



Published in final edited form as:

Cogn Neuropsychol. 2014 May ; 31(3): 237–265. doi:10.1080/02643294.2014.880676.

The analysis of perseverations in acquired dysgraphia reveals the internal structure of orthographic representations

Simon Fischer-Baum^{*} and

Department of Psychology, Rice University

Brenda Rapp

Department of Cognitive Science, Johns Hopkins University

Abstract

At a minimum, our long-term memory representations of word spellings consist of ordered strings of single letter identities. While letter identity and position must certainly be represented, it is by no means obvious that this is the only information that is included in orthographic representations, nor that representations necessarily have a one-dimensional “flat” structure. Evidence favors the alternative hypothesis that orthographic representations, much like phonological ones, are internally rich, complex multi-dimensional structures, though many questions remain regarding the precise nature of the internal complexity of orthographic representations. In this investigation, we test competing accounts of the internal structure of orthographic representations by analyzing the perseveration errors produced by an individual with acquired dysgraphia, LSS. The analysis of preservation errors provides a novel and powerful method for investigating the question of the independence of different representational components. The results provide clear support the hypothesis that letter quantity and syllabic role information are associated with, but separable from, letter identity information. Furthermore, the results indicate that digraphs, letter pairs associated with a single phoneme (e.g. the SH in FISH) are units of orthographic representation. These results contribute substantially to the further development of the multi-dimensional hypothesis, providing both new and converging evidence regarding the nature of the internal complexity of orthographic representations.

Keywords

Perseveration; spelling; orthographic representation

What do we know when we know the spelling of a word? It is generally assumed that orthographic information is mentally represented in a format-independent manner that represents the underlying commonality across letter formats (B, b, /bi/) and can be used by different input and output modalities (e.g., written and oral spelling). Following Camarazza and Miceli (1990), we refer to these abstract letter identities as graphemes.¹ The default hypothesis, the *linear string hypothesis*, assumes that our knowledge of word spellings includes only the identity of the graphemes in the word and their order. While letter segment

^{*}Corresponding author Department of Psychology, MS-25 Rice University P.O. Box 1891 Houston, TX 77251 Tel: 713-348-3841 simon.j.fischer-baum@rice.edu.

identity and position must certainly be represented, it is by no means obvious that this is the only information included in orthographic representations. The alternative hypothesis is that orthographic representations, much like phonological ones, are internally rich, complex *multi-dimensional structures* (Caramazza & Miceli, 1990; Badecker, 1996). Over the past 20 years, various lines of research, investigating both reading and spelling, have repeatedly demonstrated that orthographic representations are, in fact, more complex than would be predicted by the linear string hypothesis.

Despite general agreement regarding the internal complexity of orthographic representations, there continues to be widespread debate over the specifics of how these complex orthographic representations are structured. The goal of this paper is to move beyond demonstrating the limitations of the linear string hypothesis and to instead, evaluate a number of contrasting proposals the structure of orthographic representations in order to develop a more precise theory of orthographic knowledge.

Proposed modifications to the linear string hypothesis generally concern either the nature of the information associated with letter representations or they concern the units of orthographic representation. In terms of the types of information associated with letter representations, one view is that letter units are represented as multi-dimensional feature bundles, each of which, in addition to specifying the identity of the letter, indicates specific values on additional dimensions (e.g., Caramazza & Miceli, 1990; Badecker, 1996). Proposals differ with respect to representational dimensions that are included in these bundles. In this study we specifically examine claims regarding the representation of grapheme quantity and syllabic role. As an example, for the word BULL, the final L would be associated both with a feature indicating its syllabic role (here an orthographic coda, Badecker, 1996; Caramazza & Miceli, 1990; McCloskey et al., 1994; Ward & Romani, 2000) as well as with a feature indicating that the L is doubled ([Dbl]) (e.g., Caramazza & Miceli, 1990; Tainturier & Caramazza, 1996). In terms of the units of orthographic representation, the linear string hypothesis assumes that these correspond to simple single letter identities. However, a number of researchers have argued that digraphs – pairs of letters associated with a single sound, like the SH in BUSH – are represented as single orthographic units rather than as two adjacent letters (Tainturier & Rapp, 2004; Houghton & Zorzi, 2003; and also see Rey, Ziegler and Jacobs, 2000 for a similar proposal in the context of reading).

Below we review the evidence that has led researchers to posit the representations of grapheme quantity, syllabic role and digraph structure. We also summarize the various alternative explanations that have been proposed to account for the relevant data. In brief, the review will show that while current evidence clearly supports some additional complexity above and beyond the linear string hypothesis, new data are required to determine the more precise details of the internal structure of orthographic representations.

¹Elsewhere in the literature, the term *grapheme* has been used to describe any single letter or letter group that corresponds to a phoneme (B and SH are both graphemes, as they each map to a single phoneme, /b/ and /ʃ/ respectively). One of the goals of this paper is to examine whether multi-letter units like SH have a unitary representation at a level of abstract letter identity. We therefore will simply refer to the units at this level of representation as *graphemes* and consider the additional hypothesis that there are graphemes composed of multiple letters.

We will argue that our analysis of perseveration errors in spelling will provide an important contribution to this debate.

Evidence for the Representation of Letter Quantity, Syllabic Roles and Digraph Structure

The representation of double letters

Analysis of the errors produced by both dysgraphic (Caramazza & Miceli, 1990; McCloskey et al., 1994; Tainturier & Caramazza, 1996) and unimpaired individuals (Rumelhart & Norman, 1982) has suggested that double letters are represented differently than other pairs of adjacent letters (see also Cassar & Treiman, 1997 for evidence from children learning to spell). Several of these studies have found that errors that involve substituting both letters in a double (e.g. BALLET → BASSET) are more common than errors in which a single letter of the double is substituted (e.g. BALLET → BALKET). Double letter transpositions, where the wrong letter of the word is doubled (e.g. BALLET → BALEET), are also common. Based on detailed analyses of the dysgraphic errors produced by one individual, Tainturier and Caramazza (1996) found that errors with incorrect double letters were more likely in words with a double letter (e.g., BALLET → BALEET) than in words with consonant clusters (e.g. BASKET → BASKEET). This finding is difficult to reconcile with the linear string hypothesis that posits that double letters are represented no differently than other pairs of adjacent letters.

To account for these and other similar results, Caramazza and Miceli (1990) proposed that orthographic representations are multi-dimensional and include a dimension that represents “grapheme quantity.” Figure 1a shows the multi-dimensional representation of the word BALLET with three dimensions depicted, grapheme identity, quantity and position.² In this framework, the LL in BALLET is represented by a single bundle of features from the different dimensions of representation: the grapheme **identity** [L] is bound to the grapheme **quantity** [Dbl] and the grapheme **position** (here marked simply with the letter X, though see Fischer-Baum, McCloskey and Rapp, 2010 for a proposal regarding the nature of letter position representations).

Critical to the feature-bundling claim, information from one dimension can be lost or distorted without affecting the other dimensions. Within this framework, the double substitution errors (e.g. BALLET → BASSET shown in Figure 1b) can be explained by assuming a single substitution ([L] → [S]) in the grapheme identity dimension, while information in the quantity and position dimensions is unaffected. The geminate transposition error (e.g. BALLET → BALEET shown in Figure 1c) can be explained by assuming that, in the letter quantity dimension, the double [Dbl] feature transposes with a nearby single [Sgl] feature, resulting in the doubling of the wrong letter while, in this case, the identity and position dimensions are unaffected.

²In Figure 1, all graphemes are marked with either a [SGL] or a [DBL] in the grapheme quantity dimension. Alternatively, it is possible that [DBL] is marked in the grapheme quantity dimension, and [SGL] is the unmarked default quantity.

However, it is not clear that it is necessary to assume a quantity dimension in order to account for these patterns of double (geminate) shifts or substitutions. McCloskey and colleagues (1994) proposed an alternative explanation for these types of double letter errors (see also Badecker, 1996). In this proposal, a single grapheme identity is associated with multiple units in the grapheme position dimension. Figure 2a shows the representation of the word BALLET according to this alternative proposal, with single grapheme identity [L] associated with two positions. As before, double substitution errors (e.g. BALLET → BASSET shown in Figure 2b) can be explained by assuming a substitution at the grapheme identity level, without affecting the position dimension or the associations between the identity and position dimensions. The transposition error (e.g. BALLET → BALEET shown in Figure 2c) can be explained by assuming that the wrong grapheme identity ([E]) has become associated with two positions, resulting in the doubling of the wrong letter.

In both proposals, double letters are represented differently than other pairs of adjacent letters and, in this way, both are alternatives to the linear string hypothesis. However, they expand on the linear string hypothesis in different ways. One proposes an additional dimension of representation – the representation of grapheme quantity that is separate from both identity and position. The other includes only identity and position dimensions, but allows for many-to-one mappings between positions and identities. Given that both proposals can explain the existing data, additional data would be needed to adjudicate between them.

The representation of syllabic structure

Experimental evidence indicating a role for syllables in written language production presents another challenge to the linear string hypothesis. When children learn to spell they show sensitivity to syllabic structure, often deleting the second consonant in onset clusters and the first consonant in coda clusters (e.g. Read, 1986; Treiman, 1991; Treiman, Zukowski & Richmond-Welty, 1995). In addition, when adults and children write words, the inter-letter intervals are longer between letters in different syllables than between letters within the same syllable (e.g. Kandel & Valdois, 2006; Kandel, Alvarez & Vallee, 2006; Kandel et al., 2009; Zesiger et al., 1994). Finally, dysgraphic adults produce errors that show sensitivity to syllabic structure (e.g. Badecker, 1996; Caramazza & Miceli, 1990; Ward & Romani, 2000). For example, Caramazza and Miceli (1990) reported a dysgraphic individual, LB, who more accurately spelled words composed only of simple Onset-Nucleus syllables (e.g. TA/VO/LO) than words with more complex syllables (e.g. AL/BE/RO, SCAM/BI). Moreover, LB showed a different pattern of errors for heterosyllabic consonant clusters – clusters in which the two consonants are in different syllables – (e.g. PAR/CEL), than for homosyllabic consonant clusters – clusters in which the two consonants are in the same syllable (e.g. PA/STEL). None of these differences are predicted by the linear string hypothesis since it does not include a representation of a letter's syllabic role and would not predict that adjacent letters within a syllable should be represented any differently than adjacent letters that cross a syllable boundary, nor that more complex syllabic structures should be more difficult to produce than simpler ones.

Based on this evidence, Caramazza and Miceli (1990) proposed that orthographic representations contain an additional dimension that represents syllabic role. Figure 3a

shows the proposed representation of the word PARCEL with the grapheme identity and syllabic role dimensions. In this representation, the grapheme identity [C] is bundled with the [Ons] position, while the [R] is bundled with information specifying the [Coda] position. In contrast, Figure 3b shows the representation of the word PASTEL. Here both [S] and [T] are bundled with the [Ons] syllabic role. Given syllabic role representation, graphemic bundles can be grouped into orthographic syllables, each containing at least one nucleus that is optionally preceded by a sequence of onsets and/or followed by a sequence of codas. For example, in Figure 3a, PARCEL is syllabified into the syllables PAR and CEL, while PASTEL (Figure 3b) is syllabified into the syllables PA and STEL. Note that these syllables are orthographic rather than phonological in nature. For example, Figure 3c shows the representation of the word FACE. Phonologically, FACE is composed of a single syllable. Orthographically, both the [A] and the [E] are in the [Nuc] position, the [F] and the [C] are in the [Ons] position, and the word can be parsed into two syllables FA and CE. The syllabic effects in acquired dysgraphia that were mentioned earlier can be explained within this framework, given additional pressure to produce responses with the simplest possible syllabic structure (Caramazza & Miceli, 1990). Furthermore if we assume that the motor plans encode the hierarchical structure of the actions to be produced, then higher-level syllabic structure can be expected to influence the timing of the motor sequences.

However, it has been suggested that many of these syllabic effects can also be explained without assuming a representational dimension that encodes syllabic role. Jónsdóttir, Shallice and Wise (1996) argued that syllabic effects in spelling arise because of the syllabic structure of phonological rather than orthographic representations. Specifically, they argued that some dysgraphic individuals, when repairing damaged orthographic representations, make use of a coarse phonological representation of the stimuli, including such things as the number of phonological syllables in a word and whether or not it has an onset. The use of this type of phonological syllabic information could result in written responses that are sensitive to syllabic structure. Similarly, Marcel (1980) argued that children's spelling errors reflect a problem with phonological rather than orthographic representations. In other words, according to these accounts, the seeming sensitivity to syllabic structure in spelling arises not because orthographic representations include syllabic structure, but because phonological representations do.

