_{E-1} , yielding SLONT, or position N1 of syllable σ_{B+2} , yielding SLANTO. Neither of these predicted perseveration responses matches the actual response SLANO.

For the graded role within-syllable scheme we assumed that coding of position was graded within a syllable's onset, nucleus, and coda roles. According to this scheme, the O in SAVOL is associated most strongly with the N1 position of the syllables σ_{B+2} and σ_{E-1} , but also less strongly with the other possible nucleus positions in these syllables (e.g., N2). We therefore assumed that a source letter could perseverate into any target position having the same basic syllabic role (onset, nucleus, or coda) within a matching syllable. For example, in the case of the SAVOL–SLANTO pair, we assumed that the O in SAVOL could perseverate into the N2 position of σ_{E-1} , yielding SLAONT. Hence the predicted perseverations for the graded role-within-syllable scheme included SLAONT as well as SLONT and SLANTO.

Finally, for the graded role across-syllables scheme we assumed that graded coding of syllable-role position extended to adjacent syllables. For example, we assumed that the O in SAVOL was associated not only with nucleus positions in syllables σ_{B+2} and σ_{E-1} , but also (less strongly) with nucleus positions in the immediately adjacent syllables. As a consequence, this scheme's predicted perseverations for SAVOL–SLANO included not only SLONT, SLANTO, and SLAONT, but also OSLANT, resulting from perseveration of an O into a nucleus position of syllable σ_{E-2} . If any of the predicted perseveration responses matched the actual response, a correct prediction was tallied for the orthosyllabic scheme.

For purposes of comparing the extrinsic-anchor both-edges scheme with the orthosyllabic schemes, we also analysed the broadly graded both-edges scheme with the prospective method. For extrinsic-anchor and letter-context schemes, the prospective and retrospective methods yield virtually identical results when applied to the same data. However, the prospective analyses were restricted to single-letter substitution and insertion errors (for reasons explained in Appendix C), whereas the retrospective analyses for the both-edges scheme were not. Therefore, to allow valid comparisons between the both-edges scheme and the orthosyllabic schemes we

applied the both-edges scheme to the same potential perseveration–source pairs used in the orthosyllabic scheme analyses.

Table 6 presents the estimated pPMforTruePS values for the discrete and graded orthosyllabic schemes and for the broadly graded both-edges scheme. For both the test-session and home-practice data sets, and for both onset maximization and coda maximization syllabifications, pPMforTruePS was significantly higher for the both-edges scheme than for any of the orthosyllabic schemes ($p < .0001$). Among the orthosyllabic schemes, the highest pPMforTruePS values were obtained for the graded roles across-syllables scheme, but these values (.61 to .85) were substantially and significantly lower than the corresponding values for the graded both-edges scheme (.97 for both data sets). This result obtained despite the fact that the number of predicted perseveration responses, and hence the number of opportunities for matching the actual perseveration response, was higher for the graded roles across-syllables scheme than for the graded both-edges scheme. The mean number of predicted perseveration responses for a perseveration–source pair was 4.3 to 4.5 for the graded roles across-syllables scheme, but only 3.9 for the graded both-edges scheme.

For each of the 12 Orthosyllabic Scheme \times Data Set \times Syllabification combinations we conducted a residuals analysis in which the broadly graded both-edges scheme was applied to the residual perseveration–source pairs (i.e., those for which the orthosyllabic scheme failed in using the source position of the intruded letter to predict its position in the perseveration response). In all cases the both-edges scheme performed significantly better than chance in predicting perseveration position ($p < .0001$). Taken together with the higher pPMforTruePS values for the both-edges scheme, these results demonstrate that none of the syllabic schemes are viable alternatives to the both-edges scheme as candidates for representing letter position at the level giving rise to L.H.D.'s perseverations.

To determine whether any of the orthosyllabic schemes might nevertheless make some

Table 6. Estimated proportions of position matches for true perseveration–source pairs ($pPM_{forTruePS}$) for orthosyllabic position schemes and the graded both-edges scheme

Position scheme	Syllabification method			
	Onset maximization		Coda maximization	
	Test session	Home practice	Test session	Home practice
Discrete orthosyllabic	.46*	.52*	.60*	.64*
Graded role within-syllable	.57*	.61*	.75*	.76*
Graded role across-syllables	.61*	.64*	.85*	.82*
Graded both-edges	.97	.97	.97	.97

*Significantly lower than graded both-edges ($p < .0001$).

contribution to letter position representation, we conducted 12 residuals analyses in which the orthosyllabic schemes were applied to the potential perseveration–source pairs for which the narrowly-graded both-edges scheme failed to predict perseveration position. For the home-practice data, none of the orthosyllabic schemes showed above-chance performance with either onset maximization or coda maximization syllabifications. For the test-session data, one of the six analyses yielded a significant effect: The graded roles across-syllable scheme with coda maximization performed significantly better than chance in predicting perseveration position for the residual perseveration–source pairs ($p < .01$). However, in addition to occurring only in the test-session data set, this effect was very small: The orthosyllabic scheme accounted for only an estimated 16 true perseverations not explained by the graded both-edges scheme.

