Verbal Working Memory as Emergent from Language Comprehension and Production

Steven Schwering[1], Maryellen C. MacDonald[1*],

[1]Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA

*** Correspondence:**
Maryellen MacDonald
mcmacdonald@wisc.edu

Keywords: working memory, language comprehension, language production, serial order, long-term memory, lexical representations.

Abstract

This article reviews current models of verbal working memory and considers the role of language comprehension, and long-term memory in the ability to maintain and order verbal information for short periods of time. While all models of verbal working memory posit some interaction with long-term memory, few have considered the character of these long-term representations or how they might affect performance on verbal working memory tasks. Similarly, few models have considered how comprehension processes and production processes might affect performance in verbal working memory tasks. Modern theories of comprehension emphasize that people learn a vast web of correlated information about language and the world and must activate that information from long-term memory to cope with the demands of language input. To date, there has been little consideration in theories of verbal working memory for how this rich input from comprehension would affect the nature of temporary memory. There has also been relatively little attention to the degree to which language production processes naturally manage serial order of verbal information. The authors argue for an emergent model of verbal working memory supported by a rich, distributed long-term memory for language. On this view, comprehension processes provide encoding in verbal working memory tasks, and production processes maintenance, serial ordering, and recall. Moreover, the computational capacity to maintain and order information varies with language experience. Implications for theories of working memory, comprehension, and production are considered.

Introduction

In 1885, Hermann Ebbinghaus published his extensive verbal memory experiments and observations, establishing a new theoretical approach to cognitive psychology through the formal study of memory (Ebbinghaus, 1885/1913). In his quest to isolate the properties of memory, Ebbinghaus observed that immediate recall of verbal material was utterly contaminated by long-term knowledge of language. He found it impossible to isolate immediate memory when he probed recall of meaningful verbal memoranda such as lines of poetry or narratives, and he established critical methodological practices aimed at stripping away confounding factors. In his attempt to isolate immediate memory, Ebbinghaus developed a collection of nonwords, thousands of consonant-vowel-consonant syllables that could be used to construct lists for immediate recall. The contamination of long-term experience persisted, as certain nonwords exhibited "very important and almost incomprehensible variations as to the ease or difficulty with which they are learned." (p. 23). Moreover, Ebbinghaus noted that even these novel materials could not completely isolate immediate memory from other cognitive processes; visual, acoustic, and articulatory components of verbal perception and action necessarily affected task performance.

Over 130 years of research now contributes to answering the questions posed by Ebbinghaus, and it is useful to ask how his catalyzing observations continue to influence theoretical and methodological approaches to memory research. In this article, we critically analyze Ebbinghaus' goal of isolating immediate memory as well as his warning that such isolation may be impossible. Following some establishment of terms and definitions and a brief sketch of some current models of immediate memory, we consider several intersecting points, all of which stem from a language-based perspective on the ability to temporarily maintain verbal information. First, we consider the dependence of immediate memory on long-term language knowledge, as Ebbinghaus first observed, and consider the impact of these relationships on modern theories of working memory. These modern accounts recognize some role for long-term memory, but we argue that they have been slow to embrace more modern approaches to the nature of long-term word representations and processing. Instead, we argue that language comprehension and production processes underpin encoding, maintenance, and production of old and new verbal memoranda without the need for separable buffers that are common in some current memory models. A key development in some models of immediate memory is the assumption that memory for words is separate from memory for their orders. In contrast, we consider the many ways in which various word and order representations are intertwined in language comprehension and production research and propose a new emergent account which incorporates these representations in VWM. In closing, we consider implications of our perspective on theories of language use and on related research areas.

Working memory models and terminology

There exists a fundamental disagreement about the definition of working memory (e.g. Aben, Stapert, & Blokland, 2012; Cowan, 2008), as evidenced by a wide array of both qualitative descriptions of immediate memory and competing memory models (see Cowan, 2017). We will focus on two general classes of models for how humans can encode verbal material, maintain it for a brief period of time, and produce the memoranda by speaking or writing. Proponents of the two types of models that we discuss, the multicomponent models (e.g. Baddeley, Thomson, & Buchanan, 1975; Baddeley, 2000) and emergent models (e.g., Cowan, 1993; Postle, 2006), do not always use terms in the same way, and so we begin with some definitions.

Verbal working memory (VWM) is commonly viewed as the temporary maintenance of verbal information (i.e., some aspects of language). Some researchers distinguish VWM as an immediate memory for processing of information (converting speech to meaning, say) from short -term memory (STM), a passive temporary store. However, as Buchsbaum & D'Esposito (2019) have noted, information is always being transformed in some way in the service of goal-directed behavior, and so we will use the term VWM to refer to both storage and processing, except where we specifically refer to theories invoking a STM component. Finally, VWM researchers have increasingly investigated the ability to recall verbal material in the same order it was presented. Thus, we discuss abilities to recall a word or nonword (termed *item memory*) and recall in the correct order in a list (*order memory*).

Multicomponent models, which get their name from the distinct components posited in the working memory system (Baddeley, 1992), draw a sharp distinction between passive storage of information in "buffers" and processing mechanisms such as speech perception and production processes. In this respect, multicomponent models are aligned with classical theories of working memory advanced by Ebbinghaus. In this view, the sole function of STM is to act as a site of storage. Specifically, multicomponent models posit a short-term buffer that maintains a rapidly degrading representation of memoranda (Baddeley, Lewis, & Vallar, 1984). Critically, in this perspective, long-term memory is separate from STM (e.g. Shallice & Warrington, 1970; 1974), but via a process called *redintegration* (e.g., Hulme et al., 1997), LTM can provide cues to rebuild STM as it degrades (Lewandowsky & Farrell, 2000). LTM can interact with STM in other ways. With respect to language processing, some researchers claim that verbal STM is a buffer that stores partially processed linguistic representations (e.g. R. C. Martin & He, 2004; R. C. Martin & Romani, 1994), or is a specific subcomponent of language processing mechanisms dedicated to storage (Shallice & Papagno, 2019). Certain theories propose that the buffer holds copies of or pointers to representations derived from LTM that may require further processing in the future (Norris, 2017). Thus, whereas Ebbinghaus (1885) tried to isolate STM processes within an interacting system, the multicomponent models have converted that research goal into an architectural claim: STM is a distinct system with only the most limited, indirect contact with LTM and language processing mechanisms.

Although multicomponent accounts are the dominant perspective in VWM research, there is a long history of cautions about this approach. More than 25 years ago, Crowder (1993) predicted a wholesale reassessment of multicomponent models of VWM in favor of alternative approaches. He described the notion of a separate, dedicated short-term store (the multicomponent model) as "archaic and, to some of us, even downright quaint" and suggested that "Increasingly, the field is turning instead to a procedural attitude toward memory" (p. 143). Crowder's predictions were wildly inaccurate in their timeline, as multicomponent models of memory remain important and useful theories of VWM now many decades after Crowder predicted their demise. Nevertheless, Crowder correctly predicted the rise of alternative, emergent models of VWM that did away with separate buffers.

Emergent approaches do not generally distinguish between storage and processing mechanisms. Some earlier variants were called procedural models, defining VWM as a secondary product of procedures in support of other cognitive processes (Craik & Lockhart, 1972; Crowder, 1993; D. M. Jones, Macken, & Nicholls, 2004; Kolers & Roediger, 1984). Other early theorizing by Saffran, N. Martin, and colleagues explored relationships between aphasic patients' VWM in the context of their language production abilities, informed by Dell's (1986) spreading activation model of language production (N. Martin, Saffran, & Dell, 1996; Saffran & Martin, 1997). We advocate this "rich emergent" approach here, where VWM is the activated portion of linguistic LTM (Acheson & MacDonald, 2009a, 2009b; Buchsbaum & D'Esposito, 2018; Cowan, 1993; Hasson, Chen, & Honey, 2015; MacDonald, 2016; Postle, 2006). This approach emphasizes VWM as a complex of skills, honed by past language comprehension and production experience. On this view, knowledge of word meanings and other forms of linguistic knowledge shape performance in VWM tasks. Performance on VWM tasks co-opts language LTM, by which we mean any parts of LTM involved in language tasks, including knowledge of events, word meanings, word order, phonological form, and other information (Acheson & MacDonald, 2009b; MacDonald, 2016; MacDonald & Christiansen, 2002). LTM itself is characterized as a set of processing mechanisms employed to achieve goal-directed behavior rather than store a static set of memoranda chunked or compressed from prior experience (Buchsbaum & D'Esposito, 2019; Postle, 2006). In the case of WM for linguistic memoranda, we have proposed that the language production architecture is co-opted to maintain and order the memoranda, obviating the need for a separate memory buffer (Acheson & MacDonald, 2009b; MacDonald, 2016). Whereas, in the multicomponent model, effects of prior language knowledge in LTM have been attributed to secondary mechanisms (e.g. Hulme et al., 1997; Lewandowsky & Farrell, 2000), we see these LTM effects arising naturally from language production and comprehension processes. For example, language production is well known to favor serial orders that have been used frequently or recently (Bock, 1986a) and to group related words together in an utterance (Solomon & Pearlmutter, 2004). These biases in production may underlie effects of semantic grouping and similarity to natural language that have been observed in recall tasks (T. Jones & Farrell, 2018; Miller & Selfridge, 1950). Thus, we view temporary maintenance and ordering as the job of action systems, which must construct an action plan and maintain it before it can be executed, so that the action plan is the "memory of what is to come" (Rosenbaum, Cohen, Jax, Weiss, & van der Wel, 2007, p. 528). For language, the action planning system is language production, and the utterance plan is the memory of both what is to be produced and the order in which it will be produced at several levels, including words, phonemes, and articulatory gestures (Acheson & MacDonald, 2009b; MacDonald, 2016; N. Martin et al., 1996). On this view, VWM is simply the skill of maintaining and ordering linguistic material, and that skill, as with all subcomponents of language production and comprehension, emerges from actions of the language systems and varies with experience (MacDonald & Christiansen, 2002; MacDonald, 2016).

In contrast to the "rich emergent" account described above, some "limited emergent" accounts posit a more restricted interaction with language processes, with different systems working in parallel to support memory for items and their orders (Majerus, 2013; 2019). On this view, item memory engages ventral language pathways that process semantics, with dorsal pathways supporting order within the item (i.e. phonemes). In contrast, order memory for sequences of words themselves engages frontal-parietal networks and networks closely associated with attentional mechanisms. The item/order memory distinction has been supported by findings that word characteristics, like frequency of use (Poirier & Saint-Aubin, 1996; Saint-Aubin & Poirier, 1999) and semantics (Majerus & D'Argembeau, 2011), largely affect memory for items but not memory for order. Furthermore, memory for items and order appear to engage distinct neural populations, as indicated by neuroimaging results (Guidali, Pisoni, Bolognini, & Papagno, 2019; Majerus et al., 2006; Majerus et al., 2008) and aphasic patient data (e.g. Majerus, Attout, Artielle, Van der Kaa, 2015; Majerus, Norris, & Patterson, 2007).

The separate item/order memory of more limited emergent accounts is consistent with a multicomponent approach, namely that LTM is able to support STM only in cases where the items and order conform to prior experience. Multicomponent models are particularly emphatic about this point, arguing that this is a critical reason a STM buffer must exist distinct from LTM (e.g. Norris, 2017). Some emergent accounts also recognize that there are limitations to LTM. For example, Majerus (2013) suggests that "the representations of the language system are able to support familiar item and order information, but not unfamiliar order information" (p. 4). This distinction between familiar and unfamiliar orders is problematic because it presumes a dichotomy between novel and familiar, when similarity to prior experience is actually continuous. We consider this point further in the section entitled "Problems with Limited Emergence."

In the next sections, we contrast our rich emergent account against a variety of alternative multicomponent and more limited emergent memory models. Specifically, we describe current research on the nature of LTM language representations and the language comprehension and production processes that interact with LTM. Because all accounts of VWM must refer in some way to LTM, we argue that this characterization of language knowledge informs all theories of encoding, maintaining, and ordering verbal information.

Word Representations in VWM and Language Research: No Word is an Island

Since the time of Ebbinghaus, most VWM models have assumed discrete representations or "items" in memory. Often, verbal memory is conceptualized by the unit of the word, or word-like collections of phonemes (nonwords). For example, there are a multitude of studies investigating immediate or delayed word recall that document word accuracy across list position (e.g. Murdock, 1962; Watkins & Watkins, 1977), word omissions (e.g. Roodenrys, Hulme, Lethbridge, Hinton, & Nimmo, 2002), word intrusions (e.g. Coltheart, 1993), and so on. Furthermore, measurement of VWM capacity is often indexed by list span, or the average number of words recalled from lists (e.g., Daneman & Carpenter, 1980; Hulme & Tordoff, 1989). In part, such descriptions are a convenient shorthand for bits of information (Miller, 1956) , but they also reflect certain assumptions about the isolability of memory representations. One common assumption is that word memory is supported by fully separable phonological and semantic codes (Howard & Nickels, 2005; R. C. Martin, 1987; R. C. Martin, Lesch, & Bartha, 1999). Another is that order memory is separable from the memory for the word, itself; this view is further compounded by viewing the words in lists as separate from each other, especially in the case of novel word orders (Majerus, 2013, 2019).