There are several lines of evidence against this phonology-only account of orthographic syllable structure. Kandel and colleagues (2009) found that children showed similar timing patterns in their writing at orthosyllabic boundaries whether or not a word is phonologically bisyllabic (e.g. BARQUE: monosyllabic in the phonology vs. BALCON: bisyllabic in the phonology; both are orthographically bisyllabic). Based on evidence from letter-decision, syllabic segmentation and syllable counting tasks Chetail and Content (2012, 2013) argued that adults build orthographic syllables around all of the vowels in a visually presented letter string, even those vowels that are not realized in the pronunciation. Olson and Caramazza (2004) showed that deaf spellers also make errors that are sensitive to syllabic structure, despite having minimal or non-existent phonological representations. Given the pervasiveness of the phonology-only account, additional converging evidence would be important in clarifying the status of the syllabic role dimension in orthographic representation.

The representation of digraphs

Digraphs – multiple letters (most typically two letters) associated with a single phoneme – have also been claimed to have a special status in orthographic representations. Kandel and Spinelli (2010) found differences in the handwriting times of French words related to the presence/absence of digraphs. For example, they found that it took longer to produce the first letter of a digraph (e.g. the A in PRAIRIE) than a matched non-digraph letter (e.g. the A in CLAVIER) (see Kandel et al., 2006 for similar results with children). These systematic differences would not be predicted if the A in PRAIRIE was represented identically to the A in CLAVIER, as would be expected if digraphs did not have a special status. Further evidence comes from Tainturier and Rapp (2004). They reported on two individuals with acquired dysgraphia who had deficits in orthographic working memory (graphemic buffering), such that the probability of making an error on a letter increased with the length of the word. Tainturier and Rapp (2004) reasoned that if the graphemic buffer is sensitive to the number of orthographic units and if digraphs are indeed orthographic units, then digraphs should provide less of a burden than two-letter clusters to a damaged graphemic buffer. To test this hypothesis, they examined spelling accuracy for matched 2-letter clusters and digraphs (e.g., POCKET /BASKET). They found significantly more errors in the clusters than in the digraphs, consistent with the hypothesis that digraphs create less of a working memory burden. Furthermore, in terms of error types, this study found that constituents of digraphs were less likely to be separated than the constituents of clusters (see also Blanken et al., 1999). That is, it was less likely in digraphs than in clusters for only one of the constituent letters to be produced or for the two letters to be produced in non-adjacent positions. In other words, a digraph error such as POCKET → POCLER was less likely than the cluster error BASKET → BASLET). This latter result is consistent with the notion that letters are more tightly coupled in a digraph than in a cluster.

Evidence of this type has led some researchers to propose that, at some level of representation, the single or multiple letter sequences that (typically) correspond to single phonemes are represented as single orthographic units (Houghton & Zorzi, 2003; Tainturier & Rapp, 2004). (Note that these multiple letter sequences are referred to by some researchers as graphemes, by others as digraphs and yet others as complex graphemes.) Accordingly, the 6 letters of the 3 phoneme word “wreath” would be represented by the 3 (digraph) units WR+EA+TH, while the 6 letters of the 6-phoneme word “strict” would be represented by the 6 units S+T+R+I+C+T. By this hypothesis, the WR in WREATH enjoys a similar representational status as the S in STRICT, in the sense that, at least at some orthographic level of representation, it is an indivisible unit, with errors either affecting both the W and the R or neither of them.

Miceli & Capasso (2006) accounted for these patterns of performance with digraphs by arguing that orthographic representations are composed only of single letter units, but with varying amounts of “glue” holding together two adjacent letter units. Under this account, the letters that compose a digraph are bound together more strongly than other instances of adjacent graphemes but they are still composed of two graphemes, rather than a single unit. This orthographic texture hypothesis (see also Jones, Folk and Rapp, 2009) adds complexity to the linear string hypothesis (by positing varying degrees of connectedness between units)

and can explain the existing results without requiring the assumption that digraphs are represented as single units.

Summary of previous results

Across a range of methods and populations, the past twenty-plus years of research has clearly shown that orthographic representations are more complex than simple ordered strings of letters. What remains less clear is the precise content and structure of these complex orthographic representations. While a range of empirical results certainly can be explained by the hypothesis that orthographic representations contain dimensions representing letter quantity and syllabic roles, and that digraphs are represented as single units, the question remains whether all of these additional representational assumptions are actually necessary.

In our opinion, the debate should move away from treating the linear string hypothesis as the null hypothesis, and instead focus on adjudicating between alternative proposals regarding orthographic representation. The work presented in the paper attempts just that, presenting evidence and arguments that are directed towards developing a more complete and detailed theory of orthographic representation.

Perseverations as a tool to investigate orthographic representations

Perhaps the clearest diagnostic of separable dimensions is that the candidate dimensions can be affected independently of one another, in other words, that they show “separate modifiability.” This is a term that Sternberg (2001) used for the identification of independent processing modules and which Rapp & Fischer-Baum (in press) applied to the identification of separable dimensions of orthographic representation. Generally speaking, as well as in the specific context of orthographic representation, the separate modifiability of candidate representational dimensions may be manifested in a variety of ways, such as: differential effects on the different dimensions of manipulations of the levels of some experimental variable (e.g., frequency), differential susceptibility to disruption (or facilitation), or the possibility of independent “movement”. By this logic, the diagnostic characteristic of a single unit of representation is that the constituents of a candidate unit cannot be affected independently of one another; in other words, they show “inseparability” at least at the level of representation at which they operate as a unit. Inseparability can be expected to be manifested in many ways, for example, disruption should either affect all constituents or none of them.

In this paper, we present the case of an individual with acquired dysgraphia, LSS, who produced significantly more letters perseveration errors in his spelling than would be expected by chance. A perseveration error is the inappropriate intrusion of an element from a previous response into a subsequent response (see Table 1). The analysis of these errors provides a particularly powerful and compelling method for identifying the separable dimensions of orthographic representation under the simple logic that if information from one dimension of representation can be shown to “travel” from one response to the next without information from another dimension, then those two dimensions must be at least partially independent. Perseverations can also be informative for identifying units of

representation, based on the logic that if it can be shown that constituents either travel (or fail to travel) together across responses, then these constituents comprise a single unit, at least at the level of representation at which the perseveration errors arise. This investigation, by focusing on the analysis of perseveration errors, provides a novel source of evidence regarding the internal structure of orthographic representations.

The remainder of the paper is structured as follows. To provide a framework for the discussion, we first give a brief description of the cognitive architecture of spelling. We then provide background information on LSS, followed by four sets of analyses of his spelling errors. In the first, we establish that LSS produced perseveration errors at rates greater than chance and, furthermore, we show that these errors were generated at the level of abstract orthographic representation, ruling out alternative phonological or motor sources for the errors. In the second and third sets of analyses, we show that grapheme quantity and syllabic role can perseverate separately from grapheme identity. In the fourth set of analyses, we show that constituent letters of a digraph either perseverate together or not at all, while the letters in other consonant clusters perseverate independently of one another. Taken together, the results of these analyses provide evidence for the separable dimensions and units of orthographic representations.

Cognitive architecture of spelling

Most of the data we report come from the task of spelling to dictation (a word or nonword is dictated, and the participant produces a written spelling response). Figure 4 depicts a fairly standard cognitive spelling theory that specifies the representations and processes required for this task (see Tainturier & Rapp, 2001, and Miceli & Capasso, 2006).

According to this framework, when a familiar word (e.g., “table”) is dictated, the corresponding phonological word form is activated in the phonological lexicon. This word form, in turn, provides access to the word's lexical-semantic representation, which forms the basis for lexical selection – the retrieval of the corresponding orthographic word form in an orthographic lexicon. Some authors have also proposed direct connections between phonological and orthographic word forms (e.g., Patterson, 1986). The lexicons and semantic system correspond to long-term memory stores of knowledge of familiar words. Once an orthographic word form has been selected, it activates graphemic representations, by a process we will call grapheme encoding. These graphemic representations are assumed to be abstract in that they do not include information regarding the format of the letter to be produced. That is, the graphemic representation of TABLE includes the same T grapheme whether the word is to be typed, handwritten or spelled aloud. The goal of the current investigation is to understand how these graphemic representations are structured.

The phoneme-to-grapheme conversion process (PGC process) uses sublexical knowledge of the relationships between the sounds and letters (e.g., in English, /f/ is usually spelled with the letter F) to generate plausible spellings for phonological strings. This knowledge is applied when phonological or orthographic word forms or lexical-semantic representations are not available, such as in spelling an unfamiliar phonological string or when there is damage to these lexical components. In those cases, the phoneme-to-grapheme conversion process can generate a plausible graphemic representation

Abstract graphemic representations provide the basis for retrieving the format-specific letter representations needed for generating spelling responses. In oral spelling, letter-name representations are selected (e.g. /ti/ for the letter T) while, for written spelling, letter shapes corresponding to the intended case and font are selected. We assume that letter-shape and -name representations are selected serially. For this reason, graphemic representations are assumed to be processed by a working memory system referred to as the graphemic buffer (Caramazza, Miceli & Villa, 1987) or orthographic working memory (see Buchwald and Rapp, 2009) while letter shape or letter name selection takes place.

CASE REPORT

LSS has been reported on in several recent papers (Fischer-Baum, McCloskey & Rapp, 2010; Fischer-Baum & Rapp, 2012) which can be consulted for details of the case. Here we present a summary of the information that is most relevant for this investigation. LSS was a 58 year-old, left-handed man with a college degree who worked as a manager for a health services company prior to suffering a left middle cerebral artery infarct 33 months before this investigation. CT scans showed damage affecting the left inferior frontal gyrus, the inferior portion of the precentral gyrus, insula, putamen, caudate and the anterior pole of the left temporal lobe. He and his wife reported that he had normal pre-morbid spelling abilities and had used written language extensively at work.

As a result of the stroke, LSS suffered right-sided weakness, but no hemispatial neglect or visual deficits. Single word auditory comprehension was in the low range as evaluated by the PPVT (Dunn & Dunn, 1981) (score = 7th percentile). His word repetition was very good according PALPA Test# 9 (95% correct), though nonword repetition (PALPA Test# 8) was impaired (43% correct) (PALPA: Kay, Lesser & Coltheart, 1992). LSS had mild difficulties in the oral reading of single words: PALPA Test #31 (91% correct), and greater difficulties in reading nonwords (31% correct). He was 62% correct on the Boston Naming Test (Kaplan, Goodglass & Weintraub, 1983) which placed him significantly below age-matched controls. These word-finding difficulties were also evident in conversational speech. Visual memory was intact, with a score at the 84th percentile on the Visual Memory Index of the Wechsler Memory Scale (3rd edition; Wechsler, 1997).

A spelling evaluation revealed severe difficulties, with 19% accuracy in spelling words to dictation and 0% accuracy for nonwords. Errors for words and nonwords were similar in consisting primarily of nonword responses that were orthographically-related to the target and included a mix of letter substitutions, deletions, insertions and transpositions. LSS did produce some lexical substitutions but few errors that were phonologically plausible or semantically or morphologically related to the target. Letter accuracy was sensitive to lexical-level variables such as word frequency and concreteness, but not to word length. This contrasts with the typical pattern found in individuals with orthographic working memory deficits, and indicates impairment to the lexical spelling route (Buchwald & Rapp, 2009), most likely affecting grapheme encoding rather than lexical selection. Furthermore, his inability to produce a single phonologically plausible spelling of a nonword indicates a severe impairment to the phoneme-to-grapheme conversion system. Our current interpretation of LSS's spelling deficit is that he suffered impairments in both the lexical and

sublexical spelling routes that resulted in an inability to activate graphemes at the level of graphemic representation that is shared by both routes. In a previous investigation we also determined that LSS suffered an additional impairment that resulted in a failure to inhibit the graphemic representations activated during previous trials (see Fischer-Baum & Rapp, 2012 for discussion).

In the course of this investigation, LSS spelled 1636 words to dictation; when asked, he correctly repeated each stimulus word both before and after spelling it. This complete data set was used in all of the analyses reported below, unless otherwise indicated.

ANALYSIS 1: LETTER PERSEVERATIONS

LSS's spelling responses often contained letters that were not in the target stimulus and we refer to these as *letter intrusions*. Interestingly, many of the letter intrusions also appeared in preceding responses. A series of analyses were carried out to determine if the letter intrusions corresponded to letters produced in previous responses (perseverations) at rates greater than would be expected by chance and to determine whether the units of perseveration were graphemes, phonemes or motoric representations. These analyses are reported in a somewhat abbreviated manner here; see Fischer-Baum, McCloskey & Rapp (2010) for a more detailed description.

Analysis 1a: True perseveration errors?