The results of orthosyllabic scheme analyses may be summarized as follows: Estimated $pPM_{forTruePS}$ values were consistently higher for the graded both-edges scheme than for any of the orthosyllabic schemes, and the both-edges scheme performed significantly better than chance in predicting perseveration position in all residuals analyses conducted over perseveration–source pairs not accounted for by an orthosyllabic scheme. In contrast, residuals analyses applying orthosyllabic schemes to pairs not explained by the graded both-edges scheme yielded only a single significant effect in one of our two data

sets, and this effect was trivially small. We conclude that the results provide no evidence for a role of orthosyllable position representations at the level(s) of representation implicated in L.H.D.'s reading deficit.

Beginning–end gradients in both-edges position representation

The position analyses reported in the preceding sections strongly support the conclusion that letter position is represented by a graded both-edges position scheme at the levels of the reading process giving rise to L.H.D.'s perseverations. In a final set of analyses, we explored a potential property of the both-edges representations. In some competitive queuing (CQ) simulations of spelling, both-edges position representations are posited, and these representations are assumed to have an interesting property: Beginning-based coding plays a more important role for letters close to the beginning of a word than for those toward the end, whereas end-based coding is more important for letters near the end of the word than for those close to the beginning (e.g., Glasspool & Houghton, 2005; Houghton et al., 1994). This property could be instantiated in our illustrative representational format by assuming that activation of letter units is strongest in the $B+1$ and $E-1$ pools, decreasing systematically across pools representing positions increasingly distant from the beginning and end anchor points. For example, in a discrete both-edges representation of FLAG

we might assume that activation is strong for the F unit in the $B+1$ pool, somewhat less strong for the L unit in the $B+2$ pool, rather weak for the A unit in the $B+3$ pool, and weaker still for the G unit in the $B+4$ pool; whereas in the end-based pools activation might be strong for the G in the $E-1$ pool, less strong for the A in the $E-2$ pool, and so forth. Given this pattern of activation, beginning-based coding would play a dominant role for letters near the beginning of a word, whereas end-based coding would predominate for letters near the end. In the case FLAG, for example, the F unit would be much more strongly activated in the $B+1$ pool than in the $E-4$ pool, whereas the G unit would have stronger activation in the $E-1$ than in the $B+4$ pool.

To assess whether both-edges representations evidence beginning–end gradients at the level giving rise to L.H.D.’s errors, we applied beginning- and end-based schemes separately to perseverations into the first and second halves of responses. If beginning-based coding is more important than end-based coding near the beginning of a word, then perseverations into the first half of a response should be more likely to maintain beginning- than end-based position—that is, perseverated and source letters should be more likely to match on beginning- than on end-based position. Similarly, if end-based coding is more important than beginning-based coding near the end of a word, then perseverations into the second half of a response should maintain end-based position more often than beginning-based position.

For both discrete and graded both-edges schemes, the analyses for both data sets showed precisely this pattern. Here we report results for the narrowly graded scheme. For perseverations into the beginning half of a response, estimated $pPM_{\text{forTruePS}}$ values were significantly higher for the beginning-based scheme than for the end-based scheme: .81 vs. .67, respectively, for the test-session data, and .83 vs. .50 for the home-practice data, $ps < .001$. Perseverations into the end half of a response, however, showed the opposite pattern: $pPM_{\text{forTruePS}}$ values were significantly higher for the end- than for the beginning-based scheme: .86 vs. .74, respectively,

for the test-session data, and .78 vs. .70 for the home-practice data, $ps < .001$.

Taken together with the results of the preceding analyses, these findings suggest that the both-edges position representations at the level giving rise to L.H.D.’s perseverations exhibit two distinct forms of grading. First, each letter is associated not only with a representation of the exact position it occupies in a word, but also less strongly with representations of nearby positions, such that, for example, the F in FLAG is associated strongly with position $B+1$, but also less strongly with position $B+2$, and so forth. (Another way of saying this is that position representations are coarse, specifying not exact but rather approximate positions, so that $B+2$, for example, should be glossed as “approximately two positions from the beginning.”) Second, positions closer to the beginning or end anchor point have stronger representations than more distant positions, such that, for example, in FLAG the $B+1$ representation for the F is stronger than its $E-4$ representation. This second form of grading may provide a basis for explaining the finding that L.H.D.’s error rate is higher for letters near the end of a word than for letters near the beginning: She may be more impaired at activating abstract letter identities in end-based position representations (which are most important for letters near the end of a word) than in beginning-based representations (which are most important for letters near the beginning).