Considering that all major memory models posit some kinds of ties with language representations, it bears asking how a compartmentalized view of item and order representations, and a compartmentalized view of item components (e.g. phonology, semantics, grammatical role), accords with language research. In this section, we describe developments in both comprehension and production research that is completely antithetical to the isolated representations prevalent in much memory research. This work shows that different levels of language representation used in production and comprehension, what we refer to as language LTM, influence each other and are integrated. We suggest that this integration, and the statistical regularities between classically defined and supposedly dissociable representations that are critical for language research, have significant consequences for how verbal information is maintained. In other words, we argue that the nature of linguistic LTM representations, as revealed in research on language comprehension and production, is highly releveant to theories of VWM.

Integrated representations in language processing. Researchers' views about the nature of word representations and their use in comprehension and production have undergone enormous change in the last several decades. Initially, researchers believed that comprehension processes were modular, such that dedicated components worked independently to interpret language input (e.g., Frazier, 1987; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; see also Almeida & Gleitman, 2018, for more historical context and current views of modularity). Similarly, models of production were highly staged, with minimal interaction between different language representations (e.g. Levelt, Roelofs, & Meyer, 1999). Theories of word representation pointed to a lexicon with distinct levels (phonological, syntactic, semantic, e.g., Allport & Funnell, 1981). Importantly, these models assumed that, regardless of the nature of LTM, language processes could selectively extract and operate over subcomponents of linguistic knowledge, such as processing phonology or syntax without meaning, with some later integration stage (Forster, 1985; Frazier, 1987). While this work did not often invoke VWM, the notions of separable language components and isolated processing systems are compatible to the orientation of multicomponent models.

More recent theories of language comprehension are far less aligned with these compartmentalized approaches. Instead they have emphasized extensive interaction between different kinds of language representations. This is most clearly demonstrated behaviorally in instances where certain information cannot be "turned off", even when it is beneficial to do so (e.g., Stroop, 1935). For example, Seidenberg and Tanenhaus (1979) demonstrated that the orthographic form of a word interfered with judgments of phonological form, meaning that one form of information in LTM (orthographic information) interfered with another form of information in LTM (phonological form). While early neuropsychological studies suggested that the subcomponents of language knowledge were represented with discrete neural codes (Dapretto & Bookheimer, 1999), more recent analyses support integrated representations. For example, Siegelman, Blank, Mineroff, and Fedorenko (2019) argue against previous evidence for divisions between syntactic and semantic representations during sentence comprehension. Similalry, Dikker, Rabagliati, Farmer, & Pylkkänen (2010) found that phonological/orthographic information contributes to syntactic analyses within 100 ms, even before a word has been recognized, because the phonological form is correlated with and therefore provides information about the likely grammatical category (noun, verb, etc.) of the to-be-recognized word. Together this work and others (e.g., Pereira et al., 2018) suggest that word comprehension and LTM representations are much more interconnected than was previously recognized.

This article is not the place for a full specification of how representations are integrated, nor for the natural ongoing debates concerning how to characterize linguistic knowledge, but it is worth noting why a number of researchers now assume extensive interaction and integration among what has been traditionally described as distinct levels of linguistic information. In these more integrated accounts, multiple sources of information interact in perception and comprehension because those interactions are beneficial, essential really, to comprehend and produce language in real time. Language contains strong correlations between different levels of language knowledge, between language and the world, and between information earlier and later in a linguistic signal to be interpreted. People are voracious statistical learners, and they leverage their LTM of the statistical regularities between different kinds of information to comprehend and produce language efficiently and accurately (Seidenberg & MacDonald, 2018). Indeed, the combination of several partially informative information sources (phonology and semantics, for example) is now seen as central to accounting for the speed with which comprehenders interpret incoming language input despite the massive ambiguity known to pervade language; an individual source of information only weakly constrains interpretation alone but is highly effective in combination with other constraints (Graves, Binder, Desai, Conant, & Seidenberg, 2010; Joanisse & McClelland, 2015; MacDonald & Seidenberg, 2006; Seidenberg, 1997). Each language comprehension experience is a source of learning (Chang, Dell, Bock, & Griffin, 2000), and a consequence of learning all this combinatorial information is that any one source of information, including words, cannot be atomic or isolated (Willits, Amato, & MacDonald, 2015). Instead, words and other classically defined levels of representation are highly intertwined, because learning (and therefore LTM) must capture a complex web of statistical structure to maximize performance during language comprehension and production. Word representations can be modeled as attractors in networks comprising various types of information (phonological, semantic, etc., Hinton & Shallice, 1991), and some linguists and psycholinguists now consider discrete notions such as *word* and *phoneme* to be convenient fictions, highly useful for researchers' discussions but having more to do with people's conscious intuitions than with the way that language is actually represented and processed in the brain (Baayen, Shaoul, Willits, & Ramscar, 2016; Bybee & McClelland, 2005; Ramscar & Port, 2016).

Separated representations in memory models. These highly interactive approaches have not yet penetrated much of the theorizing in most current multicomponent and emergent models of VWM, which continue to emphasize individual "items" of memory. Multicomponent models posit specialized, separate buffers, such as the phonological loop (Baddeley & Hitch, 1974), which encode a single type of information. Initially, patient lesion data seemed to provide further support to modular memory and language approaches, as in patients who exhibited impaired memory abilities with spared language abilities (often called "STM patients", e.g., Warrington & Shallice, 1969) and in cases reporting double dissociations of phonological and semantic information in memory and language tasks, leading to a separation between phonology and semantics in multicomponent models (R. C. Martin, Shelton, & Yaffee, 1994; R. C. Martin & Romani, 1994). This dissociation between representations extends into memory for items and their order. Certain aphasic patients demonstrate apparently isolable item or order memory impairments (Attout, Kaa, George, & Majerus, 2012; Majerus, Attout, Artielle, & Van der Kaa, 2015), and this behavioral pattern is accompanied by neuroimaging evidence suggesting item and order memory are supported by distinct neural populations (Attout, Magro, Szmalec, & Majerus, 2019; Kalm & Norris, 2014).

A strict notion of "item" in memory becomes more complicated when considering the qualities of statistical information in linguistic LTM. For example, phonotactic long-term knowledge influences recall of novel words. Non-words consistent with the transitional probabilities of phonemes (or acoustic properties or articulatory gestures) in natural language are recalled better than non-words inconsistent with these patterns (Gathercole, Frankish, Pickering, & Peaker, 1999; Thorn & Frankish, 2005). Researchers have likewise extended these findings to suggest that both lexical and sublexical properties affect recall of non-words (Majerus, Van der Linden, Mulder, Meulemans, & Peters, 2004; Roodenrys et al., 2002). Tanida, Nakayama, and Saito (2019) further demonstrated an effect of forward and backward bimora transition probabilities on ordered recall. Together, these results suggest that memory of one phoneme or acoustic pattern influences memory of others via LTM of the phonological statistical structure of language. These "neighborhoods" of patterns in LTM can be quite subtle, as evidenced by improved recall for nonwords with regular pitch accent compared to irregular pitch accent, an effect moderated by phonotactic frequency (Tanida, Ueno, Lambon Ralph, & Saito, 2015; see also Yuzawa & Saito, 2006). Not only do these studies suggest that LTM is relevant for VWM, but they suggest multiple grain sizes of phonological information interact to inform performance in memory tasks.

Beyond phonological information, language users also track and leverage complex statistical regularities between different types of linguistic representations, such as between phonology and semantics. Our claim is not that phonology and semantics are completely merged (they are clearly not), but rather that they are intertwined to a degree that affects language use and VWM performance. Such regularities are not always obvious. Indeed, with some exceptions (Christiansen & Monaghan, 2016; Farmer, Christiansen, & Monaghan, 2006; Schmidtke, Conrad, & Jacobs, 2014), the mapping between phonology and semantics seems largely arbitrary. If phonology and semantics were completely distinct, then each representation could be stored in a separable buffer, consistent with multicomponent accounts. However, claims for a strict semantic-phonological divide break down when considering morphologically complex words, such as *painter, ideas, friendship,* and *working*. These words contain morphemes (*-er,- s, -ship, -ing*) for which the mapping from phonology to semantics is not arbitrary. The same mapping occurs repeatedly through the language (e.g. *worker, baker, seeker,* etc.), and words sharing these affixes form semantic-phonological neighborhoods that shape language LTM and behavior (Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997; Seidenberg & Gonnerman, 2000). These relationships also encode grammatical form (e.g., *–er* is associated with nouns, *-ing* with verbs). It might be tempting to consider morphologically complex words as marginal and not part of more "typical" language, but morphologically complex words are common in English and their phonological-semantic-grammatical regularities have been shown to affect word learning in infants (Willits, Seidenberg, & Saffran, 2014). In adults, regularities between phonological, orthographic, semantic, and grammatical knowledge drive very early stages of language comprehension, even before conscious word recognition (Dikker et al., 2010). Even so, recent reviews suggest there is a "notorious lack of consensus" (p. 37) in the imaging literature about the brain representations of phonological, semantic, and morphological relationships among more complex words (Leminen, Smolka, Duñabeitia, & Pliatsikas, 2019). As such, it is clear that many representations simultaneously impact language comprehension and production, and it is unclear how any one representation could be extricated from this web of processing.

Given these regularities in language use, it is not surprising that morphophonological regularities also impact VWM. For example, the use of morphophonological cues has been well-studied in children's nonword repetition. Nonwords with morphophonological cues are recalled better than nonwords without such cues, and children with language impairments may be less sensitive to this effect (Archibald & Gathercole, 2006; Casalini et al., 2007; Estes, Evans, & Else-Quest, 2007). Thus, experience with language, specifically the regular co-occurrences between phonology and semantics in morphologically complex words, affects VWM for nonwords (though see Szewczyk, Marecka, Chiat, & Wodniecka, 2018). These results have largely been examined with children completing single word repetition tasks. It would be worthwhile to extend this work to other tasks and populations. Incorporating regularities between phonology and semantics in stimuli (e.g. via use of affixes), could alter the apparent separability of phonology and semantics, as has been suggested by many memory and language studies (e.g. R. Martin, Shelton, & Yaffee, 1994).

The "primary systems" approach to memory and language use begins to incorporate some current insights about language representations and argues for phonology and semantics as separable yet interacting representations (Savill, Cornelissen, Pahor, & Jefferies, 2019; Ueno et al., 2014). Broadly, this approach supports emergent memory accounts, suggesting that the effects of semantics and phonology on word and non-word recall reflect a balance of processing. For example, when phonological support is weak, semantic support affects recall to a larger degree compared to when phonological support is strong (Savill, Cornelissen, Whiteley, Woollams, & Jefferies, 2019). In such accounts, the interactions between phonology and semantics emerge from processing in a quasi-regular domain, resulting in integrated representations. Ueno et al. (2014) demonstrated that words with low imageability are recalled worse than words with high imageability (i.e. effect of semantics), and this effect is exacerbated by words with an atypical pitch accent (i.e. effect of phonotactics). In line with the primary systems account, this suggests that the effect of phonotactics on recall depends in part on semantics. Interestingly, the researchers developed a neurobiologically constrained connectionist model of word comprehension, repetition, and production, demonstrating that phonological (ventral) and semantic (dorsal) language pathways are differentially engaged when processing typical and atypical phonotactic patterns. As a result, the semantic pathway was more engaged in processing atypical phonotactic patterns. Such research suggests that subtle phonological information may infiltrate a putative semantic pathway (see also Jefferies, R. W. Jones, Bateman, & Lambon Ralph, 2005).

The tracking of complex statistical patterns in support of language comprehension, production, and memory is not limited to within-word components like phonology and semantics; statistical regularities also support representation of word order. This point gets to the heart of the item vs. order distinction in VWM theorizing. Memory researchers readily agree that sentences are recalled better than scrambled lists of words (Brener, 1940), and this effect scales with list approximation to natural language sequence statistics (Miller & Selfridge, 1950). These effects are typically attributed to semantic coherence or episodic pattern recognition (Allen, Hitch, & Baddeley, 2018; Baddeley, Hitch, & Allen, 2009). However, episodic memory is not sufficient to explain the full range of results. Memory is similarly facilitated for lists of non-words that approximate natural language syntax (Epstein, 1961; 1962). Thus, meaning does not seem to be necessary to the effect. Jones & Farrell (2018) further demonstrated that people are more likely to recall sentence-like lists in an order consistent with syntactic knowledge, and that errors are more likely to conform to prior syntactic knowledge than expected by chance (for corpus analyses tying language experience to memory performance, see G. Jones et al., 2020; Perham et al., 2009). In each case, inter-item information affected memory for order by invoking long-term knowledge of language syntax, suggesting that memory for items and their order interact to support each other. For example, experience using English builds a LTM of the word *pull*. The LTM of *pull* not only encodes meaning and sound but also co-occurrence tendencies; *pull* is often is flanked by words denoting animate entities and objects involved in a pulling event (as in *The girl pulled the cart*). We are emphatically not claiming that linguistic knowledge is limited to co-occurrence, merely that such knowledge includes linear relationships and that what might be viewed as multi-word frequency knowledge shapes both language use (Seidenberg & MacDonald, 2018) and memory (Arnon & Snider, 2010). While strict chaining accounts of ordering have generally fallen out of favor in memory research (e.g. Hurlstone, Hitch, & Baddeley, 2014), these studies suggest that inter-item associations are not only encoded and leveraged for performance in memory tasks (see also Fischer-Baum & McCloskey, 2015, for discussion) but reinforced by LTM. Such effects are likely amplified by the presence of multi-morphemic words (such as *pulled*), because, as noted above, morphemes such as –*ed* also contain grammatical information and thus provide cues to inter-word relationships (see Epstein, 1961; 1962). Thus, it is unclear to what extent item knowledge can be separated from order knowledge if the source of the order benefit is derived from the information associated with the individual words.