Table 1 shows a sequence of target words and responses. On the error trial E, “under” was misspelled as UNDEL and the L is a letter intrusion. On the five trials prior to trial E (E-1 through E-5) the letter L appeared in responses on trials E-1 and E-2. If there was a causal relationship between the presence of an L in the preceding responses and the L-intrusion on trial E, then the L-intrusion would be a *perseveration error*, and the prior response it originated from would be the *perseveration source*. However, given that there are just 26 letters in the alphabet and that L is a fairly common letter, it is possible that the previous responses were not true perseveration sources, but instead contained the intruded letter simply due to chance. If at least some true perseveration errors were produced, the rate of observed perseveration sources should be greater than the rate expected by chance. To evaluate whether intruded letters appeared in the previous responses more than would be expected by chance, we applied an analysis method developed by McCloskey, Macaruso and Rapp (2006).

From the set of words LSS spelled to dictation, a computer program identified 251 single consonant intrusion errors in which the intruded consonants was surrounded by correct vowels or word edges (e.g., “easily” → EASIBY).³ For each intrusion error, the program then examined the preceding 5 trials (E-1 through E-5) and determined, across the set of intrusion errors, the proportion of responses at E-1 that contained the intruded letter. This corresponds to the proportion of potential perseveration sources. For intrusions with no potential source at E-1, it then considered the responses at E-2, and so on, through the

³These analyses are carried out over the identical corpus of spelling trials for LSS described in Fischer-Baum et al. (2010). However, we constrained the analysis here to the perseveration of only single consonant intrusions – leaving out vowel intrusions and multi-letter intrusions – as these results will serve as an appropriate comparison for the later analyses reported in this investigation.

response at E-5. The proportions of potential perseveration sources are shown in the solid line of Figure 5. As the graph indicates, the response immediately preceding an intrusion (trial E-1) contained the intruded letter 69% of the time, and the probability that an intruded letter first appeared in previous responses decreased with increasing distance from the intrusion.

To determine if LSS made any true perseveration errors we asked: Is 69% (and the other proportions observed at E-2 thru E-5) any greater than one would expect from a situation in which it is known that there is no relationship between responses produced at different time points? Our chance estimation procedure simulates just that situation and provides an estimate of the number of potential perseveration sources that are expected by chance due to the structure of the language, the types of words tested, propensities to produce certain letters, etc. It is based on the following logic: if there is no relationship between an intruded letter and the potential perseveration source, then the intruded letter should be just as likely to appear in a response that is not in the vicinity of the intrusion as it is to appear in a response that is in the vicinity of the intrusion. To determine this, the computer program identified, for each intrusion error, a pool of control responses matched to the actual responses produced at positions E-1 through E-5. Each pool of control responses consisted of all the responses of the same length as the actual responses in each of these positions and which were also produced during the same testing session⁴, but which were not produced within five responses of the intrusion error. From the pools of control responses for E-1 through E-5, the program constructed 10,000 randomly generated “control data sets”. The program then calculated the proportion of the actually intruded letters that appeared in each of the control data sets. The means of the 10,000 control data set values at positions E-1 to E-5 are shown in the dotted line of Figure 5, which also indicates the 95% confidence interval for each chance mean. This 95% confidence interval reflects the range of chance means for the middle 95% (or 9,500 control data sets) of the 10,000 control data sets. A comparison of the observed proportion of potential perseveration sources versus the distribution of chance proportions indicates that the observed values were greater than any of the 10,000 chance values for the two responses preceding an intrusion error ($p < .0001$). From this we conclude that at least some of LSS's intrusion errors had true perseverations sources in the preceding two responses.

Having established that LSS produced true perseverations, we then considered the type of information that was being perseverated. If these data are to be used to make inferences about abstract orthographic representations then we must establish that it was specifically graphemic information that was perseverated. Earlier, we argued, on the basis of the language and spelling assessments, that LSS' spelling errors resulted from impairment in processes that activate grapheme representations. To further support this conclusion, two

⁴Results similar to those reported in this paper were obtained when we allowed control responses to be drawn from the entire corpus of LSS's spelling errors. However, we believe constraining the control responses to the same testing session is a more conservative analysis. Consider a testing day in which LSS (for whatever reason) had produced (for whatever reason) a far larger number of Fs than usual. If control responses are taken only from that testing day, many of the control responses would contain the letter F. If the control responses are taken from the entire corpus, fewer of the control responses would contain the letter F. Therefore, for a single perseveration-source pair involving an F that was produced during the “F-intensive” session, an analysis that included control responses from the entire corpus might over-estimate the probability that it was a true perseveration.

additional tasks were administered to rule out two alternative perseveration sources: phonemes and motor-plans.

Analysis 1b: Phoneme persistence?

One could be concerned that phoneme persistence from trial to trial was responsible for the observed letter preservations. This is unlikely given that LSS could correctly repeat target stimuli both before and after spelling each word. However, to rule this out this we used a test developed by McCloskey et al. (2006). This involved two lists for spelling-to-dictation. Each list contained 24 consecutive word triplets: a “prime” word containing the phoneme /f/, followed by two target words that did not contain the phoneme /f/. In the F-List, the /f/ phoneme in the prime words was spelled with F (e.g. AFRAID), while in the No-F-List, the /f/ phoneme was always spelled in some other way (e.g. SPHERE). If it was the phoneme /f/ that was persisting, then the likelihood that LSS's responses would contain an intruded F should not depend on how that sound was spelled. However, if the grapheme F was persisting, then the likelihood of an intruded F should be greater for responses in the F-list than in the No F-list. The results show that LSS intruded the letter F significantly more often in his responses on the F-list (10/48) than on the No F-list (2/48), ($\chi^2=4.67, p<.05$), indicating that LSS's perseveration errors had a graphemic and not phonemic basis.

Analysis 1c: Motor-plan persistence?

A task, also from McCloskey et al. (2006), was administered to evaluate if perseveration errors were due to motor-plan persistence. In a spelling-to-dictation task (N=72 words), LSS was instructed to alternate his written responses from trial to trial between lower and upper case. If the letter perseverations occurred as a result of persisting graphemes, then the specific forms of the letters should not matter. However, if the persisting representations were motoric then they should be form-specific and we would expect intruded letters from the preceding response to be found at rates above chance only when the preceding responses were produced in the same case as in the error trial; in other words, one would not expect an “a” to appear as “A” in a later trial if the perseverated representations are motoric. We performed the perseveration analysis described above on this data set, focusing on only the immediately preceding trial that was produced in different case, and found that LSS showed a significant letter perseveration effect ($p<.0001$). Furthermore, the proportion of intrusions with a preservation source at E-1 was .71, a value comparable to the .69 proportion reported above for the larger data set (see Figure 5) at E-1, for a task in which all responses were produced in the same case. These results rule out the possibility that the perseverations involved form-specific representations such as motor plans.

Analysis 1: Summary

This set of analyses reveals that LSS produced true perseveration errors and that these errors originated at the level of abstract graphemic representation. In a recent paper, we investigated the nature of the deficits that resulted in these errors (Fischer-Baum & Rapp, 2012) and we concluded that LSS suffered both from: (1) a failure to activate the graphemic representations associated with the current target word from either the lexical or sublexical spelling routes, and (2) a failure to inhibit the graphemic representations produced during

previous response/s. Taken together, these deficits caused elements from previous responses to sometimes have higher activations than the under-activated elements from the current target, leading to elements from the previous response being selected and produced instead.

Analysis 1 demonstrated that single abstract grapheme identities from a previous response may sometimes be selected in the place of graphemes from the current target. We can also ask: Can other components of a complex multi-dimensional orthographic representation perseverate across responses? The next three analyses address this question

ANALYSIS 2: THE REPRESENTATION OF GRAPHEME QUANTITY

In this analysis, we consider whether information about grapheme quantity and grapheme identity ever perseverate independently. Consider LSS's error "tragic" spelled as TRRACE with the preceding response EXCESS. This error can be explained by assuming that a feature from the grapheme quantity dimension (Grapheme Quantity: [Dbl]) perseverated from the previous response. In other words, a letter was doubled in the response to "tragic" because a double letter was produced in the previous response. But since the identity of the letter that was doubled was different in the error ([R]) and the previous response ([S]), we could interpret this error as the perseveration of grapheme quantity without the perseveration of grapheme identity.

These "quantity-not-identity" perseveration errors are predicted by a theory of orthographic representation that posits separable dimensions of identity and quantity information such that it should be possible to disrupt information on one dimension without disrupting information on the other. For example, LSS's perseverative impairment could occasionally affect information in the quantity dimension (a failure to activate the quantity feature associated with a grapheme during the current trial, paired with a failure to inhibit quantity information from the previous trial), while leaving unaffected information in the identity dimension. Such a disruption should lead to *quantity-not-identity perseveration errors*. The separable dimensions hypothesis account also predicts *identity-not-quantity perseveration errors*, in which the perseverative deficit disrupts processing of information in the identity but not the quantity dimension (e.g., errors such as "butter" → BUZZER in which a double substitution is produced, with a preceding response such as LAZY that contains a single instance of the intruded double letter Z).

It is unclear how alternative theories of double letter representation (McCloskey et al., 1994; Badecker, 1996) would account for quantity-not-identity perseveration errors. As illustrated in Figure 2 these alternative theories include only grapheme identity and grapheme position dimensions, with doubles represented as a single identity associated with two positions. Consider again the error "tragic" → TRRACE with the previous response EXCESS. This error cannot be explained by the perseveration of grapheme identity information, as the error does not involve intruding graphemes from the previous response into the current trial. Nor can it be explained as a perseveration of grapheme position information – first because the response contains the appropriate number of positions and second because the double letter appears in different positions in TRRACE and EXCESS⁵. Of course, it is possible that the R in TRRACE was doubled for reasons other than the presence of a double letter in the

immediately preceding response. Therefore, as with the single letter intrusion errors, it is critical to establish that LSS produced these errors more often than would be expected by chance.

Analysis 2a: Quantity-not-Identity

For the quantity-not-identity analysis, a computer program identified all trials that contained a double letter in the response that did not also contain a double in the target. We refer to this as the set of *doubling errors* (e.g., “pardon” → PADDY). Note that these doubling errors included cases in which the erroneously doubled letter was already in the target (e.g., a single D appears in the target “pardon” and a doubled-D in the response PADDY). These could be included as this analysis only evaluates whether quantity (not identity) information perseverates between responses.

In total, LSS produced 155 doubling errors. For these, and subsequent, analyses we considered only the two immediately preceding responses as possible sources for the intruded elements. We did this because Analysis 1 indicated that only these responses were significant perseveration sources for LSS. We first removed from the analysis all doubling errors with a preceding response (either E-1 or E-2) containing the same geminate as appeared in the intrusion (“jacket” → JASST preceded by DRESS). These items do not permit testing whether grapheme quantity perseverates independently from grapheme identity, as the same grapheme identity appears in the same quantity in both the intrusion and (at least one of) the previous responses. The remaining 57 doubling errors consisted of items in which the E-1 and/or E-2 either did not contain a double letter (26/57), or contained a one that was different from the double in the error (PADDY preceded by KIRRY). The 31 items (54%: 31/57) in which a doubling error was preceded by response with a different double letter were potential quantity-not-identity perseveration errors.

Next, we determined whether LSS was producing quantity-not-identity perseveration errors at rates greater than would be expected by chance. We estimated the chance probability using the same logic as in Analysis 1. Control responses for each of the 57 critical doubling errors were identified for positions E-1 and E-2. These control responses might or might not have included a double letter, and they were subjected to the additional requirement that they not contain the same double letter as appeared in the intrusion error. For each of 10,000 control data sets, we calculated the proportion of the 57 doubling errors where one or both of the preceding control words contained a (different) double letter.

As indicated in Table 2, for LSS's actual data set, .54 of the doubling errors were preceded by trials containing a different double letter, while the average proportion for the control data sets was only .29. Moreover, in none of the 10,000 random data sets was a proportion as high as or higher than the observed proportion of .54 ever generated. From these results we can conclude that LSS produced “quantity-not-identity” perseveration errors at rates

⁵It may be possible to explain the error by assuming that what perseverates between one response and the next is the “fact” that there is a single grapheme associated with two positions, separate from information about either identity or the position of that double. However, if information about the fact that a grapheme is doubled is represented separately from both identity and position information, it is unclear how this is substantively different from an account that assumes separate dimensions of quantity, position and identity information.

greater than would be expected by chance ($p < .0001$), and therefore that, at least on some trials, LSS perseverated grapheme quantity independently from grapheme identity.

Analysis 2b: Identity-not-Quantity

The previous analysis demonstrated that LSS produced errors where grapheme quantity information was perseverated while grapheme identity information was not. In this analysis, we examined if grapheme identity can be perseverated in the absence of grapheme quantity. We first identified the set of errors with an intrusion of either a single or a doubled consonant. The single consonant intrusion errors were those used in Analysis 1a; the double consonant intrusion errors consisted of errors in which a double was intruded into a response although the doubled letter identity was not in the target (either as a single or a double) and (as for the single consonant intrusions) the double intrusion was surrounded by either word edges or vowels that appear in the correct spelling of the word. In all, LSS produced 251 single consonant intrusions and 27 double consonant intrusions.