GENERAL DISCUSSION

The present study reported extensive analyses of reading errors produced by patient L.H.D., who is selectively impaired in visual letter identification as a consequence of a large ventral occipital-temporal lesion. L.H.D.’s letter identification errors are often perseverations from prior responses, and we used two large corpora of her letter perseveration errors to infer how letter position is represented at prelexical levels of the visual word identification system—in particular, at levels representing allographs (i.e., stored letter

shape representations) and abstract letter identities. Our analyses considered three classes of position representation schemes: extrinsic-anchor schemes, which represent a letter's position relative to one or more anchor points that do not depend on the identities of the letters in the word; letter-context schemes, which specify a letter's position with respect to preceding and/or following letters; and orthosyllabic schemes, which represent position on the basis of a letter's role in an orthographic syllable structure. Within each class, we considered several specific positional schemes, including both discrete schemes, in which position representations have no similarity structure, and graded schemes, in which nearby positions within a word have more similar representations than distant positions.

The results of our analyses strongly supported an extrinsic-anchor both-edges scheme in which a letter's position is represented relative to both the beginning and the end of the word. For example, in this scheme the position of the L in FLAG is represented as $B+2$ (second position from the beginning of the word) and $E-3$ (third position from the end of the word). Our results also indicated that the both-edges position representations are graded rather than discrete, such that, for example, the F and L in FLAG have more similar position representations than the F and G. The present study's evidence for graded both-edges position representations in visual word recognition converges with the evidence from the illusory word paradigm obtained by Fischer-Baum et al. (2011) in their study of neurologically intact participants. This convergence of evidence from two different methods strengthens our conclusion in favour of the graded both-edges scheme.

Our results are problematic for alternative proposals concerning letter position representation in visual word recognition, including those positing beginning-based (e.g., Coltheart et al., 2001), midpoint (Caramazza & Hillis, 1990), letter-context (e.g., Dehaene et al., 2005; Grainger et al., 2006; Whitney, 2001), and orthosyllabic (e.g., Perry et al., 2010) schemes. Our findings may also pose challenges for several recent visual word-

processing theories with more complex assumptions regarding the representation and processing of positional information. The overlap model (Gomez et al., 2008), the spatial coding model (Davis, 2010b), and the noisy channel model (Norris & Kinoshita, 2012) each instantiate a graded representation of letter position that makes no reference to syllabic structure or letter context. As they currently stand, none of these theories can account for the basic phenomenon of letter perseveration errors that preserve source position, and none posits a both-edges letter position representation scheme. However, the theories could perhaps be modified to allow for position-maintaining perseveration errors, and the perseverations might then show a tendency to preserve graded both-edges source position. In the overlap model, letter strings of different lengths are aligned such that they span a common representational space. This normalized space is divided into slots depending on the number of letters in a string. For example, the space is partitioned into seven equal slots for a seven-letter string and three equal slots for a three-letter string. As a consequence, even when two strings have different lengths, their first letters and their last letters are aligned within the representational space. In the noisy channel and spatial coding models, all possible alignments are considered when strings of different lengths are compared. These alignments include one in which the beginnings of the strings are aligned and one in which the ends are aligned. Given these assumptions, the theories might well be able to accommodate perseveration errors that preserve both-edges source position.

However, the theories would perhaps encounter difficulty with some aspects of the results from L.H.D. For example, each theory would probably predict some contribution of a midpoint scheme above and beyond that of the both-edges scheme, unlike what was observed with L.H.D. In the noisy channel and spatial coding models, midpoint as well as beginning and endpoint alignments are considered when strings are compared, and in the overlap model, midpoints as well as beginnings and ends are aligned within the normalized representational space. (See the

supplemental materials of Fischer-Baum et al., 2011, for further discussion of this point in the context of letter migration errors.) The extent to which these theories could accommodate our results remains an open question; the best way to resolve this issue would be to modify the simulations of the theories to generate position-maintaining perseveration errors and then to evaluate how well the simulated errors match the observed pattern with L.H.D.

Although many of the existing theories of visual word processing encounter difficulties in explaining the results of the current study, these theories are nevertheless able to account for substantial amounts of empirical data. For example, Davis (2010b) lists more than 60 benchmark results that are explained by his spatial coding model. Given the impressive success of these theories, what should we conclude from their apparent inconsistency with the results from L.H.D.? One possibility is that different position representation schemes are implemented at different levels of the reading system, such that position is represented by a graded both-edges scheme at the peripheral processing levels that culminate in letter identification, with some alternative scheme(s) instantiated at more central levels. Given this possibility, a theory concerned only with post-letter-identification reading processes could be reconciled with the present results by assuming that these results speak to levels of representation beyond the scope of the theory. However, some of the theories positing non-both-edges positional schemes are clearly intended to address peripheral reading processes (e.g., Adelman, 2011; Caramazza & Hillis, 1990; Coltheart et al., 2001), and many of the benchmark results taken as support for other theories arguably arise from peripheral reading processes. We suggest that theorists treating our results as outside the scope of their theories should, at the least, offer explicit claims about the level(s) of representation their theories encompass, and about the level(s) giving rise to the various phenomena taken to be theory-relevant. Ideally, assumptions about how and why graded both-edges representations are

transformed into alternative forms of position representation should also be offered.