The role of language processes in performing VWM tasks. If performing a VWM task is dependent on language processes, such as comprehension for encoding (MacDonald & Christiansen, 2002), lexical production for item memory (Page et al., 2007), or sentence production skills for item ordering (Acheson & MacDonald, 2009b; MacDonald, 2016), then theories of VWM must consider how theories of language comprehension and production constrain memory performance. Here, we describe some current models of language comprehension and production with a specific eye toward describing statistical regularities in language and the integrated representations in LTM that capture those regularities. Of course, these models were not explicitly designed to model performance in VWM tasks. There is an essential tension between the complexity of LTM representations and modeling: the more complex and intertwined the representations are thought to be, the more difficult it is to capture this complexity in a computational model. Few explicit emergent models of VWM exist, as some researchers have noted (Norris, 2017), though many models adopt principles consistent with the emergent approach (e.g. Botvinick & Plaut, 2006). However, from the language emergent perspective, theories of language comprehension and production should serve as a useful analogue, continuing the role models of language use have played in shaping memory research (e.g. N. Martin, Dell, Saffran, & Schwartz, 1994).

On this view, language comprehension and production processes underlie the encoding and retrieval mechanisms posited in memory accounts, respectively. Language comprehension processes extract meaning from input by mapping an input signal to a semantic representation of the entities and events being referred to (MacDonald & Hsiao, 2018). Often, comprehension processes involve partial predictions of upcoming input (Altmann & Mirković, 2009; Federmeier, 2007; Kuperberg & Jaeger, 2016), which means that comprehension processes routinely involve not only semantic integration of words that have been encountered but also generation of serial order expectations among representations of words that are likely upcoming in the input. Similarly, interpretation of some language input can depend on material that comes later (Connine & Clifton, 1987; MacDonald, 1994). There are many language comprehension models that depend on integrated representations, variously capturing word segmentation (Christiansen, Allen, & Seidenberg, 1998), utterance interpretation without a separate word segmentation stage (Baayen et al., 2016), the learning of phonological forms (Plaut & Kello, 1999), word reading and its relationship to phonology (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989), the learning of grammatical knowledge (J. Allen & Seidenberg, 1999), behavior in the visual world paradigm (Mayberry, Crocker, & Knoeferle, 2009), disorders of comprehension in individuals with developmental language disorder (also called specific language impairment, Joanisse & Seidenberg, 2003), and more. In turn, language production models attempt to generate a well-formed utterance from a message representation, either externally motivated in the case of a repetition task or internally generated in the case of self-generated production. Several interactive models exist, capturing lexical selection (i.e., retrieving words from LTM, Dell, Schwartz, Martin, Saffran, & Gagnon, 1997) and phrase (Dell, Burger, & Svec, 1997) or sentence production (Chang, Dell, & Bock, 2006; Dell & Chang, 2014). The Lichtheim-2 model implements an account of single word comprehension and repetition as well as the degradation of those processes in aphasia (Ueno, Saito, Rogers, & Lambon Ralph, 2011). All of these models share several core features that tie them to the emergent account. In each, learning algorithms, such as backpropagation, encode statistical knowledge in the connection weights updated through experience, forming the model's LTM. Each of these models also develops a VWM through learning; for example the TRACE model of speech perception (McClelland & Elman, 1986) got its name from the claim that the STM trace of the model emerged from the interacting layers of the network. No separable STM buffers divorced from LTM are employed in any of the above models.

Critically, integrated representations are a core part of these language models, most commonly instantiated as distributed representations in a network. Distributed representations, as their name implies, spread a representation over the entire network, via connection weights between layers. Integrated representations exhibit at least two key ties to distributed representations in connectionist language models. First, integrated representations emerge in processing via bidirectional spreading activation between layers, a feature evident in models of human comprehension and production (e.g., Dell, 1986; Seidenberg & McClelland, 1989). Second, the integrated representations blend processed information across the network such that phonological, semantic, lexical, and grammatical information cannot be strictly separated from other types of information (e.g., McClelland et al., 2010). Of course, we are not claiming that language models do not develop certain specializations for phonological, semantic, lexical, grammatical, and other types of information. Instead, specialization is a matter of degree, where complete modularity and complete overlap are less likely than an intermediate state (McClelland et al., 2010). For example, in some models, impairments of a discrete representation (e.g. phonology) disrupts the use of other representations (e.g. semantics), via layers that allow interaction between those representations (e.g. Monaghan & Woollams, 2017). Such models are most consistent with primary systems accounts (e.g. Ueno et al., 2014; Savill, Cornellisen, Pahor, & Jefferies, 2019). In other models, the integrated representations are not as explicit. For example, simple recurrent networks of comprehension and production, allow information to be processed through time. Such networks cross item and order information via recurrent connections (Botvinick & Plaut, 2006; Elman, 1991; Joanisse & Seidenberg, 2003), and there is no clear way in which item and order information can be separated.

Distributed representations as they are captured in connectionist models are not the only way to characterize integrated representations. We have focused on variations in distributed connectionist approaches as examples that most clearly embrace the interconnected representations that should affect theorizing about VWM, but other computational approaches could also incorporate integrated representations in processing (e.g., Frank & Goodman, 2014). Furthermore, localist representations, like the one implemented in Dell, Schwartz, Martin, Saffran, and Gagnon (1997), also have interaction among different types of information and have proven incredibly useful in describing mechanisms by which LTM engages with VWM.

Potential research directions and predictions for a language-emergent VWM. There are several predictions for VWM research that stem from the language emergent view, the first of which emphasizes the role of language production processes in serial ordering of the items in a memory list. Previous research has argued that production processes are engaged in maintenance and recall of verbal material, specifically that the utterance plan that maintains the to-be-uttered words in order also serves the maintenance and ordering functions during VWM tasks (Acheson & MacDonald, 2009b; MacDonald, 2016). As MacDonald (2016) discussed, this claim is much more controversial for some kinds of VWM tasks and performance than others. For example, Page, et al. (2007) posited a limited role for language production processes in ordering at the item level. They argued that parallels between word production processes and word recall in VWM tasks pointed to individual, word-level utterance plans playing a role in phonological maintenance in VWM, but ordering the words themselves (order memory) must be the purview of a dedicated short-term store. Lombardi and Potter (1992) and Potter and Lombardi (1998) hypothesized a different role for language processing: in VWM tasks involving whole sentence repetition, the comprehension system interprets the meaning of the sentence and the production system regenerates it from that meaning. The model we advocate incorporates the language system for remembering individual words, whole sentences, as well as all cases in-between, including ordering of word sequences that are less than full, coherent sentences. As there are very few tests of these ideas in the existing literature, our discussion addresses the kinds of word-ordering phenomena in language production that may be relevant to performance in VWM tasks.

An essential task in language production is the creation of serial order over many levels, including messages, words, sub-lexical forms such as phonemes, and articulatory gestures that enable overt language (Dell, Burger, et al., 1997). Acheson and MacDonald (2009b) extensively reviewed how the interactivity of phonological information with other information predicted serial order phenomena through the lens of language production research. They concluded that "…one key insight about the serial ordering of verbal information in language production is that serial ordering results from interactions across multiple levels of representation over time, that is to say, as a result of recurrent connectivity" (p. 54). For example, word ordering in language production is more likely to go awry when words share features, including both grammatical features (e.g., noun) and phonological features (Dell & Reich, 1981), meaning that phonological and lexico-grammatical information are together affecting serial ordering processes. Given Acheson & MacDonald's review, we do not focus on phonological interactions with word order here, but it is worth noting a few more recent phenomena relevant to their claims. A number of studies have investigated semantic-phonological interactions termed semantic binding, the finding that lexico-semantic knowledge affects the nature of phonological representations in VWM and other tasks (e.g., Hoffman, Jefferies, Ehsan, R. W. Jones, & Lambon Ralph, 2009; Patterson, Graham, & Hodges, 1994; Savill et al., 2017). Relatedly, Acheson and colleagues conducted several studies suggesting that phonological and semantic information jointly affect serial order in VWM tasks in a way that would be expected from how information interacts in comprehension and production (Acheson, MacDonald, & Postle, 2011; Acheson, Postle, & MacDonald, 2010; see also Poirier, Saint-Aubin, Mair, Tehan, & Tolan, 2014). Similarly, Macken, Taylor, and D. M. Jones (2014) investigated the memory implications for prosody, the intonation patterns that span whole phrases and sentences in everyday language use, in VWM tasks. Like syntactic and discourse relations, prosody is another multi-word phenomenon that does not fit neatly into the item/order distinctions in memory tasks. Macken et al. found that prosodic phrasing does affect recall, which argues against individual word units in memory.

Far less research concerns the nature of sentence-level language planning and serial ordering in VWM tasks. We mention three findings from language production research that seem particularly relevant to claims about the role of language production in VWM, because all three point to the essential non-independence of words and word orders in utterance planning. First, a central tenet across essentially all approaches to language production is that lexico-semantic characteristics of individual words strongly affect their order in a sentence (Bock, 1987; Levelt, 1993). An example is that animate entities like *woman* tend to appear earlier in utterances than inanimate words like *book*. This effect is thought to reflect a more general phenomenon linked to LTM retrieval, in which early-retrieved words enter the utterance plan first and end up in earlier positions in the utterance (Bock, 1987). Semantic features such as animacy affect retrieval and, consequently, serial position in the sentence (Bock, 1987; MacDonald, 2013). Second, the word orders that people produce tend to be ones that have been recently produced (Bock, 1986b; Weiner & Labov, 1983), but the strength of this effect is modulated by the particular words in the sentence: repeated words lead to more repeated word orders (see Pickering & Ferreira, 2008, for review). Again, words and their orders are interdependent. Third, word orders and the presence/absence of optional words in sentences vary with semantic relationships between words, where semantic similarity between two words yields more word omissions and different word orders than in the absence of semantic similarity across words (Gennari, Mirković, & MacDonald, 2012; Hsiao, Gao, & MacDonald, 2014; Montag, Matsuki, Kim, & MacDonald, 2017). Thus, whereas the first two examples illustrated interactions between properties of a particular word and word order of an entire utterance, this example shows that semantic relationships between two words also affect word order. All of these examples of word and word order interdependence are broadly compatible with models of language production that represent production as activation of learned weights in a connectionist architecture; these representations arguably cross item and order memory (Chang et al., 2006; Dell & Chang, 2014; McCauley & Christiansen, 2014). On this view, language production models could serve as highly informative models of serial recall, especially when the models engage in sentence repetition (see Ueno, Saito, & Lambon Ralph, 2011 for word repetition and Fischer-Baum, 2018 for other potential commonalities in serial order representations). We see this approach as inconsistent with the currently dominant views of VWM, that memory for items (the words) and memory for their serial order are unrelated, accomplished by independent mechanisms (Guidali, Pisoni, Bolognini, & Papagno, 2019; Henson, Hartley, Burgess, Hitch, & Flude, 2003; Majerus, 2009).

These results and approaches offer several avenues for investigations of the relationship between serial ordering of words in language production and VWM tasks. For example, it is worth further consideration of the item-order distinction in some theories of VWM, particularly those that posit a role for LTM and language production for item memory but a special purpose system for ordering the items (Majerus, 2009; Page et al., 2007). From the point of language production, serial order is crucial both across items (i.e. word order) and within items (syllable, phoneme, articulatory gesture orders). It is curious that within-word serial order demands are considered "item memory" rather than another example of ordering memory. For current purposes, a key difference between the two types of serial order would seem to be their regularity, in that phonological order is much more rigid than syntactic order. For example, the phonemes and articulatory gestures must be in a particular order to produce a given word, and the semantic identity of the word "binds" the sub-lexical representations and their order together —the semantic binding hypothesis (Patterson et al., 1994). Dell & Chang (2014) posit a similar kind of binding from message-level semantics to the serial orders of words, but this binding is weaker and more variable than in the word-phoneme case; there are statistical regularities between types of messages and sentence forms, but messages can usually also be conveyed with alternative word orders (MacDonald, 2016). In other words, the item-order distinction is really one of two different kinds of serial ordering demands and LTM, and the one called "item memory" (which includes ordering of phonological codes) is much stronger and more regular than the one called "order memory". On that view, it should be possible to manipulate these contingencies in simulations or experimentally, perhaps with artificial languages in which "word" order and "phoneme" order vary in their rigidity. If, after learning the artificial language, participants had to perform a memory task, we predict that performance at both levels should respond to the regularities of past experience and thus strength of LTM constraints, in contrast to accounts positing a rigid item/order distinction (see also Acheson & MacDonald, 2009 for discussion of "item" vs. phoneme errors and Botvinick & Bylsma, 2005 for recall in artificial languages).