For 217 of the set of 278 single and double consonant intrusions, the same letter appeared with the same quantity in either E-1 or E-2 or both. For example, immediately before the double-letter error “member” → MEFFER, LSS correctly spelled the word COFFEE. These errors can be interpreted as the perseveration of both grapheme quantity and identity, and therefore are not informative for determining whether grapheme identity can perseverate independently of quantity. For the remaining 61 errors, we can ask whether there is a mismatch between the quantity of the intruded consonant/s and the quantity of the same consonant/s in preceding responses — these are the *potential identity-but-not-quantity errors*. These 61 errors included cases in which the preceding responses contained the same consonant in a different quantity (the identity-not-quantity perseverations errors) as well as cases in which the preceding responses did not contain the intruded consonant at all. LSS produced 9 identity-not-quantity perseveration errors (9/61; 15%). For example, LSS spelled “butter” as BUZZER immediately after producing the response LAZY, or spelling “while” as WHIDE immediately after the response CLIDD. In the former example, the doubled Z intruded into the spelling of the word, but the immediately preceding response contains only a single Z, suggesting that quantity did not intrude. In the latter, example a single D intruded into the spelling of word, while the previous response contained a double D.

The observed proportion of identity-not-quantity errors was then evaluated relative to chance, with the analysis carried out over the 61 intruded consonant/s where the same letter did not appear in the same quantity in either E-1 or E-2. For each intrusion, we identified a set of source control responses for both source response E-1 and E-2. These source control responses were selected in the same way as in Analysis 1, with the additional requirement that they should not contain the same letter-identity in the same quantity as the error. A Monte Carlo analysis created 10,000 control data sets composed of each intrusion error paired with random source control responses for the E-1 and E-2 response. A program then tabulated the proportion of these intrusion errors where the same letter identity appeared in a different quantity in either E-1 or E-2 (or both).

The results (see Table 2) indicated that observed value exceeded the value derived from all 10,000 control data sets. From this analysis, we conclude that LSS produced more “identity-not-quantity” perseveration errors than would be expected by chance.

Analysis 2: Summary

This set of analyses demonstrated that LSS perseverated grapheme quantity and grapheme identity separately more often than would be expected by chance. That is, for at least some of the doubling errors, there is a causal relationship between the presence of a double letter in the previous response and the intrusion of a (different) double letter in the current response. Similarly, there are some true perseverations of grapheme identity information even when grapheme quantity information does not perseverate. These results are predicted by theories that assume separable dimensions of grapheme quantity and identity (e.g. Caramazza & Miceli, 1990) and are difficult to reconcile with theories that do not include a quantity dimension in the orthographic representations (e.g., McCloskey et al., 1994).

For many of LSS doubling errors, the preceding responses did contain the same doubled letter (98/155; 63%) and for many of LSS's consonant intrusions, the preceding responses contained the intruded letter in the same quantity (217/278; 78%). Taken together these results suggest that while identity and quantity information can be disrupted separately, they often are disrupted together. This result is consistent with the claim that orthographic representations can be thought of as bundles of features from different dimensions and that, as a result, it is more likely that features from different dimensions will perseverate together than that they will perseverate independently.

ANALYSIS 3: THE REPRESENTATION OF SYLLABIC ROLE

The next set of analyses considers whether there is an additional dimension of syllabic role information that is separable from grapheme identity information (Caramazza & Miceli, 1990). These analyses follow the same logic as the previous analyses: if syllabic role information and grapheme identity information are separable dimensions, then we expect syllabic role to perseverate separately from grapheme identity (*role-not-identity perseveration errors*) and grapheme identity to perseverate separately from syllabic role (*identity-not-role perseveration errors*).

Analysis 3a: Role-not-Identity perseverations

The analysis considers errors in which syllabic role but not grapheme identity is intruded. In the error “cause” → COUST, the letter T is produced in the place of the final E in the error. This error can also be interpreted as the intrusion of a syllabic role because, as shown in Figure 6, the response differs in syllable structure from the target. In other words, no letter in the target CAUSE corresponds the coda of an orthographic syllable, while in COUST, the T is a coda position and, thus, we could characterize this error as the intrusion of the syllabic role “coda.” The previous response, DOOR, contains the coda consonant R, allowing for the possibility that the intruded coda in COUST might have been a perseveration of syllabic role (coda) from the previous response. Given that the previous response does not contain the letter T, if syllabic role did perseverate, it perseverated independently of grapheme identity.

To determine if these errors occurred at rates greater than would be expected by chance, we identified responses containing either a word-initial onset or a word-final coda when no such syllabic role appeared in the target spelling (e.g., “owl” → HOPWOL) or “cause” → COUST). We will refer to these errors as *potential syllabic role intrusion errors*. We restricted this analysis to the intrusion of word-initial onsets and word-final codas because there is no clear consensus regarding syllabification of word-medial consonants. Some have argued for an onset maximization approach to syllabification, in which as many letters as possible are assigned to syllable onsets (e.g. Badecker, 1996), while others have argued for a coda-maximization approach, in which as many letters as possible are assigned to syllable codas (e.g., Taft, 1979). We further restricted our analysis so that the letter identity that appeared in the intruded syllabic role appeared nowhere in the immediately preceding response. Since the error “owl” → HOPWOL was preceded by the response HEOPCON, the H intrusion could not be a role-not-identity perseveration error, and therefore would not be included in the analysis. In all, LSS produced 27 potential syllabic role intrusion errors in which the intruded letter did not appear in the immediately preceding response.⁶

As indicated in Table 3, for 22 of LSS's 27 syllabic role intrusion errors (81%) were potential role-not-identity perseveration errors as the immediately preceding response contained the same syllabic role (either word-initial onset or word-final coda). We carried out a chance analysis similar to those described above to determine if LSS produced role-not-identity perseverations errors at rates greater than would be expected by chance. Briefly, a computer program identified a pool of source control responses for each of the 27 syllabic role intrusions, following the same constraints as in Analysis 1 with the additional constraint that control responses not contain the letter that appeared in the intruded syllabic role. In a single run of a Monte Carlo analysis, the program randomly selected a single source control response for each of the intrusion errors and calculated the proportion of intrusion errors in which the control response contained the intruded syllabic role. In all, 10,000 runs of the Monte Carlo analysis were carried out. In only 267 of the 10,000 analysis runs did we obtain a value as high as or higher than the observed proportion of .88. We can therefore conclude that role-not-identity perseveration errors occurred at rates greater than would be expected by chance ($p < .05$). In other words, even though word-onsets and word-coda consonants are highly likely in LSS's responses, they are even more likely for the trial preceding an intrusion of one of these syllabic positions than for a trial drawn at random. This result indicates that syllabic role, or at least word-initial onsets and word-final coda roles, can perseverate independently of the grapheme identity that appears in that role.

Analysis 3b: Identity-not-Role

A second analysis examined errors in which grapheme identity was perseverated independently of syllabic role. Consider the error “wheat” → WHEAL preceded by the

⁶Unlike the other analyses reported here, we restricted our window of perseveration to a single preceding trial (E-1) rather than the two immediately preceding trials. Expanding the analysis to a larger window reduced the power of our analysis as it was very likely that one of two randomly chosen responses would contain word-initial onsets or word-final codas, increasing chance proportions to near ceiling levels. By limiting our window of perseveration to a single preceding trial, we lower the probability of observing word-onsets and word-codas in the perseveration window by chance. This more limited analysis allows us to evaluate whether word-onsets and word-codas appear in previous responses more than would be expected by chance without concerns about ceiling effects. Nonetheless, it is worth noting that when the larger perseveration window was used, the result approached significance ($p < .1$).

response HOTLEN. On trial E, the intruded L appeared in the coda position while on trial E-1, the L appeared as the onset of the second syllable. This error is an example of a potential identity-not-role perseveration error. If syllabic role is a separable dimension from grapheme identity, LSS should have produced more identity-not-role perseveration errors than would be expected by chance.

A computer program syllabified the complete set of LSS's targets and responses assigning each grapheme in each response to a syllabic role. This program assigned syllabic role using principles of onset maximization, assigning the maximum number of letters to a syllable onset such that principles of English orthotactics are not violated (see McCloskey, Macaruso & Rapp, 2006 for further discussion of this syllabification program). Strings that could not be syllabified into orthotactically legal syllables were discarded from the analysis (a total of 152 or 9% of all responses). We then reanalyzed the set of single consonant intrusion errors described in Analysis 1. From these single consonant intrusion errors, we identified 50 intrusions errors in which that intruded letter did not appear in the same syllabic role in either E-1 or E-2. These intrusions included both *potential identity-not-role perseveration errors*, in which the intruded letter appeared in a different syllabic role in a prior response, as well as intrusions that did not appear in either of the two prior responses. We then calculated the proportion of the total set of intrusions that were *potential identity-not-role perseveration errors* as well as the proportion of these that would be expected by chance.

As indicated in Table 3, 13 of the 50 (26%) of the intrusions in which the intruded letter did not appear in the same syllabic role in either E-1 or E-1, (26%), were *potential identity-not-role perseveration errors*. We determined the likelihood that these errors arose by chance using the same analysis techniques as above; here the source control responses were restricted to those responses that did not have the same graphemes in the same syllabic role as each intrusion error. A proportion as high as or higher than the observed proportion of identity-not-role perseveration errors was obtained in only 36 out of 10,000 chance analysis runs, indicating that LSS perseverated the identity of a grapheme independently of its syllabic role more often than would be expected by chance ($p < .01$).

Analysis 3: Summary

The results of this set of analyses suggest that syllabic role and grapheme identity information can perseverate separately. These results are predicted by theories that assume separable dimensions of grapheme quantity and syllabic role information (Caramazza and Miceli, 1990) and it is unclear how alternative explanations based on the influence of the phonological representations associated with the target word (e.g., Jónsdóttir, Shallice & Wise, 1996) could explain the syllabic role perseverations. First, the relevant errors are those in which the response had a different syllabic structure than the target, both orthographically and phonologically. If phonological information about the previous target influenced spelling performance, that should have worked against the production of these types of errors. Second, in Analysis 1 we demonstrated that LSS's errors were perseveration of graphemes and not of phonemes, making it unlikely that information regarding phonological syllabic roles from a previous trial would influence the upcoming spelling response. In sum, the results of the “role-not-identity” and “identity-not-role” analyses are difficult to

reconcile with the alternative proposals that lack a representation of orthographic syllabic role that is separable from the representation of grapheme identity.

The results of this analysis indicate that orthographic representations include information about syllabic structure. It is noteworthy that syllabic structure information has been posited to play a central role in representing grapheme order in spelling (e.g. Houghton & Zorzi, 2003). We will return to this issue in the General Discussion.

ANALYSIS 4: UNITARY DIGRAPH REPRESENTATIONS

In a final set of analyses, we used LSS's perseveration errors to test the hypothesis that digraphs are represented as single units in the grapheme identity dimension. We reasoned that if digraphs are represented as single, indivisible units, they should perseverate like other single, indivisible units and their constituent letters should not perseverate individually. In contrast, if digraphs are represented as two adjacent units, they should perseverate at rates comparable to other adjacent units (e.g., consonant clusters), and their constituent letters should sometimes perseverate individually. To examine these predictions, we first examined whether digraph perseveration rates were comparable to those observed for single consonants or to those observed for adjacent consonants. Second, we examined whether single consonant perseverations originate in digraphs or from consonant clusters at rates greater than would be expected by chance.

Analysis 4a: Digraph Perseverations

LSS occasionally made *digraph intrusion* errors in which he intruded a digraph into the spelling of the word (e.g., “soul” spelled as SOUCK). He also occasionally made *adjacent consonant intrusion* errors in which he intruded adjacent consonants that did not comprise a digraph into the spelling of word (e.g., “junk” spelled as JUMP). If digraphs are represented as single units, the digraph intrusion errors involve a single grapheme identity, while the adjacent consonant intrusion errors involve two adjacent grapheme identities. If so, we would expect the proportion of digraph intrusions in which the digraph perseverates from the previous response to be comparable to the proportion of single letter intrusions in which a single letter perseverates from a previous response. Further, in contrast to digraphs, adjacent consonants should perseverate from previous responses at rates different than those observed for single letters.

A computer program identified all intrusions of two adjacent, non-identical consonants into a response, surrounded either by word edges or vowels in responses in which both consonants appeared in the same orthographic syllable⁷. These intrusions were then classified as either digraphs or adjacent consonants. The program then determined the likelihood that the intruded letters were adjacent either in response E-1 or E-2, as well as the likelihood that they would be expected to be adjacent by chance. In all, LSS produced 23 digraph intrusions and 35 adjacent consonant intrusions.