Another approach to reconciling our results with successful theories positing alternative forms of position representation is to consider the possibility that the theories' successes may have little to do with their assumptions about positional schemes. Theories adopting different position representation schemes can accommodate many of the same benchmark results: Some of these results appear unrelated to issues of position representation (e.g., frequency effects), and other findings that speak to position representation can nevertheless be accommodated by theories that posit various schemes. For example, almost any graded position scheme can account for the transposition priming effect (e.g., SERVICE is primed more by SEVRICE than by SEDLICE), because adjacent letters have similar position representations in graded schemes. These points suggest that in addition to explaining the results of the present study, theories adopting graded both-edges position representations could conceivably be as successful as extant theories in accounting for Davis's (2010b) benchmark results.

Indeed, we believe that the empirical approach taken in this paper—as well as the complementary approach taken with neurologically intact adults by Fischer-Baum et al. (2011)—provides a much-needed addition to masked priming studies when it comes to distinguishing among competing theories of visual word recognition. Theories that make very different assumptions about how letter sequences are represented can account for many of the same benchmark reading results from the masked priming literature. Several authors have noted how difficult it is to identify priming conditions that distinguish between leading theories (Davis, 2010a; Whitney et al., 2012). Analysis of errors produced by individuals with acquired dyslexia or by neurologically intact individuals in the illusory word paradigm (Davis & Bowers, 2004, 2006; Fischer-Baum et al., 2011) represents an important additional tool for addressing questions of letter position representation and may provide

novel insights that are difficult to obtain from masked priming studies alone.

Position representation in spelling and other cognitive domains

Fischer-Baum et al. (2010) analysed perseveration errors in two acquired dysgraphic patients, arguing that the results supported graded both-edges representation of letter position within the cognitive spelling system (see also Glasspool, 1998; Glasspool & Houghton, 2005; Houghton et al., 1994, for computational models of spelling with both-edges position representations). In conjunction with the reading evidence from patient L.H.D. and the illusory word paradigm (Fischer-Baum et al., 2011), the spelling results suggest that the same positional scheme may underlie letter position encoding in reading and spelling. This conclusion is consistent with theories holding that reading and spelling rely on shared representations of letter identity and position (e.g., Caramazza & Hillis, 1990).

Evidence for graded both-edges representations has also been reported in other cognitive domains. For example, Henson (1999) has argued for both-edges representation of item position in immediate serial recall. The fact that the same scheme is used to represent the position of a letter in a word in reading and the position a word in a list in immediate serial recall supports theories of visual word processing that argue that the reading system is built on a series of domain-general representational principles (e.g., Norris & Kinoshita, 2012).

Manuscript received 8 October 2013

Revised manuscript received 27 December 2013

Revised manuscript accepted 2 January 2014

First published online 4 February 2014

REFERENCES

Adelman, J. S. (2011). Letter in time and retinotopic space. *Psychological Review*, 118, 570–582.

- Badecker, W. (1996). Representational properties common to phonological and orthographic output systems. *Lingua*, 99, 55–83.
- Brown, G. D. A., & Loosemore, R. P. W. (1994). Computational approaches to normal and impaired spelling. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and application* (pp. 319–335). Chichester: John Wiley & Sons.
- Caramazza, A., & Hillis, A. E. (1990). Levels of representation, coordinate frames, and unilateral neglect. *Cognitive Neuropsychology*, 7, 391–445.
- Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition*, 37, 243–297.
- Cohen, L., & Dehaene, S. (1998). Competition between past and present: Assessment and interpretation of verbal perseverations. *Brain*, 121, 1641–1659.
- 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. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256.
- Cox, G. E., Kachergis, G., Recchia, G., & Jones, M. N. (2011). Towards a scalable holographic word-form representation. *Behavioral Research Methods*, 43, 602–615.
- Davis, C. J. (2010a). SOLAR versus SERIOL revisited. *European Journal of Cognitive Psychology*, 22, 695–724.
- Davis, C. J. (2010b). The spatial coding model of visual word identification. *Psychological Review*, 117, 713–758.
- Davis, C. J., & Bowers, J. S. (2004). What do letter migration errors reveal about letter position coding in visual word recognition? *Journal of Experimental Psychology: Human Perception and Performance*, 30, 923–941.
- Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: Evidence from orthographic similarity effects. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 535–557.
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *TRENDS in Cognitive Sciences*, 9, 335–341.
- Ellis, A. W., Flude, B. M., & Young, A. W. (1987). “Neglect Dyslexia” and early visual processing of letters in words and nonwords. *Cognitive Neuropsychology*, 4, 439–464.