Another interesting domain is performance in Hebb repetition tasks, in which participants repeatedly encounter certain serial orders across lists (Guerrette, Saint-Aubin, Richard, & Guérard, 2018; Page, Cumming, Norris, McNeil, & Hitch, 2013). Performance in these tasks should at least initially be moderated by statistical regularities in the broader language (that is, in LTM, via prior experience with language), where certain words occur in certain serial orders more frequently than others. For example, we might expect that words referring to animate entities (*child, teacher*) would yield different serial order behavior than inanimate words (*book, table*) in ordered recall, because people's broader experience ordering different types of words in their history of language production would affect how rapidly repeated patterns are learned. More generally, we expect serial ordering behavior to reflect both long-term language use and also rapid adaptation to more recent ordering contexts, a phenomenon that is robust in both language comprehension (Fine, Jaeger, Farmer, & Qian, 2013) and production (Bock, 1986a). Whereas Hebb repetition effects have been described in terms of repetition of specific tokens, syntactic priming effects in language processing carry across multi-word grammatical and semantic relations. If there are interactive representations between word and grammatical roles, then classic Hebb repetition effects should carry across these abstract relational categories and be moderated by fit with the category. Indeed, some studies have begun to examine these effects in sentence repetition (Allen, Hitch, & Baddeley, 2018; T. Jones & Farrell, 2018) and in recall of lists with grammatical dependencies (Perham, Marsh, & D. M. Jones, 2009) by considering how lists consistent with grammatical knowledge are recalled better than lists inconsistent with these patterns. The emergent account described here would further predict that the effect of grammatical knowledge would be moderated by semantic information of words, such as animacy, and morphophonological cues, reflecting interrelationships in LTM. For example, recall of animate nouns should be greater than recall of inanimate nouns in the context of word lists that encourage a noun to be interpreted as an agent, because animate nouns are commonly agents of actions and inanimate nouns are not. Furthermore, this account would suggest rapid adaptation to novel orders would affect memory in a manner consistent with models of language production that learn over experience.

Challenges for the multicomponent approach. Rather than viewing memory representations as graded, integrated, and distributed, as described above, multicomponent models separate various representations into discrete components. For example, the phonological loop stores phonological representations in a buffer separate from other representations (Baddeley & Hitch, 1974). Likewise, other researchers posit separate phonological and semantic buffers stemming from language mechanisms (R. C. Martin & Romani, 1994). These models are reminiscent of older, modular models of language comprehension and production that employ discrete stores and restricted interaction of information (Forster, 1985; Frazier, 1987). To capture the fully rich and interactive tapestry of language representations that are invoked in more current language research, multicomponent models would seem to require a combinatorial explosion of additional buffers for each form of interaction. In terms of parsimony and plausibility, this seems unlikely to be a tenable solution. R. Martin and colleagues offered a possible solution in which various language representations may interact in a multicomponent memory model by passing activity through layers with phonological and semantic buffers (e.g. R. C. Martin & Freedman, 2001). This approach may allow more interaction but is also inconsistent with much language research, as it specifically implies that certain language representations are processed independently and in sequence (MacDonald & Seidenberg, 2006). As far as we are aware, no research has explicitly considered how different forms of interactive representations could be modelled in VWM in a manner consistent with language comprehension and production research. Even so, it is unclear how integrated representations and interactive processing could be implemented in a multicomponent account.

An important route for LTM effects on VWM performance in multicomponent models is redintegration, a process that rebuilds decaying memory traces from LTM (R. J. Allen & Hulme, 2006; Clarkson, Roodenrys, Miller, & Hulme, 2017; Roodenrys & Hinton, 2002; Roodenrys et al., 2002). The redintegration mechanism not only rebuilds the phonological loop with phonological information from LTM (Clarkson et al., 2017), it also is the mechanism invoked to account for other LTM effects that go beyond phonological structure, including influences of word frequency and long-term knowledge of semantics and word co-occurrences on VWM (Hulme et al., 1997; Roodenrys et al., 2002; Stuart & Hulme, 2009; Walker & Hulme, 1999). On this view, the redintegration process must use LTM outside the phonological domain to shore up decaying phonological buffers. It is not clear how that process would work if LTM representations are highly integrated. Such a process would imply that phonological representations are first stripped from their richly integrated encoding in LTM, maintained in a separate phonological buffer, and then recombined with their integrated representations at the time of recall.

Currently, empirical evidence in favor of emergent (Acheson, Hamidi, Binder, & Postle, 2011; Buchsbaum & D'Esposito, 2018; Postle, 2006) and multicomponent accounts (for review see Shallice & Papagno, 2019; Yue, Martin, Hamilton, & Rose, 2019) has established little consensus. We recognize that many of the claims above are logical arguments, and further empirical evidence could prove some of our assumptions faulty. Proponents of emergent models should see language comprehension and production mechanisms as consistent with VWM systems that stem from a richly structured and integrated LTM (Acheson & MacDonald, 2009b; Hughes, Chamberland, Tremblay, & D. M. Jones, 2016; G. Jones & B. Macken, 2015). Proponents of multicomponent models, however, may see these discussions of a rich language LTM and the processes that operate with it as simply more evidence for the sorts of information that could be encoded via language processes or that redintegration could use to reconstruct memory traces. Regardless, defining LTM representations is important for the advancement of memory models, and language models should provide insight into these LTM representations.

Challenges for limited emergence. Perhaps one of the most persistent complaints against the emergent account is its inability to handle aphasic patient data (Shallice & Papagno, 2019). Classically, patterns of behavior by patients with aphasia have been seen as evidence for the notion that STM and LTM are supported by distinct neural populations. Lesions to the medial temporal lobe have appeared to yield deficits of LTM with spared STM, typically assessed using lexical decision tasks and digit span tasks, respectively (Baddeley & Warrington, 1970; Cave & Squire, 1992; Penfield & Milner, 1958; Scoville & Milner, 1957; Warrington, Logue, & Pratt, 1971). In contrast, damage to left parietal regions have been interpreted to cause impairments in verbal recognition tasks and digit span tasks with spans greater than 1 or 2 while sparing other cognitive functions and LTM (e.g. Shallice & Warrington, 1970; Shallice & Warrington, 1974; Vallar & Baddeley, 1984a; Warrington & Shallice, 1969; Warrington & Shallice, 1972). Thus, these studies of patients appeared to show a double dissociation of STM and LTM.

Some patient data may also support a dissociation between language processing and STM. For example, the patient K.F. reported in Warrington and Shallice (1969) exhibited strong repetition deficits with spared word knowledge, which would typically classify the patient as having conduction aphasia. However, given that the patient exhibited recognition deficits even when no verbal output was required by the task (i.e. pointing), Warrington and Shallice concluded that the patient's impairment was not limited to language repetition. Later work reinforced this notion in patients with impaired phonological discrimination with spared word recognition and short sentence comprehension (Basso, Spinnler, Vallar, & Zanobio, 1982; Silveri & Cappa, 2003; Vallar & Baddeley, 1984) as well as in patients with dissociable speech and STM deficits (R. C. Martin & Breedin, 1992). In a similar way, more recent research has attempted to unconfound item and order memory (Attout et al., 2012; Majerus et al., 2015).

However, the putative pure deficits of STM are frequently tainted by subtle language impairments (N. Martin & Saffran, 1992). For example, Warrington, Logue, & Pratt (1971) described a selective impairment of STM, yet those same patients exhibited difficulty in repetition of abstract words, reading, and fluent speech. Vallar and Baddeley (1984a; 1984b) claimed to have found a pure deficit of STM, yet that same patient exhibited impaired comprehension of longer sentences compared to other participants (1984a). Even the patients identified with fluent speech also exhibited abnormalities. For example, the patient described by Shallice and Butterworth (1977) exhibited paraphasic errors in speaking names and had difficulty comprehending spoken discourse and written text. Furthermore, comprehension difficulty was exacerbated for complex sentences. Jacquemot, Dupoux, Decouche, & Bachoud-Lévi (2006) claimed to have found patients with a specific STM impairment, yet those same patients also exhibited difficulty in language comprehension tasks and sentence repetition tasks, resulting in phonological paraphasias. A truly pure deficit has proven quite elusive (though see Martin & Breedin, 1992). Rather than see these language deficits as stemming from a specific STM impairment, we see both as being driven by deficits in LTM. A complementary pattern is seen in other lines of research. For example, Hannula et al. (2006) found that hippocampal deficits cause impairments in relational processing at both short and long durations, upsetting prominent research suggesting that hippocampal activity is associated only with LTM. A strongly emergent perspective accords neatly with this data.

A recurrent theme in this review has been that the relationship between VWM and LTM depends on the nature of language LTM. Patient data is no exception. Reference to models of language production and comprehension reveal how apparent STM deficits could be captured by damage to LTM. N. Martin and Saffran (1992) presented the case of a patient with deep dysphasia who exhibited apparent errors of STM: difficulty producing nonwords and semantic errors in repetition. This patient exhibited fluent speech with semantic and phonological paraphasias. The researchers evaluated this patient's performance through the lens of the Dell (1986) interactive spreading activation model of lexical retrieval. This model employs discrete representations of phonology, lexical entries, and semantics that interact in a bidirectional network. The model was able to produce human-like lexical selection behaviors without any storage buffer separate from LTM. Critically, the model was able to capture putatively pure STM patient data solely through perturbation of the model parameters and without the inclusion of a distinct memory buffer. In this specific case, an increased decay rate reduced the ability of lexical representations to support lexical selection. The predictions afforded by this model were later confirmed in additional analyses of patient data by N. Martin, Dell, Saffran, & Schwartz (1994; see also Dell, Schwartz, N. Martin, Saffran, & Gagnon, 1997), and patient recovery was also able to be modelled using the same framework (N. Martin et al., 1996). These results suggest that a specification of the LTM representations relevant to language comprehension and production may help test claims about the representational basis of VWM and its relationship to LTM.

Findings such as these point to the need for contact between theories of VWM and perspectives on long-term representations of serial order in language. That is, the extent to which the above or similar results affect VWM models depends on the hypothesized nature of LTM, particularly the extent to which LTM could contribute to representations of novel memoranda and their order. Language LTM captures relations between words and levels of linguistic representation and therefore allows generalization to new cases. Indeed, any linguistic input is novel in many ways, such as a new word order, new speaker, new acoustic environment, and so on. By definition, the goal of language comprehension processes is to cope with novel input, and language production processes constantly generate novel utterances. The VWM literature offers a different perspective, with some claiming that buffers are needed explicitly to represent novel material (Norris, 2017). One challenge for memory research is the need to characterize a clear divide between "old" and "new," especially given that novelty means very different things in different memory models. Distributed language models provide a key demonstration of the emergent perspective. In such models, novel stimuli are processed with respect to their similarity to prior experience, without any need for separate systems dedicated to handling the particularities of novel items or orders. In parallel, emergent models of VWM are capable of producing novel sequences just using LTM, without dedicated short-term buffers (e.g. Botvinick & Bylsma, 2005; Botvinick & Plaut, 2006, 2009). Perhaps greater adoption of graded representations of novelty could bridge the divide between language emergent and pure memory accounts. Important behavioral data linking graded phonotactic LTM to VWM (e.g. Tanida, Ueno, Lambon Ralph, & Saito, 2015) and graded grammatical LTM to VWM (e.g. T. Jones & Farrell, 2018) already speaks to the usefulness of this approach.

Implications for Language and VWM research

We have cited a broad range of work in both VWM and in language comprehension and production, and one of the striking features of that work is how very little the fields have to say about each other. For example, it is completely uncontroversial that language comprehension and production processes are constrained by what is commonly called "verbal working memory capacity" in those fields, and yet the specific mechanisms posited in classic VWM models are, with only a few exceptions, absent from theorizing about how limited capacities shape language processes (see Caplan & Waters, 2013, for review and a different perspective). Similarly, while VWM accounts assume that VWM abilities must be used in everyday activities, the connection to actual theories of language use is equally scant. Here we discuss several fronts with more potential for interaction among the fields.

Implications for relating WM assessments to other measures

The approach that we have advocated, in which performance on VWM tasks is heavily supported by language processes, which are themselves dependent on long-term knowledge, naturally leads to questions about what VWM tasks actually measure. This question is not only central to theories of working memory but also has enormous practical significance, because there is wide usage of tasks that are described as VWM assessments in clinical and educational contexts—in typical and atypical child development, young adults, older adults, and patients with brain injury. Whereas some researchers have considered poor VWM performance as a cause of poor language skill, potentially ameliorated by working memory training (e.g., Ingvalson, Dhar, Wong, & Liu, 2015), our language-emergent VWM view suggests that poor VWM performance is a symptom associated with poor language skill. In other words, the abilities to encode, maintain, and order verbal information are skills that emerge from language use, and individuals who have higher language skills have richer LTM representations and more practiced comprehension and production processes (see also G. Jones et al., 2020). Thus, we can view tasks that are described as VWM tasks not as assessments of a separate VWM capacity but rather as measures of a person's skill in encoding and maintaining verbal information. Consistent with this approach, there are now a number of reassessments of tasks that have previously been called "working memory tasks," with arguments that they are better viewed as assessments of language skill, including but not limited to encoding, maintenance, and ordering. Tasks that have been reinterpreted in this way include reading span (MacDonald & Christiansen, 2002), digit span (G. Jones & B. Macken, 2015), nonword repetition (Edwards, Beckman, & Munson, 2004; Graf Estes, Evans, & Else-Quest, 2007), sentence repetition (Klem et al., 2015), and immediate serial recall of word lists (Perham et al., 2009). In each of these examples, the argument has the same character. The apparent "verbal working memory task" does not measure a separate memory capacity but instead measures the quantity and quality of language skill and experience relevant to the specific demands of the task (see also D. Jones & Macken, 2018). Thus, nonword repetition performance can be traced to knowledge of phonological patterns and vocabulary (Edwards et al., 2004; Gupta & Tisdale, 2009), digit span performance can be linked to prior experience with and statistical learning of digit sequences (G. Jones & Macken, 2015), and so on. The over-arching conclusion from this work is that the skill or computational capacity to perform some task is not independent of long-term language knowledge and experience (MacDonald & Christiansen, 2002). That is an essential claim of an emergent perspective.