⁷We restricted intrusions to those that appeared in the same syllable because of the possibility that pairs of consonants that are pronounced with a single phoneme when they appear in the same syllable (e.g. the MB in LAMB), may be pronounced with two phonemes when the same letter pair crosses a syllable boundary (e.g. the MB in AM.BER), creating some ambiguity as to whether or not they would be represented as digraphs by the orthographic system.

To evaluate the relevant predictions, it was necessary for this analysis (unlike for the previous ones) to compare perseveration proportions for different types of intrusions (digraphs, adjacent consonants, single consonants), all of which have different chance rates. To make these comparisons we need to calculate, for each intrusion type, what has been referred to as the *true perseveration proportion*, a value that incorporates the specific chance rates for each intrusion type (for a full description of this calculation see Appendix 1 or Fischer-Baum, McCloskey & Rapp, 2010). The true perseveration proportion values for each intrusion type are reported in Table 4. Digraph intrusions are the most likely to be true perseverations either from E-1 or E-2 (85%), followed by single consonant intrusions (69%) and then adjacent consonant intrusions (46%). These values alone suggest that digraph intrusions behave more like single consonant intrusions than like adjacent consonant intrusions, as would be predicted by the proposal that digraphs are represented as single units. These true perseveration proportions were evaluated statistically by comparing the values for digraphs and adjacent consonants to each other and to the true perseveration value for single consonant intrusions using a Monte Carlo permutation analysis which evaluates whether the magnitude of the observed differences in true perseveration proportions are greater than what would be expected across a range of random samples from the relevant error set. For the Monte Carlo permutation test, for each of 10,000 analysis runs, we combined digraph and adjacent consonant intrusions into one set and then randomly resplit the set into two groups, each the size of observed groups (e.g. 23/35 split for digraphs vs. adjacent consonants). We then calculated the difference in the true perseveration proportions for each of the two randomly reassigned groups and then calculated the proportion of the 10,000 runs in which the difference in true perseveration values for the two groups was as high as or higher than the observed difference. We rejected the null hypothesis if a 2-tailed $p < .05$ was obtained.

Using the same resplitting procedure with different sets of intrusions errors, we found significant differences in true perseveration rates between single consonant intrusions and adjacent consonant intrusions ($p < .05$), as well as between digraph intrusions and adjacent consonants ($p < .05$), but not between single consonant and digraph intrusions. The results of these analysis showed that the true perseveration proportion for digraph intrusions was not significantly different than the proportion observed for single consonant intrusions, but that both single and digraph intrusions were significantly more likely to be true perseverations than were adjacent consonant intrusions.

We also compared the true perseveration proportion of digraph and adjacent consonant perseverations with the expected joint probability that two single consonant intrusions were both perseverations. To calculate the expected probability that two adjacent consonant intrusions were both single perseverations from either E-1 or E-2 or both, we assumed that the likelihood that one intruded letter was a perseveration was independent of the likelihood that a second intruded letter was a perseveration. On this basis, we calculated the predicted probability as follows: The probability that two intrusions were perseverations from E-1 was taken to be the square of the probability that a single intrusion was a perseveration from E-1, and the same calculation was carried out for E-2. To estimate that the probability that two intruded letters had perseverated together from either E-1 or E-2 or both, we simply added

our estimates for E-1 and E-2 (the squares of the single intrusions for each) and subtracted the product of our estimates for E-1 and E-2. Given this calculation, we estimated the predicted proportion that two adjacent intrusions could both be perseverations from either E-1 or E-2 or both to be .44.

Monte Carlo permutation tests were carried out comparing the true perseveration proportions for digraph and adjacent consonant intrusions to this predicted proportion. This analysis used a similar random resplitting approach as in the analysis described above. For example, for the comparison of digraphs to the predicted probability of two adjacent consonant intrusions based on the rate of single consonant intrusions, all digraphs ($n=23$) and single consonant ($n=251$) intrusions were combined into one set, and then randomly resplit into two groups each the size of the observed groups (a random 23/251 split), with each group including both single consonant and digraph intrusions. This random resplitting was carried out in 10,000 runs of the analysis program, and for each run, the program calculated the difference in the true perseveration proportion between the two randomly resplit groups. We then calculated the proportion of the 10,000 runs in which this difference in perseveration proportion for the two groups was as high as or higher than the observed difference between the actual perseveration proportion calculated for digraph intrusions (.85) and the expected joint probability that two single consonant intrusions were both perseverations calculated based on the set of single consonant intrusions (.44). We rejected the null hypothesis if a 2-tailed $p < .05$ was obtained. An identical analysis was carried out with adjacent consonant intrusions. The results of these analyses indicated that digraphs were more likely to be perseverations from the same response than would be predicted by assuming two independent single consonant perseverations ($p < .05$), but that adjacent consonants were no more likely to be perseverations than predicted by our joint probability model.

The results of these analyses are precisely as would be predicted if digraphs were represented as single units in the grapheme identity dimension. Namely, the digraph intrusion errors exhibit a pattern similar to that of the single consonants rather than the adjacent consonants (either the observed or predicted proportions). These results are difficult to reconcile with a theory in which digraphs are represented as two adjacent grapheme identities, even one that proposes uneven texture in the orthographic representation (e.g. Miceli & Capasso, 2006). In contrast, other adjacent consonants intrusions are no more likely to have perseverated together than would be predicted by a model that assumes independence between the consonants.

Analysis 4b: Digraph-Decomposition

A second analysis was carried out to further test the digraphs-as-units hypothesis. If digraphs are represented as single, non-decomposable units in the grapheme identity dimension, then we would not expect the perseveration of constituent letters of a digraph into subsequent responses. Consider the sequence of trials, “fresh” → FRESH followed by an attempt to spell the word “cat”. If the SH in FRESH is represented as a single unit, we would expect either the whole digraph to perseverate (e.g. “cat” → CASH) or neither constituent of the digraph to perseverate (e.g. “cat” → CAT). We would not expect errors in

which only one of the two constituent letters of the digraph perseverated (e.g. “cat” → CAS; “cat” → CAH). In contrast, since consonant clusters that are not digraphs (e.g. SK in FLASK) are represented by two grapheme identities, we would expect the S and the K to sometimes perseverate independently into subsequent responses.

To test these predictions, we evaluated whether single consonant intrusion errors (e.g. the S in “cat” → CAS) are more likely than expected by chance to be perseverations from digraphs produced in the previous response (e.g. FRESH), and also whether they are more likely than expected by chance to be perseverations from consonant clusters produced in the previous response (e.g. FLASK).

To carry out this analysis, we identified the subset of single consonant intrusions used in Analysis 1a that could be a constituent of either a cluster or a digraph – that is all consonant intrusions except for the letters J, Q, X, V and Z, as these letters cannot serve as constituents for clusters or digraphs attested in English⁸. In all, 178 such consonant intrusions were identified. We then calculated the likelihood that these single consonant intrusions appeared in either E-1 or E-2 as a part of a digraph or as a part of a consonant cluster. As indicated in Table 5, only 3% of the single consonant intrusions appeared in digraphs in either E-1 or E-2, while 19% appeared in consonant clusters in either E-1 or E-2. We then compared these observed proportions to the proportions expected by chance, using the analysis procedures applied in Analysis 1a. We found that single consonant intrusions were no more likely than would be expected by chance to appear in digraphs in E-1 or E-2 (Chance Mean: 3%, $p > .5$) but were significantly more likely than would be expected by chance to appear in a consonant cluster in E-1 or E-2 (Chance Mean: 10%, $p < .001$). From this analysis it is evident that the constituent letters of digraphs do not perseverate alone, while individual letters of consonant clusters do.

Analysis 4: Summary

These two analyses support the hypothesis that digraphs are represented as single, non-decomposable units at least at some level of orthographic representation. Evidence for the unitary nature of digraphs comes from: (1) the finding that intruded digraphs perseverated at rates comparable to those observed for single letters, while other adjacent consonants perseverated at rates that would be predicted for two independent single letter perseverations and (2) the finding that digraphs, unlike other consonant clusters, were no more likely than chance to have been the source of single consonant intrusions.

GENERAL DISCUSSION

In this investigation, we first showed that an individual with acquired dysgraphia produced spelling responses containing incorrect letters that were related to previous responses -- true perseverations errors. The analysis of the perseveration errors, along with other lines of evidence in the literature, challenge the view that orthographic representations constitute a

⁸Consonant clusters that could either be digraphs or not depending on the word (e.g. SC in SCOUT versus SCIENCE) were excluded from the analysis. The set of possible consonant clusters was: {bl; br; cl; cr; ct; dr; fr; ft; gl; gr; kl; kr; lb; ld; lg; lk; lp; lt; mp; nd; nk; nt; pl; pr; pt; rb; rd; rf; rg; rk; rl; rm; m; rp; rt; sk; sl; sm; sn; sp; st; sw; tr; wn}. The set of possible consonant digraphs was: {ch; ck; dg; gh; gn; kn; mb; mn; ng; ph; ps; sh; th; wh; wr}

simple concatenation of letter identities, providing support instead for a specific version of the multi-dimensional representational hypothesis. The proposed multi-dimensional orthographic representational system includes syllabic role, grapheme quantity and grapheme identity dimensions, as well as the representation of digraphs as single units in the grapheme identity dimension. To illustrate the proposal, the representation of the word SHALL is depicted in Figure 7. The digraph [SH] in the grapheme identity dimension is associated with the quantity [Sgl] in the grapheme quantity dimension, and the role [Ons] in the syllabic role dimension. The grapheme [A] in the grapheme identity dimension is associated with the quantity [Sgl] in the quantity dimension and the role [Nuc] in the syllabic role dimension. The grapheme [L] in the grapheme identity dimension is associated with the quantity [Dbl] in the quantity dimension and the role [Coda] in the syllabic role dimension. Orthographic representations also presumably contain additional dimensions beyond those discussed in this paper. For example, it has been shown that orthographic representations contain a consonant-vowel dimension (Buchwald & Rapp, 2006) and grapheme position information (Fischer-Baum, McCloskey & Rapp, 2010). A full theory of orthographic representation incorporates all of these representational components (Rapp & Fischer-Baum, in press).

Perseverations reveal orthographic structure

Analyses 1B and 1C demonstrated that LSS's perseveration errors involved graphemic representations and not phonemic or motor plan representations and, according to the spelling theory outlined in Figure 4, the level of graphemic spelling representations are activated by both the lexical and the sublexical routes (Folk & Rapp, 2004; Rapp, Epstein, & Tainturier, 2002). Lexical processing involves lexical representations of word spellings activating their corresponding to letters at the graphemic level of representation. By this account, knowledge about the double L in the word SHALL is encoded in the connections between the lexical representation [SHALL] at the level of the orthographic lexicon, and grapheme representations of the letter L and the quantity [Dbl]. The representations of L and the quantity [Dbl] could also be activated via the sublexical route, by processes that determine the likely sequence of graphemes given a sequence of phonemes.

Within this framework, the interpretation of LSS's deficit is relatively straightforward. Fischer-Baum and Rapp (2012) argued that LSS produced letter perseveration errors because grapheme representations from previous responses persisted abnormally in an activated state after they were produced and also because the activation of a target's graphemes was sometimes abnormally weak. We have argued here that accounting for the detailed set of results reported in the current investigation, requires a number of representational assumptions, namely that: the perseverating representations include not only letter identity representations, but also quantity and syllabic role representations and also that digraphs are represented as single perseverating units. Given these representational claims, quantity and syllabic role information can persist from a previous response even if the activation of letter-identity information does not, causing the observed quantity-not-identity and role-not-identity perseverations described above. The representation of digraphs as single units in the letter-identity dimension provides an explanation for our finding that either both letters in the digraph perseverate together or neither of them do.

While this proposal assumes a shared level of orthographic representation for lexical and sublexical spelling processes, it is possible that there may be different types of representational structure in different components of the spelling system. For example, some researchers make a distinction between long-term and working memory graphemic spelling representations (e.g., Houghton & Zorzi, 2003) and, accordingly, one could posit different types of orthographic representational structure for these different systems. Therefore, while our results clearly favor this complex, multi-dimensional orthographic representation at some level of graphemic representation, we cannot rule out the possibility that orthographic information is represented differently in different components of the spelling process.