- Fischer-Baum, S., Charny, J., & McCloskey, M. (2011). Both-edges representation of letter position in reading. *Psychonomic Bulletin and Review*, *18*, 1083–1089.
- Fischer-Baum, S., McCloskey, M., & Rapp, B. (2010). Representation of letter position in spelling: Evidence from acquired dysgraphia. *Cognition*, *115*, 466–490.
- Friedmann, N., Kerbel, N., & Shvimer, L. (2010). Developmental attentional dyslexia. *Cortex*, *46*, 1216–1237.
- Glasspool, D. W. (1998). *Modelling serial order in behaviour: Studies of spelling*. Doctoral Thesis, University College London.
- Glasspool, D. W., & Houghton, G. (2005). Serial order and consonant-vowel structure in a graphemic output buffer model. *Brain and Language*, *94*, 304–330.
- Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*, 577–600.
- Grainger, J., Granier, J. P., Farioli, F., Van Assche, E., & van Heuven, W. (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., & van Heuven, W. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental lexicon: "Some words to talk about words"* (pp. 1–23). New York: Nova Science Publishers.
- Hannagan, T., Dupoux, E., & Christophe, A. (2011). Holographic string encoding. *Cognitive Science*, *35*, 79–118.
- Hannagan, T., & Grainger, J. (2012). Protein analysis meets visual word recognition: A case for string kernels in the brain. *Cognitive Science*, *36*, 575–606.
- Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*, 491–528.
- Henson, R. N. A. (1999). Positional information in short-term memory: Relative or absolute? *Memory and Cognition*, *27*, 915–927.
- Houghton, G., Glasspool, D., & Shallice, T. (1994). Spelling and serial recall: Insights from a competitive queueing model. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 365–404). Chichester: John Wiley & Sons.
- Jacobs, A., Rey, A., Ziegler, J., & Grainger, J. (1998). MROM-p. An interactive activation, multiple readout model of orthographic and phonological processes in visual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 147–188). Hillsdale, NJ: Lawrence Erlbaum Associates Ltd.
- Kahn, D. (1976). *Syllable-based generalizations in English phonology*. Doctoral dissertation, MIT.
- Kinoshita, S., & Norris, D. (2009). Transposed-letter priming of prelexical orthographic representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1–18.
- Kinoshita, S., & Norris, D. (2013). Letter order is not coded by open bigrams. *Journal of Memory and Language*, *69*, 135–150.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception, part I: An account of basic findings. *Psychological Review*, *88*, 375–407.
- McCloskey, M., Macaruso, P., & Rapp, B. (2006). Grapheme-to-lexeme feedback in the spelling system: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*, *23*, 278–307.
- Miozzo, M., & Caramazza, A. (1998). Varieties of pure alexia: The case of failure to access graphemic representations. *Cognitive Neuropsychology*, *15*(1), 203–238.
- Mozer, M. (1987). Early parallel processing in reading: A connectionist approach. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 83–104). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there is nothing special about the perception of orthography. *Psychological Review*, *119*, 517–545.
- Norris, D., Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language*, *62*, 254–271.
- Olson, A., & Caramazza, A. (2004). Orthographic structure and deaf spelling errors: Syllables, letter frequency, and speech. *Quarterly Journal of Experimental Psychology A*, *57*, 385–417.
- Perea, M., & Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 97–120). Hove, UK: Psychology Press.
- 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, 106–151.
- 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.
- Prinzmetal, W., Treiman, R., & Rho, S. H. (1986). How to see a reading unit. *Journal of Memory and Language*, 25, 461–475.
- Rapp, B. (1992). The nature of sublexical information: The bigram trough hypothesis examined. *Journal of Memory and Language*, 31, 33–53.
- Rapp, B. C., & Caramazza, A. (1989). Letter processing in reading and spelling: Some dissociations. *Reading and Writing: An Interdisciplinary Journal*, 1, 3–23.
- Schubert, T., & McCloskey, M. (2014). Pre-lexical representations and processes in reading: Evidence from an acquired dyslexia. *Cognitive Neuropsychology* 30, 360–395.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 532–568.
- Seymour, P. H. K. (1979). *Human visual cognition*. New York: St. Martin's Press.
- Spoehr, K., & Smith, E. (1973). The role of syllables in perceptual processing. *Cognitive Psychology*, 5, 71–89.
- Taft, M. (1979). Lexical access via an orthographic code: The basic orthographic syllabic structure (BOSS). *Journal of Verbal Learning and Verbal Behavior*, 18, 21–39.
- Taft, M. (1987). Morphographic processing: The BOSS re-emerges. In M. Coltheart (Ed.), *Attention and performance, XII* (pp. 265–279). London: Lawrence Erlbaum Associates Limited.
- Taft, M. (1992). The body of the BOSS: Subsyllabic units in the lexical processing of polysyllabic words. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 1004–1014.
- Taft, M. (2001). Processing of orthographic structure by adults of different reading ability. *Language and Speech*, 44, 351–376.
- Taft, M. (2002). Orthographic processing of polysyllabic words by native and non-native English speakers. *Brain and Language*, 81, 532–544.
- Taft, M., & Krebs-Lazendic, L. (2013). The role of orthographic syllabic structure in assigning letters to their position in visual word recognition. *Journal of Memory and Language*, 68, 85–97.
- Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, 8, 221–243.
- Whitney, C., Bertrand, D., & Grainger, J. (2012). On coding the position of letters in words: A test of two models. *Experimental Psychology*, 59(2), 109–114.
- 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 and Performance*, 24, 1131–1161.