A related claim is that if the working memory abilities are emergent from language and other systems, then training on working memory tasks themselves should have little inherent benefit (e.g. Soveri et al., 2017), unless they have extra components that increase vocabulary or other relevant language skills. VWM training has been applied to therapeutic contexts, such as with aphasic patients, but the effectiveness of such interventions is unclear, driven in part by methodological limitations of single case studies (Zakarias et al., 2019). VWM treatments almost always employ linguistic stimuli of some sort, meaning they inherently provide some language practice. Therefore, VWM is rarely divorced from linguistic LTM in the training. VWM training research could benefit from a consideration of the emergent perspective defined here by further developing language skill as opposed to a separate memory capacity.

Implications for Attention, Task Subcomponents, and Domain Generality

All theories of VWM have some mix of domain-specific and domain-general components. For example, the multicomponent model has the domain-specific phonological loop but also the general Central Executive, which guides behavior beyond maintenance of phonological forms. Similarly, emergent views have domain-general attention and other cognitive control processes, but the nature of the LTM activated can be domain specific, in that linguistic knowledge need not have the same properties as memory for smell or spatial relations. The specific emergent approach advocated here, in which language LTM and language comprehension and production processes underlie VWM functions, might initially seem more strongly domain-specific in character, given the strong modular perspective that has pervaded language research. However, "emergent from language processes" need not be "domain specific." Indeed, there has been new interest in investigating how language use is supported by domain general processes of attention and episodic memory (Hepner & Nozari, 2019; Nozari, Trueswell, & Thompson-Schill, 2016; Van de Cavey & Hartsuiker, 2016), and interest in how distinct brain networks must coordinate to accomplish language comprehension and other complex cognitive processes (Fedorenko, 2014; Fedorenko, Behr, & Kanwisher, 2011). Close ties with attention have long been a component of emergent models (e.g. Cowan, 1993), and researchers are now considering the interrelationships between language and attention mechanisms with respect to VWM (Majerus, 2019). More generally, there is real interest in considering the extent to which language production processes are related to or are themselves emergent from more general action planning processes or domain-general sequencing systems (Anderson & Dell, 2018; Guidali et al., 2019; Van de Cavey & Hartsuiker, 2016). Long-term ordering knowledge across domains (e.g. Kaiser, 2012; Van de Cavey & Hartsuiker, 2016) may inform sequence ordering, further tying together domain general perspectives, emergent models, and language research. If language research continues to embrace more domain general processes, this development could have substantial consequences for debates about the relationship between language processes and VWM, including distinctions between multicomponent and emergent accounts. That is, if VWM and language researchers both incorporate the same domain general processes, then the distinction between multicomponent models and emergent models becomes less theoretically important.

Perhaps one of the most compelling examples of how domain-general processes affect language use and temporary maintenance may be seen in conversational turn taking, which draws on episodic memory (Duff & Brown-Schmidt, 2012; Rubin, Watson, Duff, & Cohen, 2014) and cognitive control. Using data from recordings of conversations in 10 languages, Stivers et al. (2009) found that speakers typically begin speaking less than 500 ms after the previous speaker has ended their conversational turn. A number of researchers have argued that this closely time-locked behavior requires extensive attention, maintenance, and cognitive control, because the next speaker simultaneously juggles a number of disparate tasks, some of which bear close similarity to demands of VWM tasks. The conversational demands on the person who will soon speak include: comprehending the person currently speaking; planning a response and maintaining that utterance plan until time to speak; predicting the timing of the current speaker's end point, which often involves predicting the actual words that the current speaker is likely to end on; and triggering an anticipatory in-breath and then exhalation to allow the speech to begin (De Ruiter, Mitterer, & Enfield, 2006; Levinson, 2016; Torreira, Bögels, & Levinson, 2015). Not surprisingly, turn taking and planning before speaking have high processing loads, as measured in a variety of methods (Barthel & Sauppe, 2019; Boiteau, Malone, Peters, & Almor, 2014; Kemper, Hoffman, Schmalzried, Herman, & Kieweg, 2011). Thus, while a participant's overall goals in a conversation and a VWM task are very different, it should be clear that the task demands of both activities overlap, including simultaneously encoding input while developing and maintaining plans to generate a response. Researchers are actively investigating the attention and cognitive control demands of language planning in advance of speaking, including serial ordering and monitoring of utterance plans (see Nozari & Novick, 2017, for review and Fischer-Baum, 2018 for potential implications for VWM tasks). Some methods manipulating selective attention to individual words in a list could prove to be useful for new studies of both VWM tasks and more typical language production (e.g., Nozari & Dell, 2012; Nozari & Thompson-Schill, 2013). We see this research as complicating the domain specific/general debates but also as an important arena for collaboration between VWM and language researchers.

Implications for Language Production Research

The view that language production is the engine of maintenance of verbal information has significant implications for language production research, because every VWM study can be seen as a particular subtype of language production, and therefore this work has the potential to inform theories of language production. Interaction between the fields has long been evident at phonological levels. There has been keen interest in phonological level speech errors as important data for theories of serial ordering in language production (Dell, 1984; Dell, Burger, et al., 1997), and there are extensive discussions of relationships between speech errors and recall errors in VWM tasks (Acheson & MacDonald, 2009a; Ellis, 1980; Hartley & Houghton, 1996; Page et al., 2007). In addition, VWM research has increasingly investigated the Hebb Repetition effect, the improved recall of repeated lists (Hebb, 1961; Oberauer, T. Jones, & Lewandowsky, 2015). In parallel, production researchers have investigated the effects of learning on serial ordering and speech errors in production (Anderson, Holmes, Dell, & Middleton, 2019; Dell, Reed, Adams, & Meyer, 2000). These investigations may be mutually informative, especially when placed in the context of computational models of ordering in VWM and models of language production which produce ordered sequences. As we have noted, some of these models have already suggested some parallels in ordering mechanisms between the two domains (Hartley, Hurlstone, & Hitch, 2016; Page & Norris, 2009).

There are also potential parallels beyond the phonological level, relevant to questions concerning the relationship between words and their production in ordered sequences. MacDonald (2016) argued that of the three most obvious task demand differences between immediate serial recall and everyday language production (item list vs. coherent message, recall signal vs. spontaneous production, and producing exact list order vs. flexible language production), the latter was particularly important for understanding relationships between language production and VWM. Whereas serial recall, by definition, must be in the presented order, a hallmark of language production at the phrase or sentence level is serial order flexibility —that almost any message can be conveyed via several different words and word orders. This difference is informative when considering how interference among similar words can affect performance in language production and VWM tasks. Interference among list items leads to item omissions and re-ordering of list items in recall; these are naturally treated as ordering errors, given the task demands in immediate serial recall (Baddeley, 1966; Page, et al., 2007; though see Saint-Aubin & Poirier, 1999). Language production is also subject to interference among words, which leads to omissions and alternative word orders, compared to production conditions without interference (Gennari et al., 2012; Hsiao et al., 2014). These shifts and omissions are not considered errors but in some sense evidence of production skill, that is, evidence for how the speaker uses alternative ordering to maintain fluency in the face of interference. What is missing in this literature is a better understanding of interference during production planning and maintenance, and how alternative word orders emerge in the face of this interference. These questions seem ripe for insight from and collaboration with VWM research.

Implications for language comprehension

Theories of language comprehension aim to explain how language percepts are recognized and interpreted. Important data in this endeavor has been measures of comprehension difficulty, or, more specifically, the relative difficulty of some kind of language compared to another. In the case of sentence-level comprehension research, the focus has been on why some kinds of sentences are harder than others, and VWM capacity has been a common explanatory factor in this field (MacDonald & Hsiao, 2018). Many researchers have invoked decay in VWM to explain comprehension difficulty of certain kinds of sentences, as the difficult sentences require integration over distant information that has degraded in working memory (Babyonyshev & Gibson, 1999; Gibson, 1998; Grodner & Gibson, 2005; Just & Carpenter, 1992). An alternative approach suggests that VWM and comprehension difficulty are constrained by interference rather than decay or capacity limitations (Glaser, Martin, Van Dyke, Hamilton, & Tan, 2013; Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & Johns, 2012). This work emphasizes that both encoding and retrieval of information becomes more difficult with increased semantic similarity between words, meaning that sentences with more interfering elements are more difficult to comprehend (see Van Dyke & Johns, 2012, for review). This area is therefore another in which VWM research could inform comprehension, particularly the influence of decay and/or interference (Oberauer, Farrell, Jarrold, & Lewandowsky, 2016). More generally, though, while language comprehension researchers have often invoked VWM limitations in accounts of comprehension difficulty, they have not necessarily aligned themselves with particular VWM models of encoding, maintenance, and retrieval processes (for some exceptions, see Caplan & Waters, 2013; Just & Carpenter, 1992; Lewis, Vasishth, & Van Dyke, 2006; Martin & Romani, 1994).

At least initially, very few accounts of language comprehension ascribed a major role for experience in language comprehension difficulty, and thus these accounts were at least in principle aligned with a multicomponent perspective, in which a separate temporary store, separate from language knowledge, provided a bottleneck in encoding and maintenance that could explain comprehension difficulty. More recently, a number of researchers have suggested that both VWM capacity and language experience are important components in processing difficulty (Demberg & Keller, 2008; Staub, 2010). In a more fully emergent approach of VWM, the capacity to encode and maintain information (whether for everyday language use or a working memory task) is not independent of long-term memory, and thus not independent of experience with language (Acheson & MacDonald, 2009a; Botvinick & Plaut, 2006; G. Jones & Macken, 2015; MacDonald & Christiansen, 2002; McClelland & Elman, 1986). We see this emphasis on experience-based capacity as a basis for investigating parallels between comprehension processes and VWM. Moreover, the emphasis on experience also casts language use and memory as intertwined, learned skills, as noted in the discussion of revised interpretations of VWM tasks above. For example, memory researchers have noted relationships between novel word learning and the Hebb repetition effect (Szmalec, Duyck, Vandierendonck, Mata, & Page, 2009). If word representations are highly intertwined, as our emergent perspective claims, then sensitivity to the Hebb repetition effect and novel word learning should exhibit exploitation of statistical regularities between different sources of information (e.g. Cassidy & Kelly, 1991; Nygaard, Cook, & Namy, 2009) rather than mere memory capacity of the learner.

Conclusions

In this article, we have aimed to describe the rich nature of linguistic LTM and its consequences for VWM. While Ebbinghaus (1885) had inklings that LTM could not be fully set aside in studying VWM, we have suggested that the linkage between language LTM and VWM is far stronger than he imagined, in part because the LTM has a different quality than he and many others had hypothesized. A more thorough understanding of the nature of language processing, attention, and LTM, we claim, will accelerate the advancement of both VWM and language research. We have argued that words are not unrelated islands in LTM representations, and therefore they should not be treated as isolated items in VWM research. We have further argued that the processes of language comprehension and production underlie a person's ability to encode, maintain, and order verbal information. These skills are essential for everyday language use, change with experience and the richness of LTM, and are brought to bear on

VWM tasks. On this view, VWM and language research should be mutually informative.

References

Aben, B., Stapert, S., & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00301

Acheson, D. J., Hamidi, M., Binder, J. R., & Postle, B. R. (2011). A common neural substrate for language production and verbal working memory. *Journal of Cognitive Neuroscience*, *23* (6), 1358–1367. https://doi.org/10.1162/jocn.2010.21519

Acheson, D. J., & MacDonald, M. C. (2009a). Twisting tongues and memories: Explorations of the relationship between language production and verbal working memory. *Journal of Memory and Language*, *60*(3), 329–350. https://doi.org/10.1016/j.jml.2008.12.002.

Acheson, D. J., & MacDonald, M. C. (2009b). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, *135*(1), 50–68. https://doi.org/10.1037/a0014411

Acheson, D. J., MacDonald, M. C., & Postle, B. R. (2011). The effect of concurrent semantic categorization on delayed serial recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(1), 44–59. https://doi.org/10.1037/a0021205

Acheson, D. J., Postle, B. R., & Macdonald, M. C. (2010). The interaction of concreteness and phonological similarity in verbal working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(1), 17–36. https://doi.org/10.1037/a0017679

Allen, J., & Seidenberg, M. S. (1999). The emergence of grammaticality in connectionist networks. In B. MacWhinney (Ed.), *The emergence of language.* (pp. 115–151). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, *55*(1), 64–88. https://doi.org/10.1016/j.jml.2006.02.002

Allen, R. J., Hitch, G. J., & Baddeley, A. D. (2018). Exploring the sentence advantage in working memory: Insights from serial recall and recognition. *Quarterly Journal of Experimental Psychology*, *71*(12), 2571–2585. https://doi.org/10.1177/1747021817746929

Allport, DA., & Funnell, E. (1981). Components of the mental lexicon. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *295*, 397–410.