Separable features and feature bundles

While the present analyses highlight the separability of the different representational components of orthographic representation – identity, quantity, syllabic role, position and (based on previous work) consonant/vowel status – we would like to be clear that we are proposing that these representational components are, nonetheless, bound together within multi-dimensional orthographic representations. In fact, although the various features can be separately modified, they are more likely to travel together than they are to “go their separate ways”. For example, although we have shown that a doubling feature may travel from one word to the next independently of letter identity, the results of Analysis 2A show that it is, in fact, more common for the double feature and letter identity to travel together. Another example of this type of boundedness comes from Fischer-Baum et al. (2010), who showed that the position of LSS's letter-identity perseveration errors tended to match the position of their respective source letters. This pattern is consistent with the idea that letter identity and position features are bound together and are more likely to travel together than to travel apart. Along these lines, we would predict that perseveration of information from other dimensions should also match in terms of the position of the source letter more often than would be expected by chance; for example, the position of the double letter in LSS' quantity-not-identity perseverations (e.g. PADDY preceded by KIRRY) are likely match the position of the double in the source response. Unfortunately, LSS produced too few quantity-not-identity perseverations to test this prediction.

The possible role/s of graphosyllabic representations

It is worth noting that there may be some redundancy across the set of dimensions of orthographic representation that have been proposed. For example, it is possible that evidence that has been proposed for the consonant/vowel distinction could be accounted for the syllabic role dimension argued for here and in previous work (Caramazza & Miceli, 1990; Badecker, 1996) given that all consonants are associated with either an onset or coda syllabic role, while vowels can only be associated with the nucleus syllabic role. Whether consonant/vowel representations can be reduced to syllabic role representations or whether there is a role for both representational types will require further investigation.

It has also been argued that the syllabic role dimension could be used to represent letter position, eliminating the need to posit other types of position representations. In a graphosyllabic position representation scheme, a letter's position is defined relative to both the syllable in which it appears and its role within the syllable. However, based on an

analysis of the same corpus of spelling errors produced by LSS that were reported on here, Fischer-Baum et al. (2010) found no evidence that graphosyllabic schemes served to represent letter position in spelling. Instead, they showed that the position LSS' letter-identity perseverations tended to match the source letter's position according to a representational scheme in which a letter's position is defined by a distance, in number of letters, from both the beginning and the end of the word. Nonetheless, the current findings do provide clear evidence of graphosyllabic role representations.

If syllabic role information is sufficient for representing letter position, why should both syllabic roles and edge-based position be represented? An answer to this question may be forthcoming if we consider the computations that operate over graphemic representations in spelling. According to the theory outlined in Figure 4, the level of graphemic representation provides input to a process that serially selects letters for production in the correct order. However, the complex orthographic representations proposed here require a more sophisticated serial selection system than would be required by the linear string hypothesis. For example, the system must operate such that given a doubling feature in the quantity dimension, whatever the grapheme identity is, it must be produced twice. Along similar lines, given a unitary digraph representation, two letter-shapes must be produced in the correct order. Various sophisticated serial selection systems in spelling have been proposed to navigate these types of complexity (e.g., Houghton, Glasspool & Shallice, 1994). Interestingly, the solution Houghton and colleagues provided for this problem – the Competitive Queuing computational model of spelling – requires a both-edges coding of letter position to accurately produce letters in order. Therefore, a both-edges representation of letter position may be critical in allowing for the accurate serial selection of letters.

However, serial letter selection may not be the only computation operating at this level of the spelling system. Badecker (1996) argued that orthotactic well-formedness computations also operate at this level. These increase the likelihood that a spelling response will be a possible or “legal” string and, unlike the serial selection computation, orthotactic well-formedness computations have been argued to rely on syllabic role information. For example, Badecker (1996) noted that double consonants can occur in word-final positions but not in word-initial positions (i.e. biss is a possible written word but ssib is not), while double vowels can appear in either word-initial (e.g. ooze) or word-final (e.g. tree) positions. Cassar and Treiman (1997) showed that the performance of children learning to read and write is sensitive to these patterns. Thus, the representation of syllabic role information at the graphemic level may be necessary for the operation of constraints that favor the production of orthographically well-formed strings (much as has been argued for the phonological system, Kahn, 1976).

Orthographic representations in reading and spelling

Questions regarding the nature of orthographic knowledge have been widely discussed and investigated in the literature on written language comprehension (reading) as well as written language production (spelling). Reading, as well as spelling, requires various long-term and working memory mechanisms that operate over letters and word spellings. These are responsible for the translation between letters/words and their corresponding sounds and, for

words, there is the additional mapping between word spellings and word meanings. One possibility is that spelling and reading rely on entirely distinct orthographic representations. Given that LSS's deficit resulted in a disruption of his spelling processes it is possible that the conclusions reached in this study specifically concern orthographic representations in spelling. However, there is fair amount of evidence from investigations of reading indicating that the orthographic representations used in both reading and spelling have similar internal structure. For example, evidence from a number of paradigms supports a role for syllabic structure in reading: letter identification tasks (e.g., Smith & Spoehr, 1973), word identification tasks (e.g., Katz & Baldasare, 1983), priming (e.g., Taft, 1979; Taft & Forster, 1976), transposition priming (e.g., Lee & Taft, 2009), the illusory conjunction paradigm (e.g., Prinzmetal, Treiman & Rho, 1986; Rapp, 1992), lexical decision and naming latencies (e.g., Carreiras, Alvarez & de Vega, 1993), among others. Furthermore, there are a number of findings (Olson & Nickerson, 2001 and Rapp, 1992) indicating that these results are not simply phonological and that they do not simply reflect low-level statistical regularities in orthographic strings. There is also evidence that digraphs are represented differently from other letter pairs in the orthographic reading representations. For example, Rey and colleagues (2000; see also Rey & Schiller, 2005) showed that letter identification for briefly presented words is more difficult when a letter appears in a digraph than when the letter appears alone (e.g. A in GREAT versus GRAPE). Pring (1981) reported that case-alternated words were recognized more slowly when the two elements of a digraph appeared in different cases, as in GreAT, than when they appeared in the same case, as in GReaT (see Martensen et al., 2003 for similar results, but cf. Lupker et al., 2012). These results, as well as recent findings suggesting that letter position is represented in the same way in both reading and spelling (Fischer-Baum, McCloskey & Rapp, 2010; Fischer-Baum, Charny & McCloskey, 2011), support the claim that reading and spelling rely on similarly structured orthographic representations.

A straightforward account of these similarities is that reading and spelling share orthographic representations and processes. While there are relatively few studies specifically examining the relationship between orthographic production and comprehension, existing behavioral, neuropsychological and neuroimaging evidence most clearly supports the notion that at least some representations and processes are shared in reading and spelling, with the clearest evidence specifically indicating shared lexical processes and representations (but also see Tainturier and Rapp, 2003 for a review of evidence regarding shared orthographic working memory processes). For example, behavioral studies with neurologically intact participants have found word-specific performance similarities/differences in spelling and reading or reading-spelling priming (e.g., Holmes & Carruthers, 1998; Burt & Tate, 2002; Monsell, 1987). In addition, the neuropsychological literature has provided a number of reports of well-studied individuals with lexical deficits in both reading and spelling (Rapcsak & Beeson, 2004; Philipose et al., 2009; Tsapkini & Rapp, 2010). Finally, with regard to neuroimaging, a number of recent studies (Rapp & Lipka, 2011; Rapp & Dufor, 2011, Purcell et al., 2011) identified two different neuroanatomical regions in the left hemisphere (the mid-fusiform gyrus and the inferior frontal junction), in which the locations of activation peaks for reading and spelling

(in the same individuals) were statistically indistinguishable, providing strong evidence of components shared by reading and spelling.

The relationship between orthography and phonology

The conception of orthographic structures as multidimensional feature bundles was initially proposed by analogy to representational structures in autosegmental phonological theory (Goldsmith, 1976) and proposals about orthographic syllable structure borrowed directly from the phonological notion of the syllable (Badecker, 1996). The idea that quantity and identity can be represented separately play an active role in phonological theory, where this idea had been proposed for geminated consonants and long vowels, as well as to account for certain morphophonological phenomena (Hayes, 1989). Even the notion that single units at one level of representation may correspond to multiple units at another has a parallel in phonological theory. Affricates, like the first sounds of CHURCH or JURY, are argued to be complex unitary segments in the phonology (see Clements, 1999 for review) just as has been proposed for the orthographic representation of digraphs.

Given these similarities between proposed orthographic and phonological representations, it is natural to wonder if the structures inferred from various written language investigations are appropriately described as orthographic or if they are not, instead, actually phonological. Such concerns are unsurprising given that written language develops later than spoken language and clearly builds upon it, both as it developed during human history and also as it is acquired in each individual. Furthermore, even in the adult expert reader and writer, spoken language knowledge and processes are often active during written language processing (see Rastle and Brysbaert, 2006 for review). However, even though there are parallels between the representational structures that organize orthographic and phonological representations, the content of the representations is certainly modality specific (e.g., graphemes vs. phonemes or phonetic features). Furthermore, even if doubling exists in both the orthography and the phonology, doubled consonants in English orthography do not correspond to geminated consonants in the phonology. Also, while orthographic and phonological representations make use of the same syllabic roles (Onset, Nucleus and Coda), the same word may be syllabified differently in the two modalities. For example, in English (and French) the final E, which is phonologically silent in many words, would be considered to be an orthographic syllable nucleus such that “France” would be monosyllabic phonologically, but bi-syllabic orthographically (FRAN/CE; Chetail & Content, 2012, 2013). Similarly, complex unitary segments may exist in both the orthography and the phonology, but a complex phoneme unit may be associated with a simple orthographic segment in the orthography (e.g. the affricate /dʒ/ maps onto the simple grapheme J), and vice-versa (e.g., the simple phoneme /ʃ/ maps onto the digraph SH).

In sum, orthographic representations are not simply parasitic on phonological representation. Instead, the similarities between orthography and phonology suggest that the same principles of representation – notions of multi-dimensional representations, a separation of identity and quantity information and complex units – apply in the different domains (see Badecker, 1996). Interestingly, it has been argued these same principles may apply more broadly across the cognitive sciences, applying even in non-linguistic domains. For

example, it has been claimed that representing item identity and quantity as separable dimensions may be a general property of the cognitive system (Endress, Nespor & Mehler, 2009). Further work is required to investigate the extent to which other representational structures and principles identified in the phonology and orthography appear in non-language cognitive domains.

Conclusions

To summarize, the findings from this investigation, along with other lines of evidence, challenge the view that orthographic representations consist of simple concatenations of letter identities. The results reported here support a rich conception of orthographic representation, similar to representational structures proposed in spoken language and potentially other non-language cognitive domains. This investigation provides an example of how the detailed analysis of perseveration errors in spelling can reveal the specific ways in which our long-term memories for the spellings of words are internally complex and structured, identifying both the separable dimensions of representation as well as shedding light on the nature of orthographic representational units. As a result, we come closer to understanding what it is that we know when we know the spellings of words.

Acknowledgments

We are grateful to LSS for his enthusiastic participation and to his wife for her invaluable support. We thank Mike McCloskey for comments and generous guidance with data analysis techniques, and Jenna Moray and Tony Pastor for help with data analysis. This research was supported by NIH Grant DC006740.

Appendix 1

The observed proportion of perseverations is a combination of some true perseveration errors from true sources, and some errors that are not true perseverations from true sources (here called pseudo-perseverations). The chance proportions at the preceding positions (E-1 and E-2) cannot be used to directly estimate the proportions of pseudo and true perseverations because this chance value refers to a data set in which all intrusions were unrelated to their sources – something which we have determined does not apply to LSS's data set. Nonetheless, this rate can still be used indirectly to estimate the pseudo-perseveration rates in an actual data set. We take the single consonant intrusions as an example to illustrate the procedure. The set of single consonant intrusions contains two intrusion types that are not true perseverations: (a) intrusions that do not even appear to be perseverations (they have no sources in preceding responses) – in LSS's single consonant intrusion set this proportion is .31 (1.0-.69) and (b) pseudo-perseverations, which are the subset of the intrusions that appear to have perseveration sources but for which the relationship between intrusion and source is due to chance – we refer to this proportion as X. Thus the total proportion of intrusions that are not true perseverations is $X + .31$. This intrusion set is comparable to the control data set in that it contains no true perseverations. Given that the chance analysis estimates that approximately .21 of a set of unrelated items should, by chance, appear to be perseverations we can now solve for X: $.21 = X/(X+.31)$. When we solve this we find that $X = .08$. This means that for LSS's set of single consonant intrusions, 8% of the intrusions that appear to have a source at E-1 are false perseverations.

Using this value to adjust the apparent perseveration rate we estimate the rate of perseverations with a true source at E-1 as .61.