APPENDIX A

Estimating numbers of true and pseudo perseveration–source pairs

This appendix describes the method for estimating the numbers of true and pseudo perseveration–source pairs (henceforth TruePS and PseudoPS, respectively) in a set of potential perseveration–source pairs (PotentialPS). PotentialPS may be expressed as a sum of TruePS and PseudoPS:

$$\text{PotentialPS} = \text{TruePS} + \text{PseudoPS} \quad (\text{A1})$$

TruePS and PseudoPS are both unknowns, but we can substitute an equivalent expression for PseudoPS and thereby eliminate this unknown. The substitution requires two values from a perseveration analysis that considers all and only the prior responses within the perseveration window (three prior responses for L.H.D.): IP, the total number of intrusion–prior–response pairs,

and pChanceLinP, the estimated probability of the intruded letter appearing in the prior response by chance.

PseudoPS can be expressed in terms of these values according to the following logic: A pseudo perseveration–source pair is any intrusion–prior–response pair that is not a true–perseveration–source pair, but includes the intruded letter in the prior response by chance. Accordingly, the expected number of pseudo pairs is the number of intrusion–prior–response pairs that are not true perseveration–source pairs (IP – TruePS) times the probability of the intruded letter being present in a prior response by chance:

$$\text{PseudoPS} = (\text{IP} - \text{TruePS}) \times \text{pChanceLinP} \quad (\text{A2})$$

Substituting this expression for PseudoPS into Equation (A1) we obtain

$$\text{PotentialPS} = \text{TruePS} + (\text{IP} - \text{TruePS}) \times \text{pChanceLinP} \quad (\text{A3})$$

Solving for TruePS is then trivial.

Consider as an example the set of 2900 potential perseveration–source pairs from L.H.D.’s test-session data. The relevant perseveration analysis provided values of 7722 for IP and .214 for pChanceLinP. Inserting these values into Equation (A3) yields

$$\text{TruePS} = 2900 - (7722 - \text{TruePS}) \times .214 \quad (\text{A4})$$

Solving for TruePS gives us 1586 as the estimated number of true perseveration–source pairs. The estimated value for PseudoPS may then be calculated as

$$\begin{aligned} \text{PseudoPS} &= \text{PotentialPS} - \text{TruePs} = 2900 - 1586 \\ &= 1314 \end{aligned} \quad (\text{A5})$$

APPENDIX B

Estimating the proportion of position matches for true perseverations

For any given positional scheme we can estimate the number of position matches for true perseverations (PsnMtchforTruePS) using values from the perseveration and position analyses. We first express the number of observed position matches for the full set of potential perseveration–source pairs (PsnMtchforPotentialPS) as a sum of position matches for true perseveration–source pairs and position matches for pseudo perseveration–source pairs:

$$\begin{aligned} \text{PsnMtchforPotentialPS} &= \text{PsnMtchforTruePS} \\ &+ \text{PsnMtchforPseudoPS} \end{aligned} \quad (\text{B1})$$

For pseudo perseveration–source pairs, position matches will occur only by chance. Hence, the number of position matches for the pseudo pairs should equal the number of pseudo pairs (PseudoPS) times the probability of a position match by chance (pChancePsnMtch):

$$\begin{aligned} \text{PsnMtchforPseudoPS} &= \text{PseudoPS} \\ &\times \text{pChancePsnMtch} \end{aligned} \quad (\text{B2})$$

Substituting this expression into Equation (B1) and solving for PsnMtchforTruePS yields

$$\begin{aligned} \text{PsnMtchforTruePS} &= \text{PsnMtchforPotentialPS} \\ &- \text{PseudoPS} \times \text{pChancePsnMtch} \end{aligned} \quad (\text{B3})$$

We have values for all of the variables on the right side of the equation. The PsnMtchforPotentialPS and pChancePsnMtch values come from the position analysis: For each position scheme these values correspond to the observed number of position matches and the estimated proportion of position matches expected by chance, respectively. The value for PseudoPS comes from the perseveration analysis and does not depend upon the position scheme.

Consider, for example, the test-session results for the end-based scheme. For this scheme the position analysis yielded 1141 as the observed number of position matches across the

full set of potential perseveration–source pairs, and .2641 as the proportion of position matches expected by chance; the perseveration analysis provided 1314 as the estimated number of pseudo perseveration–source pairs. Substituting these values into Equation (B3), we obtain the estimated the number of position matches for true perseveration–source pairs:

$$\text{PsnMtchforTruePS} = 1141 - 1314 \times .2641 = 794 \quad (\text{B4})$$

Finally, we convert the number of position matches for true perseveration–source pairs to a proportion by dividing by the estimated number of true perseverations:

$$\text{pPMforTruePS} = \text{PsnMtchforTruePS} / \text{TruePS} \quad (\text{B5})$$

For the test-session data the estimated number of true perseverations (TruePS) was 1596, and so the end-based scheme’s estimated proportion of position matches for true perseverations was 794/1596, or .5. In other words, we estimate that in 50% of L.H.D.’s true perseverations the perseverated letter occupied the same end-based position in the perseveration response as in the true source response.