Almeida, R. G. D., & Gleitman, L. R. (2018). *On Concepts, Modules, and Language: Cognitive Science at Its Core*. Oxford University Press.

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science*, *33*(4), 583–609. https://doi.org/10.1111/j.1551-6709.2009.01022.x

Anderson, N. D., & Dell, G. S. (2018). The role of consolidation in learning context-dependent phonotactic patterns in speech and digital sequence production. *Proceedings of the National Academy of Sciences*, 201721107. https://doi.org/10.1073/pnas.1721107115

Anderson, N. D., Holmes, E. W., Dell, G. S., & Middleton, E. L. (2019). Reversal shift in phonotactic learning during language production: Evidence for incremental learning. *Journal of Memory and Language*, *106*, 135–149. https://doi.org/10.1016/j.jml.2019.03.002

Archibald Lisa M. D., & Gathercole Susan E. (2006). Nonword Repetition: A Comparison of Tests. *Journal of Speech, Language, and Hearing Research*, *49*(5), 970–983. https://doi.org/10.1044/1092-4388(2006/070)

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*(1), 67–82. https://doi.org/10.1016/j.jml.2009.09.005

Attout, L., Kaa, M.-A. V. der, George, M., & Majerus, S. (2012). Dissociating short-term memory and language impairment: The importance of item and serial order information. *Aphasiology*, *26*(3–4), 355–382. https://doi.org/10.1080/02687038.2011.604303

Attout, L., Magro, L. O., Szmalec, A., & Majerus, S. (2019). The developmental neural substrates of item and serial order components of verbal working memory. *Human Brain Mapping*, *40*(5), 1541–1553. https://doi.org/10.1002/hbm.24466

Baayen, R. H., Shaoul, C., Willits, J., & Ramscar, M. (2016). Comprehension without segmentation: A proof of concept with naive discriminative learning. *Language, Cognition and Neuroscience*, *31*(1), 106–128. https://doi.org/10.1080/23273798.2015.1065336

Babyonyshev, M., & Gibson, E. (1999). The Complexity of Nested Structures in Japanese. *Language*, *75*(3), 423–450.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology*, *18*(4), 302–309. https://doi.org/10.1080/14640746608400047

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*(3), 438–456. https://doi.org/10.1016/j.jml.2009.05.004

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning & Verbal Behavior*, *9*(2), 176–189. https://doi.org/10.1016/S0022-5371(70)80048-2

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. https://doi.org/10.1016/S0022-5371(75)80045-4

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the Articulatory Loop. *The Quarterly Journal of Experimental Psychology Section A*, *36*(2), 233–252. https://doi.org/10.1080/14640748408402157

Barthel, M., & Sauppe, S. (2019). Speech Planning at Turn Transitions in Dialog Is Associated With Increased Processing Load. *Cognitive Science*, *43*(7), e12768. https://doi.org/10.1111/cogs.12768

Basso, A., Spinnler, H., Vallar, G., & Zanobio, M. E. (1982). Left hemisphere damage and selective impairment of auditory verbal short-term memory. A case study. *Neuropsychologia*, *20*(3), 263–274. https://doi.org/10.1016/0028-3932(82)90101-4

Bock, J. K. (1986a). Meaning, Sound, and Syntax: Lexical Priming in Sentence Production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 575–586. https://doi.org/10.1037/0278-7393.12.4.575

Bock, J. K. (1986b). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K. (1987). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, *26*(2), 119–137. https://doi.org/10.1016/0749-596X(87)90120-3

Boiteau, T. W., Malone, P. S., Peters, S. A., & Almor, A. (2014). Interference between conversation and a concurrent visuomotor task. *Journal of Experimental Psychology. General*, *143*(1), 295–311. https://doi.org/10.1037/a0031858

Botvinick, M., & Bylsma, L. M. (2005). Regularization in Short-Term Memory for Serial Order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 351–358. https://doi.org/10.1037/0278-7393.31.2.351

Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review*, *113*(2), 201–233. https://doi.org/10.1037/0033-295X.113.2.201

Botvinick, M. M., & Plaut, D. C. (2009). Empirical and computational support for context-dependent representations of serial order: Reply to Bowers, Damian, and Davis (2009). *Psychological Review*, *116*(4), 998–1001. https://doi.org/10.1037/a0017113

Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology, 26*(5), 467–482. https://doi.org/10.1037/h0061096

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. *Cognitive Foundations of Interpretation*, 69–94.

Buchsbaum, B. R., & D'Esposito, M. (2018). A sensorimotor view of verbal working memory. *Cortex*, *112*, 134–148. https://doi.org/10.1016/j.cortex.2018.11.010

Buchsbaum, B. R., & D'Esposito, M. (2019). A sensorimotor view of verbal working memory. *Cortex*, *112*, 134–148. https://doi.org/10.1016/j.cortex.2018.11.010

Bybee, J., & McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *Linguistic Review*, *22* (2–4), 381–410. https://doi.org/10.1515/tlir.2005.22.2-4.381

Caplan, D., & Waters, G. (2013). Memory mechanisms supporting syntactic comprehension. *Psychonomic Bulletin and Review,* 20(2), 243–268. doi:10.3758/s13423-012-0369-9

Casalini, C., Brizzolara, D., Chilosi, A., Cipriani, P., Marcolini, S., Pecini, C., Roncoli, S., Burani, C. (2007). Non-Word Repetition in Children with Specific Language Impairment: A Deficit in Phonological Working Memory or in Long-Term Verbal Knowledge? *Cortex*, *43*(6), 769–776. https://doi.org/10.1016/S0010-9452(08)70505-7

Cassidy, K. W., & Kelly, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language*, *30*(3), 348–369. https://doi.org/10.1016/0749-596X(91)90041-H

Cave, C. B., & Squire, L. R. (1992). Intact verbal and nonverbal short-term memory following damage to the human hippocampus. *Hippocampus*, *2*(2), 151–163. https://doi.org/10.1002/hipo.450020207

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*(2), 217–230.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming Syntactic. *Psychological Review*, *113*(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998, June 1). Learning to Segment Speech Using Multiple Cues: A Connectionist Model [Text]. https://doi.org/info:doi/10.1080/016909698386528

Christiansen, M. H., & Monaghan, P. (2016). Division of Labor in Vocabulary Structure: Insights From Corpus Analyses. *Topics in Cognitive Science*, *8*(3), 610–624. https://doi.org/10.1111/tops.12164

Clarkson, L., Roodenrys, S., Miller, L. M., & Hulme, C. (2017). The phonological neighbourhood effect on short-term memory for order. *Memory*, *25*(3), 391–402. https://doi.org/10.1080/09658211.2016.1179330

Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term-memory recall of repeated and novel word lists. *Memory & Cognition*, *21*(4), 539–545. https://doi.org/10.3758/BF03197185

Connine, C. M., & Clifton, C. Jr. (1987). Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(2), 291–299. https://doi.org/10.1037/0096-1523.13.2.291

Cowan, N. (1993). Activation, attention, and short-term memory. *Memory & Cognition*, *21*(2), 162–167. https://doi.org/10.3758/BF03202728

Cowan, N. (2008). Chapter 20 What are the differences between long-term, short-term, and working memory? In W. S. Sossin, J.-C. Lacaille, V. F. Castellucci, & S. Belleville (Eds.), *Progress in Brain Research* (pp. 323–338). https://doi.org/10.1016/S0079-6123(07)00020-9

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition*, *21*(2), 142–145. https://doi.org/10.3758/BF03202725

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Dapretto, M., & Bookheimer, S. Y. (1999). Form and Content: Dissociating Syntax and Semantics in Sentence Comprehension. *Neuron*, *24*(2), 427–432. https://doi.org/10.1016/S0896-6273(00)80855-7

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, 515–535.

Dell, G. S. (1984). Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(2), 222–233. https://doi.org/10.1037/0278-7393.10.2.222

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language Production and Serial Order: A Functional Analysis and a Model. *Psychological Review*, *104*(1), 123–147.

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1634), 20120394. https://doi.org/10.1098/rstb.2012.0394

Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech Errors, Phonotactic Constraints, and Implicit Learning: A Study of the Role of Experience in Language Production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26* (6), 1355–1367.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 611–629. https://doi.org/10.1016/S0022-5371(81)90202-4

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review*, *104*(4), 801–838.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*(2), 193–210. https://doi.org/10.1016/j.cognition.2008.07.008

Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, *21*(5), 629–634. https://doi.org/10.1177/0956797610367751

Duff, M. C., & Brown-Schmidt, S. (2012). The hippocampus and the flexible use and processing of language. *Frontiers in Human Neuroscience*, *6*, 69. https://doi.org/10.3389/fnhum.2012.00069

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*. New York, N.Y.: Dover.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research: JSLHR*, *47*(2), 421–436. https://doi.org/10.1044/1092-4388(2004/034)

Ellis, A. W. (1980). Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior*, *19*(5), 624–634.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, *7*(2–3), 195–225.

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences of the*

United States of America, *103*(32), 12203–12208. https://doi.org/10.1073/pnas.0602173103

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505. https://doi.org/10.1111/j.1469-8986.2007.00531.x

Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Language Sciences*, *5*, 335. https://doi.org/10.3389/fpsyg.2014.00335

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, *108*(39), 16428–16433.

Ferreira, F. (1994). Choice of Passive Voice is Affected by Verb Type and Animacy. *Journal of Memory and Language*, *33*(6), 715–736. https://doi.org/10.1006/jmla.1994.1034

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid Expectation Adaptation during Syntactic Comprehension. *PloS One*, *8*(10), e77661. https://doi.org/10.1371/journal.pone.0077661

Fischer-Baum, S. (2018). Chapter Two—A Common Representation of Serial Position in Language and Memory. In K. D. Federmeier & D. G. Watson (Eds.), *Psychology of Learning and Motivation* (pp. 31–54). https://doi.org/10.1016/bs.plm.2018.08.002

Fischer-Baum, S., & McCloskey, M. (2015). Representation of Item Position in Immediate Serial Recall: Evidence From Intrusion Errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, No Pagination Specified. https://doi.org/10.1037/xlm0000102

Forster, K. I. (1985). Lexical acquisition and the modular lexicon. *Language and Cognitive Processes*, *1*(2), 87–108. https://doi.org/10.1080/01690968508402073

Frank, M. C., & Goodman, N. D. (2014). Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology*, *75*, 80–96. https://doi.org/10.1016/j.cogpsych.2014.08.002

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII* (pp. 559–586). Retrieved from http://doi.apa.org/psycinfo/1987-98557-025

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 84–95. https://doi.org/10.1037/0278-7393.25.1.84

Gennari, S. P., Mirković, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology*, *65*(2), 141–176. https://doi.org/10.1016/j.cogpsych.2012.03.002

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Glaser, Y. G., Martin, R. C., Van Dyke, J. A., Hamilton, A. C., & Tan, Y. (2013). Neural basis of semantic and syntactic interference in sentence comprehension. *Brain and Language*, *126* (3), 314–326.

Epstein, W. (1961). The Influence of Syntactical Structure on Learning. *The American Journal of Psychology, 74*(1), 80-85. doi:10.2307/1419827

Epstein, W. (1962). A Further Study of the Influence of Syntactical Structure on Learning. *The American Journal of Psychology, 75*(1), 121-126. doi:10.2307/1419550

Estes, K. G., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the Nonword Repetition Performance of Children With and Without Specific Language Impairment: A Meta-Analysis. *Journal of Speech, Language, and Hearing Research*, *50*(1), 177–195. https://doi.org/10.1044/1092-4388(2007/015)

Graves, W. W., Binder, J. R., Desai, R. H., Conant, L. L., & Seidenberg, M. S. (2010). Neural correlates of implicit and explicit combinatorial semantic processing. *NeuroImage*, *53*(2), 638–646. https://doi.org/10.1016/j.neuroimage.2010.06.055

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentenial complexity. *Cognitive Science*, *29*(2), 261–290.

Guerrette, M.-C., Saint-Aubin, J., Richard, M., & Guérard, K. (2018). Overt language production plays a key role in the Hebb repetition effect. *Memory & Cognition*, *46*(8), 1389–1397. https://doi.org/10.3758/s13421-018-0844-2

Guidali, G., Pisoni, A., Bolognini, N., & Papagno, C. (2019). Keeping order in the brain: The supramarginal gyrus and serial order in short-term memory. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *119*, 89–99. https://doi.org/10.1016/j.cortex.2019.04.009

Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language*, *61*(4), 481–502.

Hannula, D. E., Tranel, D., & Cohen, N. J. (2006). The long and the short of it: relational memory impairments in amnesia, even at short lags. *Journal of Neuroscience*, *26*(32), 8352-8359. https://doi.org/10.1523/JNEUROSCI.5222-05.2006

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language*, *35*(1), 1–31. https://doi.org/10.1006/jmla.1996.0001

Hartley, T., Hurlstone, M. J., & Hitch, G. J. (2016). Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism. *Cognitive Psychology*, *87*, 135–178. https://doi.org/10.1016/j.cogpsych.2016.05.001

Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: Memory as an integral component of information processing. *Trends in Cognitive Sciences*, *19*(6), 304–313. https://doi.org/10.1016/j.tics.2015.04.006

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye, (Ed.), *Distinctive features of learning in the higher animal. Brain mechanisms and learning Blackwell 37–46.* (pp. 37–46). Blackwell.