REFERENCES

- Badecker W. Representational properties common to phonological and orthographic output systems. *Lingua*. 1996; 99:55–83.
- Blanken G, Schäfer C, Tucha O, Lange KW. Serial processing in graphemic encoding. Evidence from letter exchange errors in a multilingual patient. *Journal of Neurolinguistics*. 1999; 12:13–39.
- Buchwald A, Rapp B. Consonants and vowels in orthographic representation. *Cognitive Neuropsychology*. 2006; 23:308–337. [PubMed: 21049333]
- Buchwald A, Rapp B. Distinctions between orthographic long-term memory and working memory. *Cognitive Neuropsychology*. 2009; 26:724–751. [PubMed: 20425660]
- Burt JS, Tate H. Does a reading lexicon provide orthographic representations for spelling? *Journal of Memory and Language*. 2002; 46:518–543.
- Caramazza A, Miceli G. The structure of graphemic representations. *Cognition*. 1990; 37:243–297. [PubMed: 2282774]
- Caramazza A, Miceli G, Villa G. The role of the graphemic buffer in spelling: Evidence from a case of acquired dysgraphia. *Cognition*. 1987; 26:59–85. [PubMed: 3608396]
- Carreiras M, Alvarez CJ, de Vega M. Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*. 1993; 32:766–780.
- Cassar M, Treiman R. The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*. 1997; 89:631–44.
- Chetail F, Content A. The internal structure of chaos: Letter category determines visual word perceptual units. *Journal of Memory and Language*. 2012; 67:371–388.
- Chetail F, Content A. Segmentation of Written Words in French. *Language and Speech*. 2013; 56:125–144. [PubMed: 23654119]
- Clements, GN. Affricates as noncontoured stops.. In: Fujimura, O.; Joseph, BD.; Palek, B., editors. *Proceedings of LP '99: Item Order in Language and Speech*. The Karolinum Press; Prague: 1999. p. 271-299.
- Dunn, LM.; Dunn, LM. *Peabody Picture Vocabulary Test-Revised*. American Guidance Service; Circle Pines (MN): 1981.
- Endress AD, Nespors M, Mehler J. Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*. 2009; 13:348–35. [PubMed: 19647474]
- Folk J, Rapp B. Interaction of lexical and sublexical information in spelling: Evidence from nonword priming. *Applied Psycholinguistics*. 2004; 25:565–585.
- Fischer-Baum S, Charny J, McCloskey M. Both-edges representation of letter position in reading. *Psychonomic Bulletin & Review*. 2011; 18:1083–1089. [PubMed: 21935735]
- Fischer-Baum S, McCloskey M, Rapp B. The representation of grapheme position: Evidence from acquired dysgraphia. *Cognition*. 2010; 115:466–490. [PubMed: 20378104]
- Fischer-Baum S, Rapp B. Underlying cause(s) of letter perseveration errors. *Neuropsychologia*. 2012; 50:305–318. [PubMed: 22178232]
- Goldsmith, J. *Autosegmental phonology*. Doctoral dissertation. MIT; 1976.
- Hayes B. Compensatory lengthening in moraic phonology. *Linguistic Inquiry*. 1989; 20:253–306.
- Holmes VM, Carruthers J. The relation between reading and spelling in skilled adult readers. *Journal of Memory and Language*. 1998; 39:264–289.
- Houghton, G.; Glasspool, D.; Shallice, T. Spelling and serial recall: Insights from a competitive queuing model.. In: Brown, GDA.; Ellis, NC., editors. *Handbook of spelling: Theory, process and intervention*. John Wiley & Sons; Chichester: 1994.
- Houghton G, Zorzi M. Normal and impaired spelling in a connectionist dual-route architecture. *Cognitive Neuropsychology*. 2003; 20:115–163. [PubMed: 20957568]

- Jones AC, Folk JR, Rapp B. All letters are not equal : Subgraphemic texture in orthographic working memory. *Journal of Experimental Psychology : Learning, Memory and Cognition*. 2009; 35:1389–1402.
- Jónsdóttir MK, Shallice R, Wise R. Phonological mediation and the graphemic buffer disorder in spelling: cross-language differences? *Cognition*. 1996; 59:169–197. [PubMed: 8681510]
- Kahn, D. Syllable-based generalizations in English phonology. Ph.D.Dissertation. MIT; 1976.
- Kandel S, Alvarez C, Vallée N. Syllables as processing units in handwriting production. *Journal of Experimental Psychology: Human Perception and Performance*. 2006; 32:18–31. [PubMed: 16478323]
- Kandel S, Héroult L, Grosjacques G, Lambert E. Orthographic vs. phonologic syllables in handwriting production. *Cognition*. 2009; 110:440–444. [PubMed: 19150536]
- Kandel S, Soler O, Valdois S, Gros C. Graphemes as motor units in the acquisition of writing skills. *Reading & Writing: An Interdisciplinary Journal*. 2006; 19:313–337.
- Kandel S, Spinelli E. Processing complex graphemes in handwriting production. *Memory & Cognition*. 2010; 38:762–770. [PubMed: 20852239]
- Kandel S, Valdois S. Syllables as functional units in a copying task. *Language and Cognitive Processes*. 2006; 21:432–452.
- Kaplan, E.; Goodglass, H.; Weintraub, S. *The Boston Naming Test*. Lea & Febiger; Philadelphia: 1983.
- Katz L, Baldasare J. Syllable coding in printed word recognition by children and adults. *Journal of Educational Psychology*. 1983; 75:245–256.
- Kay, J.; Lesser, R.; Coltheart, M. *Psycholinguistic assessment of language processing in aphasia*. Lawrence Erlbaum; Hillsdale, NJ: 1992.
- Lee CH, Taft M. Are onset and codas important in processing letter position? A comparison of TL effects in English and Korean. *Journal of Memory and Language*. 2009; 60:530–532.
- Lupker SJ, Acha J, Davis CJ, Perea M. An investigation of the role of grapheme units in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*. 2012; 38:1491–1516. [PubMed: 22309089]
- Marcel, T. Phonological awareness and phonological representation: Investigation of a specific spelling problem.. In: Frith, U., editor. *Cognitive processes in spelling*. Academic Press; London: 1980.
- Martensen HE, Maris E, Dijkstra AFJ. Phonological ambiguity and context sensitivity: On sublexical clustering in visual word recognition. *Journal of Memory and Language*. 2003; 49:375–395.
- McCloskey M, Badecker W, Goodman-Schulman RA, Aliminosa D. The structure of graphemic representations in spelling: Evidence from a case of acquired dysgraphia. *Cognitive Neuropsychology*. 1994; 11:341–392.
- McCloskey M, Macaruso P, Rapp B. Grapheme-to-lexeme feedback in the spelling system: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*. 2006; 23:278–307. [PubMed: 21049332]
- Miceli G, Capasso R. Spelling and dysgraphia. *Cognitive Neuropsychology*. 2006; 23:110–134.
- Monsell, S. On the relation between lexical input and output pathways for speech.. In: Allport, DA.; MacKay, DG.; Prinz, W.; Scheerer, E., editors. *Language perception and production: Relationships among listening, speaking, reading and writing*. Academic Press; London: 1987.
- Olson AC, Caramazza A. Orthographic structure and deaf spelling errors: syllables, letter frequency, and speech. *Quarterly Journal of Experimental Psychology A*. 2004; 57:385–417.
- Olson AC, Nickerson J. Syllabic organization and deafness: Orthographic structure or letter frequency in reading? *The Quarterly Journal of Experimental Psychology A*. 2001; 54:421–438.
- Patterson K. Lexical but nonsemantic spelling? *Cognitive Neuropsychology*. 1986; 3:341–367.
- Philipose LE, Gottesman RF, Newhart M, Kleinman JT, Herskovits EH, Pawlak MA, et al. Neural regions essential for reading and spelling of words and pseudowords. *Annals of Neurology*. 2007; 62:481–492. [PubMed: 17702036]
- Pring L. Phonological codes and functional spelling units: Reality and implications. *Perceptual Psychophysics*. 1981; 30:573–8.

- Prinzmetal W, Treiman R, Rho SH. How to see a reading unit. *Journal of Memory and Language*. 1986; 25:461–475.
- Purcell J, Turkeltaub P, Eden G, Rapp B. Examining the central and peripheral processes of written word production through meta-analysis. *Frontiers in Language Sciences*. 2011; 2 article 239.
- Rapcsak SZ, Beeson PM. The role of left posterior inferior temporal cortex in spelling. *Neurology*. 2004; 62:2221–2229. [PubMed: 15210886]
- Rapp B. The nature of sublexical information: The bigram trough hypothesis examined. *Journal of Memory and Language*. 1992; 31:33–53.
- Rapp B, Dufor O. The neurotopography of written word production: An fMRI investigation of the distribution of sensitivity to length and frequency. *Journal of Cognitive Neuroscience*. 2011; 23:4067–4081. [PubMed: 21812571]
- Rapp B, Epstein C, Tainturier MJ. The integration of information across lexical and sublexical processes in spelling. *Cognitive Neuropsychology*. 2002; 19:1–29. [PubMed: 20957529]
- Rapp, B.; Fischer-Baum, S. The structure of orthographic representations.. In: Ferreria, V.; Goldrick, M.; Miozzo, M., editors. *Oxford Handbook of Language Production*. Oxford University Press; New York: in press
- Rapp B, Lipka K. The literate brain: The relationship between spelling and reading. *Journal of Cognitive Neuroscience*. 2011; 23:5, 1180–1197.
- Rastle K, Brysbaert M. Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*. 2006; 53:97–145. [PubMed: 16554045]
- Read, C. *Children's creative spelling*. Routledge and Kegan Paul; London: 1986.
- Rey A, Schiller NO. Graphemic complexity and multiple print-to-sound associations in visual word recognition. *Memory & Cognition*. 2005; 33:76–85. [PubMed: 15915794]
- Rey A, Ziegler JC, Jacobs AM. Graphemes are perceptual reading units. *Cognition*. 2000; 75:B1–B12. [PubMed: 10815779]
- Rumelhart DE, Norman DA. Simulating a skilled typist: a study of skilled cognitive-motor performance. *Cognitive Science*. 1982; 6:1–36.
- Spoehr K, Smith E. The role of syllables in perceptual processing. *Cognitive Psychology*. 1973; 5:71–89.
- Sternberg S. Separate modifiability, mental modules, and the use of pure and composite measures to reveal tem. *Acta Psychologica*. 2001; 106:147–246. [PubMed: 11256336]
- Taft M. Lexical access via an orthographic code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behavior*. 1979; 18:21–39.
- Taft M, Forster KI. Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behaviour*. 1976; 15:607–620.
- Tainturier MJ, Caramazza A. The status of double letters in graphemic representations. *Journal of Memory and Language*. 1996; 35:53–73.
- Tainturier, MJ.; Rapp, B. The spelling process.. In: Rapp, B., editor. *The handbook of cognitive neuropsychology: What deficits reveal about the human mind*. Psychology Press; Philadelphia: 2001.
- Tainturier MJ, Rapp B. Is a single graphemic buffer used in reading and spelling? *Aphasiology*. 2003; 17:537–562.
- Tainturier MJ, Rapp B. Complex graphemes as functional spelling units: Evidence from acquired dysgraphia. *Neurocase*. 2004; 10:122–131. [PubMed: 15788251]
- Treiman R. Children's spelling errors on syllable-initial consonant clusters. *Journal of Educational Psychology*. 1991; 83:346–360.
- Treiman R, Zukowski A, Richmond-Welty ED. What happened to the “n” of sink? Children's spellings of final consonant clusters. *Cognition*. 1995; 55:1–38. [PubMed: 7758269]
- Tsapkini K, Rapp B. The orthography-specific functions of the left fusiform gyrus: evidence from modality and category specificity. *Cortex*. 2010; 46:185–205. [PubMed: 19428003]
- Ward J, Romani C. Consonant-vowel Encoding and Ortho-syllables in a case of Acquired Dysgraphia. *Cognitive Neuropsychology*. 2000; 17:641–663. [PubMed: 20945199]
- Wechsler, D. *Wechsler Memory Scale*. 3rd ed.. The Psychological Corporation; London: 1997.

Zesiger, P.; Orliaguet, JP.; Boë, LJ.; Mounoud, P. The influence of syllabic structure in handwriting and typing production.. In: Faure, C.; Lorette, G.; Vinter, A., editors. *Advances in Handwriting and Drawing: A multidisciplinary approach*. Europa; Paris: 1994.

Position	X	X	X	X	X
Identity	[B]	[A]	[L]	[E]	[T]
Quantity	[Sgl]	[Sgl]	[Dbl]	[Sgl]	[Sgl]

(a)

Position	X	X	X	X	X
Identity	[B]	[A]	[S]	[E]	[T]
Quantity	[Sgl]	[Sgl]	[Dbl]	[Sgl]	[Sgl]

(b)

Position	X	X	X	X	X
Identity	[B]	[A]	[L]	[E]	[T]
Quantity	[Sgl]	[Sgl]	[Sgl]	[Dbl]	[Sgl]

Figure 1.

Depiction of the multi-dimensional representations (including grapheme identity, position and quantity) for: (a) the target word BALLET (b) the double-letter substitution error BASSET and (c) the geminate transposition error BALEET.