APPENDIX C

The prospective position analysis method

When applied to syllabic schemes, the position analysis method we used for evaluating extrinsic anchor and letter context schemes suffers from a self-fulfilling prophecy problem. For each potential perseveration–source pair, the method first determines what position the intruded letter occupies in the potential perseveration response and on that basis predicts where the source letter will be found in the source response. For extrinsic-anchor and letter-context schemes, the identity of an intruded letter plays no role in defining its position within the potential perseveration response. Consider, for example, the hypothetical error *slant* → SLAN?, where the question mark stands for an unspecified intruded letter. For any given extrinsic-anchor or letter-context scheme, the position of the intruded letter in no way depends on the letter’s identity. For instance, the intruded letter occupies position $E - 1$ according to an end-based scheme and position N_{-} by a preceding-letter scheme, regardless of what specific letter occupies the position.

For syllabic schemes, however, the position of a letter is determined in part by its identity. If, for example, the error were *slant* → SLAND, the intruded D would occupy a coda position in the first, and only, syllable ($\sigma_{B+1/E-1}$). However, in the actual error *slant* → SLANO, the intruded O occupies a nucleus position in the second syllable ($\sigma_{B+2/E-1}$). As this example illustrates, the identity of an intruded letter may play a role in determining both the syllable and the syllabic role occupied by the letter.

If we now use the position of the intruded O to predict the position of the source O in the source response, we are very likely to be successful, because the source O must by definition occupy a nucleus position. As long as the source O appears in

the first nucleus position of a syllable with position σ_{B+2} or σ_{E-1} , the perseveration and source positions will match by any orthosyllabic position scheme. For the *slant* \rightarrow SLANO error, the source response was SAVOL (a misreading of *savor*). The source O is in position N1 of syllable $\sigma_{B+2/E-1}$ and so matches the position of the O in SLANO. Clearly, this successful prediction of source position is something of a self-fulfilling prophecy; the position of the intruded O in SLANO is a nucleus position only because intrusion of an O makes the position a nucleus position.

To circumvent this problem, we developed a different position analysis procedure. Instead of using the location of the intruded letter in the perseveration response to predict the location of the source letter in the source response, we did the opposite, using the position of the source letter to predict the position of the perseverated letter. For purposes of explaining the method, we use a simplified syllable structure with two possible onset positions (O1 and O2), two possible nucleus positions (N1, N2), and two possible coda positions (C1, C2); however, the actual analyses encompassed three onset, three nucleus, and four coda positions. Consider the SLANO–SAVOL pair, and consider first the end-based representation of syllable position illustrated in Figure C1b. In the source response SAVOL, the O occupies the N1 position in syllable σ_{E-1} . Hence, the O unit in the N1 pool for syllable σ_{E-1} should be activated in the course of generating the response SAVOL (indicated in Figure C1b by the italicized O in the N1 column for syllable σ_{E-1}). When *slant* is then presented on the next trial, processing of the A should lead to activation of the A unit in the N1 pool for syllable σ_{E-1} . However, persisting activation from the O in SAVOL may compete with the bottom-up activation from the A in *slant*, leading to substitution of the O for the A. Therefore, the discrete syllabic position scheme predicts SLONT as a possible perseveration response.

Now consider beginning-based syllable position coding, illustrated in Figure C1a. The O unit in the N1 pool for syllable σ_{B+2} should be activated in generating the SAVOL response. When *slant* is presented, the corresponding letter units should be activated in syllable σ_{B+1} pools, and hence persisting activation from the O in SAVOL should not compete directly with any of the target letters. Nevertheless, this persisting activation could lead to perseveration of the O into the N1 position of syllable σ_{B+2} , yielding SLANTO. Accordingly, the orthosyllabic position scheme with both-edges coding of syllable position predicts that perseveration of the O from SAVOL into the response for the subsequent target *slant* should yield either SLONT or SLANTO. However, the actual response was SLANO, and hence the syllabic scheme fails to predict the position of the perseverated O for the SAVOL–SLANO pair.

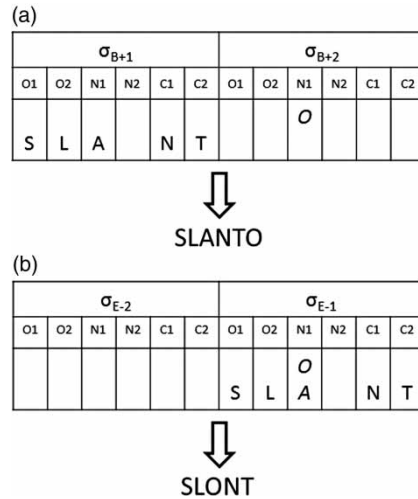


Figure C1. Example illustrating the synthesis of predicted perseveration responses, for a discrete orthosyllabic position scheme with (a) beginning-based and (b) end-based based coding of syllable position.

Note that this result contrasts with that obtained from the retrospective analysis method that predicts the source position of a perseverated letter based on position in the perseveration response. In the retrospective method, the identity of the perseverated O in SLANO defines the perseveration position as a nucleus position, leading to a position match between that O and the O in the source response SAVOL. Clearly, the result from the prospective method is the more sensible. There is no obvious reason for expecting that persisting activation of the O from SAVOL would compete with the T in *slant* and hence no obvious reason for expecting a perseverating O to replace the T. The synthesis method avoids the self-fulfilling prophecy problem.