Henson, R., Hartley, T., Burgess, N., Hitch, G., & Flude, B. (2003). Selective interference with verbal short-term memory for serial order information: A new paradigm and tests of a timing-signal hypothesis. *Quarterly Journal of Experimental Psychology Section A*, *56* (8), 1307–1334.

Hepner, C. R., & Nozari, N. (2019). Resource allocation in phonological working memory: Same or different principles from vision? *Journal of Memory and Language*, *106*, 172–188. https://doi.org/10.1016/j.jml.2019.03.003

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, *98*(1), 74–95. https://doi.org/10.1037/0033-295X.98.1.74

Hoffman, P., Jefferies, E., Ehsan, S., Jones, R. W., & Lambon Ralph, M. A. (2009). Semantic memory is key to binding phonology: Converging evidence from immediate serial recall in semantic dementia and healthy participants. *Neuropsychologia*, *47*(3), 747–760. https://doi.org/10.1016/j.neuropsychologia.2008.12.001

Howard, D., & Nickels, L. (2005). Separating input and output phonology: Semantic, phonological, and orthographic effects in short-term memory impairment. *Cognitive Neuropsychology*, *22*(1), 42–77. https://doi.org/10.1080/02643290342000582

Hsiao, Y., Gao, Y., & MacDonald, M. C. (2014). Agent-patient similarity affects sentence structure in language production: Evidence from subject omissions in Mandarin. *Frontiers in Psychology*, *5*, 1015.

Hughes, R. W., Chamberland, C., Tremblay, S., & Jones, D. M. (2016). Perceptual-motor determinants of auditory-verbal serial short-term memory. *Journal of Memory and Language*, *90*, 126–146.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1217.

Hulme, C., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology*, *47*(1), 72–87. https://doi.org/10.1016/0022-0965(89)90063-5

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, *140*(2), 339–373. https://doi.org/10.1037/a0034221

Ingvalson, E. M., Dhar, S., Wong, P. C. M., & Liu, H. (2015). Working memory training to improve speech perception in noise across languages. *The Journal of the Acoustical Society of America*, *137*(6), 3477–3486. https://doi.org/10.1121/1.4921601

Jacquemot, C., Dupoux, E., Decouche, O., & Bachoud-Lévi, A.-C. (2006). Misperception in sentences but not in words: Speech perception and the phonological buffer. *Cognitive Neuropsychology*, *23*(6), 949–971. https://doi.org/10.1080/02643290600625749

Jefferies, E., Jones, R. W., Bateman, D., & Ralph, M. A. L. (2005). A semantic contribution to nonword recall? Evidence for intact phonological processes in semantic dementia. *Cognitive Neuropsychology*, *22*(2), 183–212. https://doi.org/10.1080/02643290442000068

Joanisse, M. F., & McClelland, J. L. (2015). Connectionist perspectives on language learning, representation and processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, n/a-n/a. https://doi.org/10.1002/wcs.1340

Joanisse, M. F., & Seidenberg, M. S. (2003). Phonology and syntax in specific language impairment: Evidence from a connectionist model. *Brain and Language*, *86*(1), 40–56.

Jones, T., & Farrell, S. (2018). Does syntax bias serial order reconstruction of verbal short-term memory? *Journal of Memory and Language*, *100*, 98–122. https://doi.org/10.1016/j.jml.2018.02.001

Jones, G., Justice, L. V., Cabiddu, F., Lee, B. J., Iao, L. S., Harrison, N., & Macken, B. (2020). Does short-term memory develop?. *Cognition*. https://doi.org/10.1016/j.cognition.2020.104200

Jones, D., & Macken, W. (2018). In the beginning was the deed: Verbal short-term memory as object-oriented action. *Current Directions In Psychological Science*, *27*, 351–356. https://doi.org/10.1177/0963721418765796

Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition*, *144*, 1–13.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(3), 656–674. https://doi.org/10.1037/0278-7393.30.3.656

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122.

Kaiser, E. (2012). Taking Action: A Cross-Modal Investigation of Discourse-Level Representations. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00156

Kalm, K., & Norris, D. (2014). The Representation of Order Information in Auditory-Verbal Short-Term Memory. *Journal of Neuroscience*, *34*(20), 6879–6886. https://doi.org/10.1523/JNEUROSCI.4104-13.2014

Kemper, S., Hoffman, L., Schmalzried, R., Herman, R., & Kieweg, D. (2011). Tracking talking: Dual task costs of planning and producing speech for young versus older adults. *Aging, Neuropsychology, and Cognition*, *18*(3), 257–279. https://doi.org/10.1080/13825585.2010.527317

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S.-A. H., Gustafsson, J.-E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, *18*, 146–154. https://doi.org/10.1111/desc.12202

Kolers, P. A., & Roediger III, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, *23*(4), 425–449. https://doi.org/10.1016/S0022-5371(84)90282-2

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Leminen, A., Smolka, E., Duñabeitia, J. A., & Pliatsikas, C. (2019). Morphological processing in the brain: The good (inflection), the bad (derivation) and the ugly (compounding). *Cortex*, *116*, 4–44. https://doi.org/10.1016/j.cortex.2018.08.016

Levelt, W. J. (1993). *Speaking: From intention to articulation* (Vol. 1). Retrieved from https://books-google-com.ezproxy.library.wisc.edu/books?hl=en&lr=&id=LbVCdCE-NQAC&oi=fnd&pg=PR13&dq=willem+levelt&ots=JZgFOf6cxM&sig=RJgg78qN4IqG-ZlDkHjZ3YQqE0k

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1–38.

Levinson, S. C. (2016). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, *20*(1), 6–14.

https://doi.org/10.1016/j.tics.2015.10.010

Lewandowsky, S., & Farrell, S. (2000). A redintegration account of the effects of speech rate, lexicality, and word frequency in immediate serial recall. *Psychological Research*, *63*(2), 163–173.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454.

Lombardi, L., & Potter, M. C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language*, *31*(6), 713–733. https://doi.org/10.1016/0749-596X(92)90036-W

MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. *Handbook of Psycholinguistics*, *2*, 581–611.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*(2), 157–201. https://doi.org/10.1080/01690969408402115

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 226. https://doi.org/10.3389/fpsyg.2013.00226

MacDonald, M. C. (2016). Speak, act, remember: The language-production basis of serial order and maintenance in verbal memory. *Current Directions in Psychological Science*, *25*(1), 47–53. https://doi.org/10.1177/0963721415620776

MacDonald, M. C, & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54; discussion 55-74.

MacDonald, M. C, & Hsiao, Y. (2018). Sentence Comprehension. In S.-A. Rueschemeyer & G. Gaskell (Eds.), *Oxford Handbook of Psycholinguistics* (2nd Edition, pp. 171–196). Oxford, UK: Oxford University Press.

Macken, W. J., Taylor, J. C., & Jones, D. M. (2014). Language and short-term memory: The role of perceptual-motor affordance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1257–1270. doi:10.1037/a0036845

Majerus, S. (2009). Verbal short-term memory and temporary activation of language representations: The importance of distinguishing item and order information. In *Interactions between short-term and long-term memory in the verbal domain* (pp. 244–276). New York, NY, US: Psychology Press.

Majerus, S. (2013). Language repetition and short-term memory: An integrative framework. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00357

Majerus, S. (2019). Verbal working memory and the phonological buffer: The question of serial order. *Cortex*, *112*, 122–133. https://doi.org/10.1016/j.cortex.2018.04.016

Majerus, S., Attout, L., Artielle, M.-A., & Van der Kaa, M.-A. (2015). The heterogeneity of verbal short-term memory impairment in aphasia. *Neuropsychologia*, *77*, 165–176.

Majerus, S., Belayachi, S., De Smedt, B., Leclercq, A. L., Martinez, T., Schmidt, C., Weekes, B., & Maquet, P. (2008). Neural networks for short-term memory for order differentiate high and low proficiency bilinguals. *Neuroimage*, *42*(4), 1698-1713. https://doi.org/10.1016/j.neuroimage.2008.06.003

Majerus, S., & D'Argembeau, A. (2011). Verbal short-term memory reflects the organization of long-term memory: Further evidence from short-term memory for emotional words. *Journal of Memory and Language*, *64*(2), 181–197. https://doi.org/10.1016/j.jml.2010.10.003

Majerus, S., Norris, D., & Patterson, K. (2007). What does a patient with semantic dementia remember in verbal short-term memory? Order and sound but not words. *Cognitive Neuropsychology*, *24*(2), 131-151. https://doi.org/10.1080/02643290600989376

Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., & Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language*, *51*(2), 297–306.

Martin, N., Dell, G. S., Saffran, E. M., & Schwartz, M. F. (1994). Origins of Paraphasias in Deep Dysphasia: Testing the Consequences of a Decay Impairment to an Interactive Spreading Activation Model of Lexical Retrieval. *Brain and Language*, *47*(4), 609–660. https://doi.org/10.1006/brln.1994.1061

Majerus, S., Poncelet, M., Van der Linden, M., Albouy, G., Salmon, E., Sterpenich, V., Vandewalle, G., Collette, F., & Maquet, P. (2006). The left intraparietal sulcus and verbal short-term memory: Focus of attention or serial order?. *Neuroimage*, *32*(2), 880-891. https://doi.org/10.1016/j.neuroimage.2006.03.048

Martin, N., & Saffran, E. M. (1992). A computational account of deep dysphasia: Evidence from a single case study. *Brain and Language*, *43*(2), 240–274. https://doi.org/10.1016/0093-934X(92)90130-7

Martin, N., Saffran, E. M., & Dell, G. S. (1996). Recovery in deep dysphasia: Evidence for a relation between auditory–verbal STM capacity and lexical errors in repetition. *Brain and Language*, *52*(1), 83–113. https://doi.org/10.1006/brln.1996.0005

Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language Processing and Working Memory: Neuropsychological Evidence for Separate Phonological and Semantic Capacities. *Journal of Memory and Language*, *33*(1), 83–111. https://doi.org/10.1006/jmla.1994.1005

Martin, R. C. (1987). Articulatory and phonological deficits in short-term memory and their relation to syntactic processing. *Brain and Language*, *32*(1), 159–192. https://doi.org/10.1016/0093-934X(87)90122-2

Martin, R. C. (1993). Short-term memory and sentence processing: Evidence from neuropsychology. *Memory & Cognition*, *21*(2), 176–183. https://doi.org/10.3758/BF03202730

Martin, R. C., & Breedin, S. D. (1992). Dissociations between speech perception and phonological short-term memory deficits. *Cognitive Neuropsychology*, *9*(6), 509–534. https://doi.org/10.1080/02643299208252070

Martin, R. C., & Freedman, M. L. (2001). Short-term retention of lexical-semantic representations: Implications for speech production. *Memory (Hove, England)*, *9*(4), 261–280. https://doi.org/10.1080/09658210143000173

Martin, R. C, & He, T. (2004). Semantic short-term memory and its role in sentence processing: A replication. *Brain and Language*, *89*(1), 76–82. https://doi.org/10.1016/S0093-934X(03)00300-6

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of Input and Output Phonology in Word Processing and Short-Term Memory. *Journal of Memory and Language*, *41*(1), 3–29. https://doi.org/10.1006/jmla.1999.2637

Martin, R. C., & Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple-components view. *Neuropsychology*, *8*(4), 506–523. https://doi.org/10.1037/0894-4105.8.4.506

Mayberry, M. R., Crocker, M. W., & Knoeferle, P. (2009). Learning to attend: A connectionist model of situated language comprehension. *Cognitive Science*, *33*(3), 449–496.

McCauley, S. M., & Christiansen, M. H. (2014). Acquiring formulaic language: A computational model. *The Mental Lexicon*, *9*(3), 419–436. https://doi.org/10.1075/ml.9.3.03mcc

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences*, *14*(8), 348–356. https://doi.org/10.1016/j.tics.2010.06.002

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Miller, G. A., & Selfridge, J. A. (1950). Verbal context and the recall of meaningful material. *The American Journal of Psychology*, *63*, 176–185. https://doi.org/10.2307/1418920

Monaghan, P., & Woollams, A. M. (2017). Implementing the "Simple" model of reading deficits: A connectionist investigation of interactivity. In *Neurocomputational Models of Cognitive Development and Processing: Proceedings of the 14th Neural Computation and Psychology Workshop* (pp. 69–81).

Montag, J. L., Matsuki, K., Kim, J. Y., & MacDonald, M. C. (2017). Language Specific and Language General Motivations of Production Choices: A Multi-Clause and Multi-Language Investigation. *Collabra: Psychology*, *3*(1). https://doi.org/10.1525/collabra.94

Murdock Jr., B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*(5), 482–488. https://doi.org/10.1037/h0045106

Norris, D. (2017). Short-term memory and long-term memory are still different. *Psychological Bulletin*, *143*(9), 992–1009. https://doi.org/10.1037/bul0000108

Nozari, N., & Dell, G. S. (2012). Feature migration in time: Reflection of selective attention on speech errors. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(4), 1084–1090. https://doi.org/10.1037/a0026933

Nozari, N., & Novick, J. (2017). Monitoring and Control in Language Production. *Current Directions in Psychological Science*, *26*(5), 403–410. https://doi.org/10.1177/0963721417702419

Nozari, N., & Thompson-Schill, S. L. (2013). More attention when speaking: Does it help or does it hurt? *Neuropsychologia*, *51*(13), 2770–2780. https://doi.org/10.1016/j.neuropsychologia.2013.08.019

Nozari, N., Trueswell, J. C., & Thompson-Schill, S. L. (2016). The interplay of local attraction, context and domain-general cognitive control in activation and suppression of semantic distractors during sentence comprehension. *Psychonomic Bulletin & Review*, *23*(6), 1942–1953. https://doi.org/10.3758/s13423-016-1068-8

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition*, *112*(1), 181–186. https://doi.org/10.1016/j.cognition.2009.04.001

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, *142*(7), 758–799. https://doi.org/10.1037/bul0000046

Oberauer, K., Jones, T., & Lewandowsky, S. (2015). The Hebb repetition effect in simple and complex memory span. *Memory & Cognition*, *43*(6), 852–865. https://doi.org/10.3758/s13421-015-0512-8

Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3737–3753. https://doi.org/10.1098/rstb.2009.0173

Page, M. P., Madge, A., Cumming, N., & Norris, D. G. (2007). Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus. *Journal of Memory and Language*, *56*(1), 49–64.