Position X X X X X X
 Identity [B] [A] [L] [E] [T]

(a)

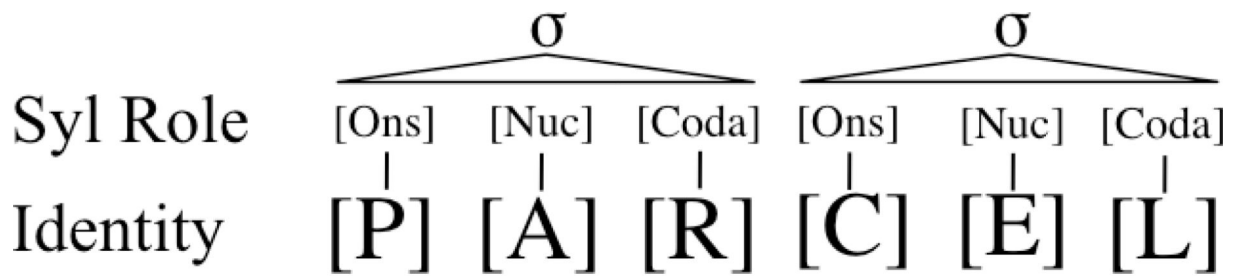
Position X X X X X X
 Identity [B] [A] [S] [E] [T]

(b)

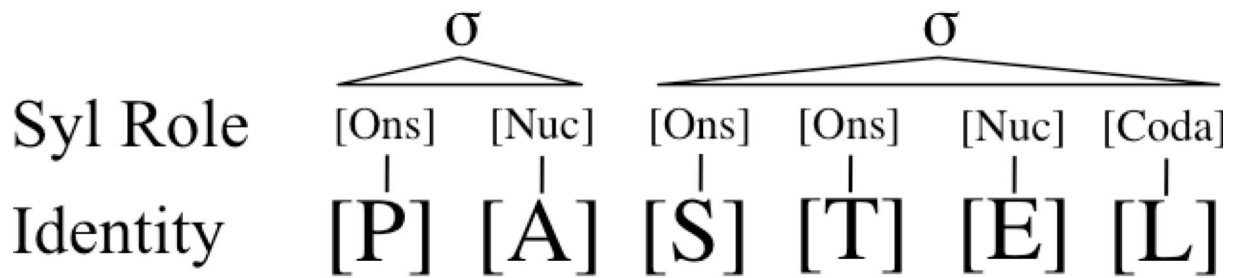
Position X X X X X X
 Identity [B] [A] [L] [E] [T]

Figure 2.

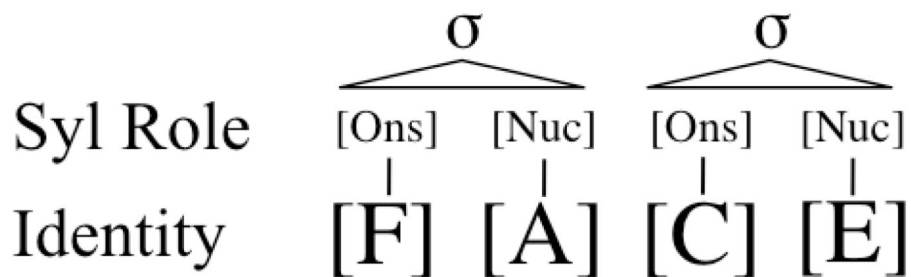
Based on the McCloskey et al., 1994, proposal including only grapheme identity and position dimensions, a depiction of representations for: (a) the target word BALLET (b) the double substitution error BASSET and (c) the geminate transposition error BALEET.



(a)



(b)

**Figure 3.**

Depiction of the grapheme identity and syllabic role representations for: (a) the word the word PARCEL (b) PASTEL and (c) the word FACE.

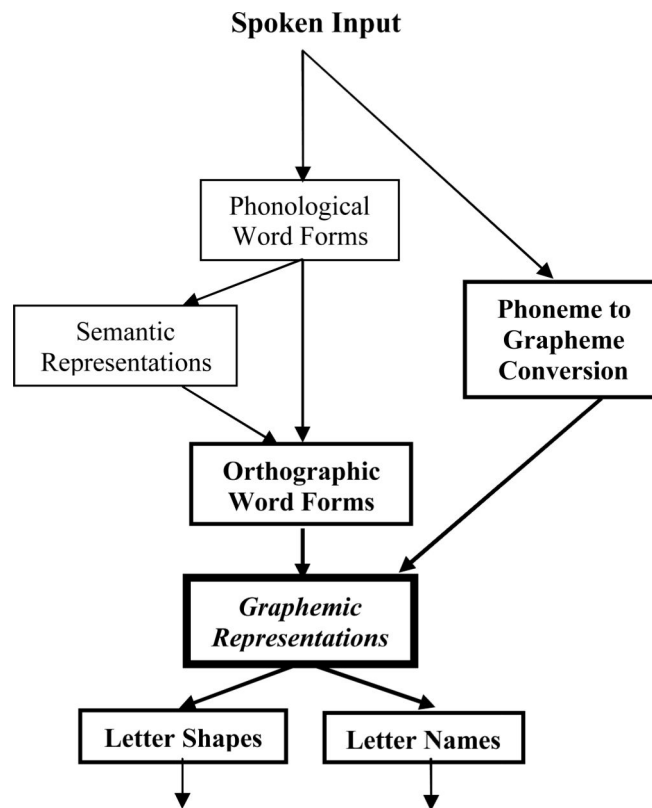


Figure 4. The cognitive architecture of spelling, highlighting (in bold) inputs and outputs to and from the level of graphemic representation.

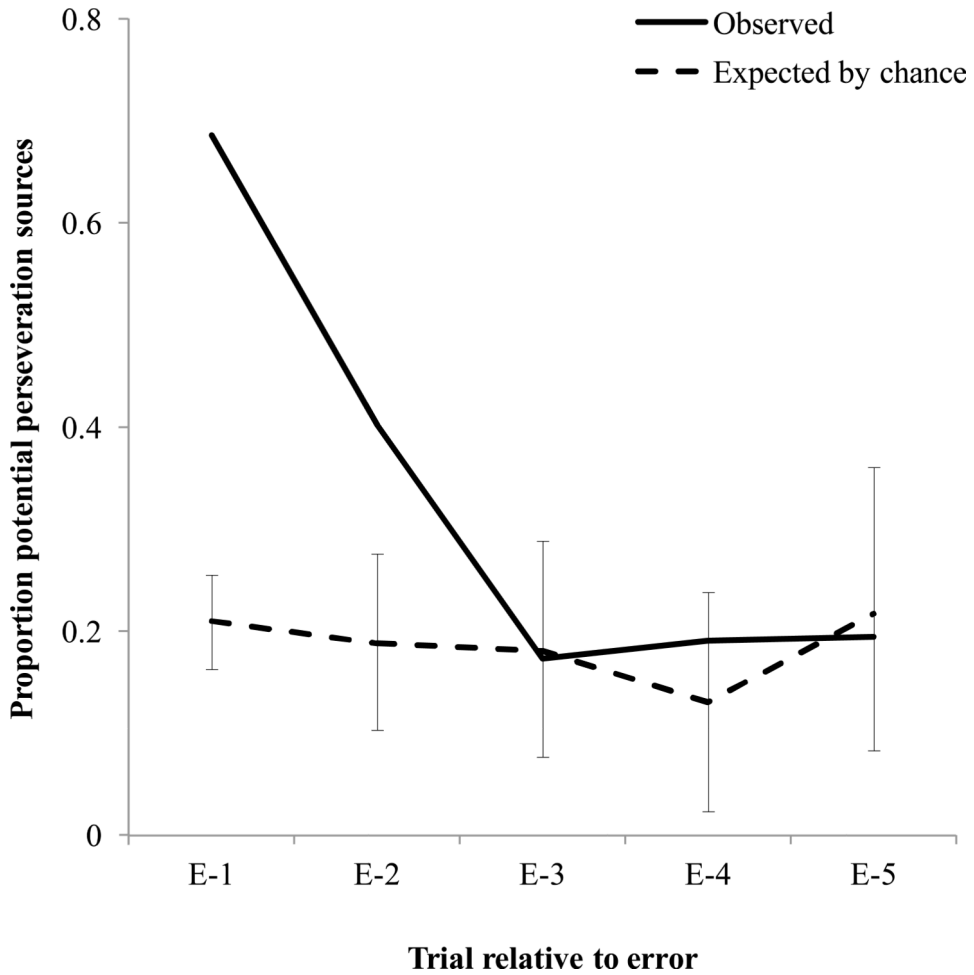


Figure 5.

For all single consonant intrusions produced at trial E, the solid line depicts the proportion of the trials in the preceding responses (E-1 through E-5) in which the intruded consonant first appeared. The dotted line corresponds to the proportion of trials that would be expected by chance to contain the intruded consonants. The bars indicate the 95% confidence intervals of the chance values. (See text for details.)

Syl Role	[Ons]	[Nuc]	[Nuc]	[Ons]	[Nuc]
Identity	[C]	[A]	[U]	[S]	[E]

(a)

Syl Role	[Ons]	[Nuc]	[Nuc]	[Coda]	<i>[Coda]</i>
Identity	[C]	[O]	[U]	[S]	<i>[T]</i>

Figure 6. Depiction of the syllabic structure of the target word CAUSE and the coda-intrusion error in the response COUST.

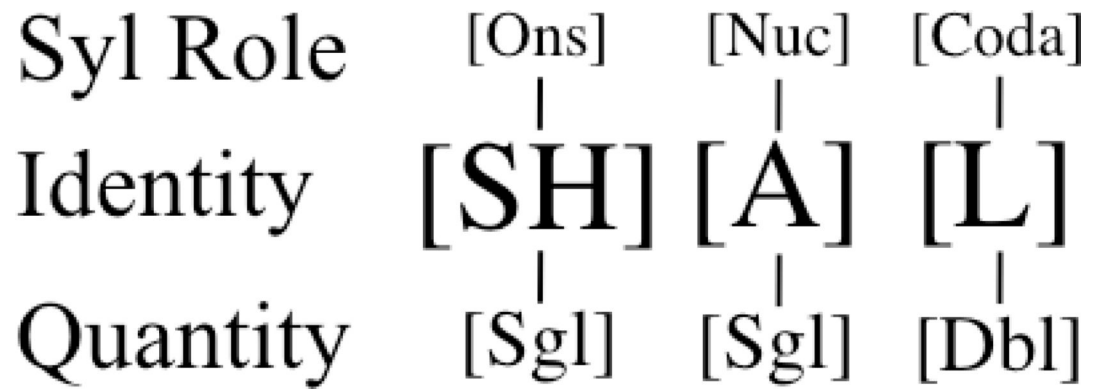


Figure 7.

Depiction of the multi-dimensional representation of the word SHALL, including the graphemeidentity, quantity, and syllabic role dimensions.

Table 1

Example of an intrusion error (trial E) produced by LSS along with the five preceding responses

Trial	Target	Response
E	UNDER	UNDEL
E-1	MOTEL	MOL <u>DEL</u>
E-2	SHOULD	SHOU <u>L</u> D
E-3	VULGAR	CHEU <u>P</u> IVE
E-4	CHEAP	CHEAP
E-5	CERTAIN	ABSEVE

Table 2

Observed proportions of quantity-not-identity and identity-not-quantity perseverations errors produced by LSS compared with the proportions predicted to occur by chance for each preservation type.

	Observed Proportion	Chance Proportion	p-value
Quantity-not- Identity	.54	.29	<i><.0001</i>
Identity-not-Quantity	.15	.04	<i><.0001</i>

Table 3

Observed proportions of role-not identity and identity-not-role perseveration errors produced by LSS compared with the proportions predicted to occur by chance for each perseveration type

	Observed Proportion	Chance Proportion	p-value
Role-not-Identity	.81	.65	<.05
Identity-not-Role	.26	.14	<.01

Table 4

Comparisons between the true perseveration proportions for single consonant intrusions, digraph intrusions and adjacent consonant intrusions, as well as the expected true perseveration proportion if the two adjacent intrusions were the result of independent perseverations from previous response/s.

	True Perseveration Proportion	vs. Single	vs. Expected Adjacent
Single Intrusions	.69		
Expected Adjacent Intrusions	.44		
Digraphs	.85*	ns	<.05
Adjacent Consonants	.46	<.05	ns

* Digraph vs. Adjacent Consonants, $p < .05$

Table 5

Proportion of single consonant intrusions whose source is either a previous consonant digraph or a consonant cluster compared to the proportions expected by chance.

	Observed Proportion	Chance Proportion	p-value
From Digraphs	.03	.03	ns
From Clusters	.19	.10	<i><.0001</i>