In synthesizing predicted perseveration responses, we assumed that if the source position of the intruded letter was occupied in the orthosyllabic representation of the target, the target and perseverating letters would compete, resulting in a substitution of the perseverating letter for the target letter (as for the SLONT example in Figure C1b). If the source position of the perseverating letter was unoccupied in the target's syllabic representation, we assumed that the perseverating letter would be inserted at that position in the target response (as for SLANTO, Figure C1a).³ If any of the predicted perseveration responses matched the actual response, a correct prediction was tallied for the orthosyllabic scheme. To estimate the proportions of position matches expected by chance, we

³ We also considered an alternative synthesis method that generated insertions as well as substitutions when the source position was occupied. This method greatly increases the number of synthesized responses, but has no effect on the pattern of results. However, the alternative method might be appropriate for patients who (unlike L.H.D.) make many insertion errors.

synthesized predicted perseveration responses from source control responses (responses that had the same length as the source response but occurred on trials distant from the perseveration response).

Synthesizing predicted perseveration responses is straightforward when the erroneous response on the potential perseveration trial is a single-letter substitution (e.g., *slant* → SLANO) or insertion (e.g., *steady* → STREADY). However, for more complex errors (e.g., *abroad* → ABVONIT), synthesis of predicted errors requires arbitrary assumptions about the nature and interactions of error processes. For example, given the *abroad* → ABVONIT error, we cannot predict from a syllabic position scheme where a perseverating T from the preceding response *catch* should appear in the response to *abroad* unless we make arbitrary assumptions about the nature of the other errors contributing to the response (e.g., those giving rise to the N and I in ABVONIT) and the interactions of these errors with the perseveration. Consequently, analyses using the synthesis method were restricted to potential perseveration–source pairs in which the potential perseveration response involved only a single-letter substitution or insertion error. Pairs in which the potential source response could not be parsed into orthographically acceptable syllables (e.g., ASSITM) were also excluded. For the test-session and home-practice data sets, respectively, the analyses included 1501 and 3700 potential perseveration–source pairs, with 837 and 2262 of these pairs estimated to be true perseveration–source pairs.

We evaluated a discrete orthosyllabic position scheme, as well as two graded schemes. For the discrete scheme, we assumed that the source letter could perseverate only into an exactly matching orthosyllabic position in the target (e.g., position N1 in σ_{B+2} or σ_{E-1}), as discussed above for the SAVOL–SLANTO pair. For the graded role within-syllable scheme we assumed that coding of position was graded within a syllable’s onset, nucleus, and coda roles. According to this scheme, the O in SAVOL would be represented by activating the O unit strongly in the N1 pool for the appropriate syllable, but also by activating the O less strongly in the N2 pool. Figure C2a illustrates this activation pattern for the end-based representation of syllable position. In synthesizing predicted perseverations for the graded role within-syllable scheme we therefore assumed that a source letter could perseverate into any target position having the same basic syllabic role (onset, nucleus, or

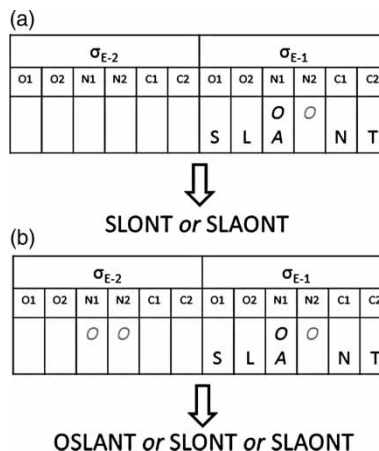


Figure C2. Example illustrating the synthesis of predicted perseveration responses, for (a) the graded within-syllable orthosyllabic position scheme and (b) the graded across-syllable scheme, with end-based coding of syllable position in both cases.

coda) within a matching syllable. For the SAVOL–SLANTO pair the analysis assumed that the O in SAVOL, which occupies position N1 in syllable $\sigma_{B+2/E-1}$, could perseverate into either the N1 or N2 position of syllable σ_{B+2} or σ_{E-1} . As shown in Figure C2a, a perseveration into position N2 of syllable σ_{E-1} would yield SLAONT, and hence the predicted perseverations for the graded role-within-syllable scheme included SLAONT as well as SLONT and SLANTO.

Finally, for the graded role across-syllables scheme we assumed that graded coding of syllable-role position extended to adjacent syllables. For example, we assumed that the O in SAVOL was represented by activating not only the O units in the N1 and N2 pools for syllables σ_{B+2} and σ_{E-1} , but also (to a lesser extent) the N1 and N2 units in the pools for immediately adjacent syllables (see Figure C2b). As a consequence, this scheme’s predicted perseverations for SAVOL–SLANO included not only SLONT, SLANTO, and SLAONT, but also OSLANT, resulting from perseveration of an O into a nucleus position of syllable σ_{E-2} .