Page, M. P. A., Cumming, N., Norris, D., McNeil, A. M., & Hitch, G. J. (2013). Repetition-spacing and item-overlap effects in the Hebb repetition task. *Journal of Memory and Language*, *69*(4), 506–526. https://doi.org/10.1016/j.jml.2013.07.001

Patterson, K., Graham, N., & Hodges, J. R. (1994). The impact of semantic memory loss on phonological representations. *Journal of Cognitive Neuroscience*, *6*(1), 57–69. https://doi.org/10.1162/jocn.1994.6.1.57

Penfield, W., & Milner, B. (1958). Memory Deficit Produced by Bilateral Lesions in the Hippocampal Zone. *A.M.A. Archives of Neurology & Psychiatry*, *79*(5), 475–497. https://doi.org/10.1001/archneurpsyc.1958.02340050003001

Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., Botvinivk, M., Fedorenko, E. (2018). Toward a universal decoder of linguistic meaning from brain activation. *Nature Communications*, *9*(1), 963. https://doi.org/10.1038/s41467-018-03068-4

Perham, N., Marsh, J. E., & Jones, D. M. (2009). Syntax and serial recall: How language supports short-term memory for order. *The Quarterly Journal of Experimental Psychology*, *62*(7), 1285–1293. https://doi.org/10.1080/17470210802635599

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*(3), 427–459. https://doi.org/10.1037/0033-2909.134.3.427

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

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56–115.

Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology/Revue Canadienne de*

Psychologie Expérimentale, *50*(4), 408–412. https://doi.org/10.1037/1196-1961.50.4.408

Poirier, M., Saint-Aubin, J., Mair, A., Tehan, G., & Tolan, A. (2014). Order recall in verbal short -term memory: The role of semantic networks. *Memory & Cognition*, *43*(3), 489–499. https://doi.org/10.3758/s13421-014-0470-6

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38. https://doi.org/10.1016/j.neuroscience.2005.06.005

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language*, *38*(3), 265–282.

Ramscar, M., & Port, R. F. (2016). How spoken languages work in the absence of an inventory of discrete units. *Language Sciences*, *53*, 58–74. https://doi.org/10.1016/j.langsci.2015.08.002

Roodenrys, S., & Hinton, M. (2002). Sublexical or lexical effects on serial recall of nonwords? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 29–33. https://doi.org/10.1037/0278-7393.28.1.29

Roodenrys, S., Hulme, C., Lethbridge, A., Hinton, M., & Nimmo, L. M. (2002). Word-frequency and phonological-neighborhood effects on verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1019.

Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J., & van der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. *Human Movement Science*, *26*(4), 525–554. https://doi.org/10.1016/j.humov.2007.04.001

Rubin, R. D., Watson, P. D., Duff, M. C., & Cohen, N. J. (2014). The role of the hippocampus in flexible cognition and social behavior. *Frontiers in Human Neuroscience*, *8*, 742. https://doi.org/10.3389/fnhum.2014.00742

Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C. S., & Mars, F. (1997). Morphological Priming, Fragment Completion, and Connectionist Networks. *Journal of Memory and Language*, *36*(3), 382–405. https://doi.org/10.1006/jmla.1996.2489

Saffran, E. M., & Martin, N. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, *14*(5), 641–682.

Saint-Aubin, J., & Poirier, M. (1999). The Influence of Long-term Memory Factors on Immediate Serial Recall: An Item and Order Analysis. *International Journal of Psychology*, *34*(5–6), 347–352. https://doi.org/10.1080/002075999399675

Savill, N., Cornelissen, P., Whiteley, J., Woollams, A., & Jefferies, E. (2019). Individual differences in verbal short-term memory and reading aloud: Semantic compensation for weak phonological processing across tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), 1815–1831. https://doi.org/10.1037/xlm0000675

Savill, N., Ellis, R., Brooke, E., Koa, T., Ferguson, S., Rojas-Rodriguez, E., Arnold, D., Smallwood, J., Jefferies, E. (2017). Keeping it together: Semantic coherence stabilizes phonological sequences in short-term memory. *Memory & Cognition*, 1–12. https://doi.org/10.3758/s13421-017-0775-3

Savill, N. J., Cornelissen, P., Pahor, A., & Jefferies, E. (2019). RTMS evidence for a dissociation in short-term memory for spoken words and nonwords. *Cortex*, *112*, 5–22.

https://doi.org/10.1016/j.cortex.2018.07.021

Schmidtke, D., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00080

Scoville, W. B., & Milner, B. (1957). Loss of Recent Memory After Bilaterla Hippocampal Lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*(1), 11–21.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568.

Seidenberg, M. S., Plaut, D. C., Petersen, A. S., McClelland, J. L., & McRae, K. (1994). Nonword pronunciation and models of word recognition. *Journal of Experimental Psychology. Human Perception and Performance*, *20*(6), 1177–1196.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. (Cover story). *Science*, *275*(5306), 1599. Retrieved from aph.

Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, *4*(9), 353–361. https://doi.org/10.1016/S1364-6613(00)01515-1

Seidenberg, M. S., & MacDonald, M. C. (2018). The Impact of Language Experience on Language and Reading: A Statistical Learning Approach. *Topics in Language Disorders*, *38*(1), 66–83. https://doi.org/10.1097/TLD.0000000000000144

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 546–554. https://doi.org/10.1037/0278-7393.5.6.546

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, *14*(4), 489–537.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology*, *22*(2), 261–273. https://doi.org/10.1080/00335557043000203

Shallice, T., & Butterworth, B. (1977). Short-term memory impairment and spontaneous speech. *Neuropsychologia*, *15*(6), 729–735. https://doi.org/10.1016/0028-3932(77)90002-1

Shallice, T., & Papagno, C. (2019). Impairments of auditory-verbal short-term memory: Do selective deficits of the input phonological buffer exist? *Cortex*, *112*, 107–121. https://doi.org/10.1016/j.cortex.2018.10.004

Shallice, T., & Warrington, E. K. (1974). The dissociation between short term retention of meaningful sounds and verbal material. *Neuropsychologia*, *12*(4), 553–555. https://doi.org/10.1016/0028-3932(74)90087-6

Siegelman, M., Blank, I. A., Mineroff, Z., & Fedorenko, E. (2019). An Attempt to Conceptually Replicate the Dissociation between Syntax and Semantics during Sentence Comprehension. *Neuroscience*, *413*, 219–229. https://doi.org/10.1016/j.neuroscience.2019.06.003

Silveri, M. C., & Cappa, A. (2003). Segregation of the Neural Correlates of Language and Phonological Short-Term Memory. *Cortex*, *39*(4), 913–925. https://doi.org/10.1016/S0010-9452(08)70870-0

Solomon, E. S., & Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, *49*(1), 1–46. https://doi.org/10.1016/j.cogpsych.2003.10.001

Soveri, A., Antfolk, J., Karlsson, L. Salo, B., Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin and Review* 24(4), 1077–1096, doi:10.3758/s13423-016-1217-0

Stallings, L. M., & MacDonald, M. C. (2011). It's not just the "heavy NP": Relative phrase length modulates the production of heavy-NP shift. *Journal of Psycholinguistic Research*, *40*(3), 177–187. https://doi.org/10.1007/s10936-010-9163-x

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*(1), 71–86. https://doi.org/10.1016/j.cognition.2010.04.002

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Stuart, G. P., & Hulme, C. (2009). Lexical and semantic influences on immediate serial recall: A role for redintegration. *Interactions between Short-Term and Long-Term Memory in the Verbal Domain*, 157–76.

Szewczyk, J. M., Marecka, M., Chiat, S., & Wodniecka, Z. (2018). Nonword repetition depends on the frequency of sublexical representations at different grain sizes: Evidence from a multi-factorial analysis. *Cognition*, *179*, 23–36. https://doi.org/10.1016/j.cognition.2018.06.002

Szmalec, A., Duyck, W., Vandierendonck, A., Mata, A. B., & Page, M. P. A. (2009). The Hebb repetition effect as a laboratory analogue of novel word learning. *Quarterly Journal of Experimental Psychology (2006)*, *62*(3), 435–443. https://doi.org/10.1080/17470210802386375

Tanida, Y., Nakayama, M., & Saito, S. (2019). The interaction between temporal grouping and phonotactic chunking in short-term serial order memory for novel verbal sequences. *Memory*, *27*(4), 507–518. https://doi.org/10.1080/09658211.2018.1532008

Tanida, Y., Ueno, T., Lambon Ralph, M. A., & Saito, S. (2015). The roles of long-term phonotactic and lexical prosodic knowledge in phonological short-term memory. *Memory & Cognition*, *43*(3), 500–519. https://doi.org/10.3758/s13421-014-0482-2

Thorn, A. S. C., & Frankish, C. R. (2005). Long-Term Knowledge Effects on Serial Recall of Nonwords Are Not Exclusively Lexical. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(4), 729–735. https://doi.org/10.1037/0278-7393.31.4.729

Torreira, F., Bögels, S., & Levinson, S. C. (2015). Breathing for answering: The time course of response planning in conversation. *Language Sciences*, *6*, 284. https://doi.org/10.3389/fpsyg.2015.00284

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

Ueno, T., Saito, S., Saito, A., Tanida, Y., Patterson, K., & Lambon Ralph, M. A. (2014). Not Lost in Translation: Generalization of the Primary Systems Hypothesis to Japanese-specific Language Processes. *Journal of Cognitive Neuroscience*, *26*(2), 433–446. https://doi.org/10.1162/jocn_a_00467

Vallar, G., & Baddeley, A. D. (1984). Phonological short-term store, phonological processing and sentence comprehension: A neuropsychological case study. *Cognitive Neuropsychology*, *1*(2), 121–141. https://doi.org/10.1080/02643298408252018

Van de Cavey, J., & Hartsuiker, R. J. (2016). Is there a domain-general cognitive structuring system? Evidence from structural priming across music, math, action descriptions, and language. *Cognition*, *146*, 172–184. https://doi.org/10.1016/j.cognition.2015.09.013

Van Dyke, J. A., & Johns, C. L. (2012). Memory Interference as a Determinant of Language Comprehension. *Language and Linguistics Compass*, *6*(4), 193–211. https://doi.org/10.1002/lnc3.330

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1256–1271. https://doi.org/10.1037/0278-7393.25.5.1256

Warrington, E.K., & Shallice, T. (1969). *The selective impairment of auditory verbal short-term memory*. *92*, 885–896.

Warrington, E, K., Logue, V., & Pratt, R. T. C. (1971). The anatomical localisation of selective impairment of auditory verbal short-term memory. *Neuropsychologia*, *9*(4), 377–387. https://doi.org/10.1016/0028-3932(71)90002-9

Warrington, Elizabeth K., & Shallice, T. (1972). Neuropsychological Evidence of Visual Storage in Short-term Memory Tasks. *Quarterly Journal of Experimental Psychology*, *24*(1), 30–40. https://doi.org/10.1080/14640747208400265

Watkins, O. C., & Watkins, M. J. (1977). Serial recall and the modality effect: Effects of word frequency. *Journal of Experimental Psychology: Human Learning and Memory*, *3*(6), 712–718. https://doi.org/10.1037/0278-7393.3.6.712

Weiner, E. J., & Labov, W. (1983). Constraints on the Agentless Passive. *Journal of Linguistics*, *19*(1), 29–58.

Willits, J. A., Amato, M. S., & MacDonald, M. C. (2015). Language knowledge and event knowledge in language use. *Cognitive Psychology*, *78*, 1–27. https://doi.org/10.1016/j.cogpsych.2015.02.002

Willits, J. A., Seidenberg, M. S., & Saffran, J. R. (2014). Distributional structure in language: Contributions to noun-verb difficulty differences in infant word recognition. *Cognition*, *132*(3), 429–436. https://doi.org/10.1016/j.cognition.2014.05.004

Yue, Q., Martin, R. C., Hamilton, A. C., & Rose, N. S. (2019). Non-perceptual Regions in the Left Inferior Parietal Lobe Support Phonological Short-term Memory: Evidence for a Buffer Account? *Cerebral Cortex (New York, N.Y.: 1991)*, *29*(4), 1398–1413. https://doi.org/10.1093/cercor/bhy037

Yuzawa, M., & Saito, S. (2006). The role of prosody and long-term phonological knowledge in Japanese children's nonword repetition performance. *Cognitive Development*, *21*(2), 146–157. https://doi.org/10.1016/j.cogdev.2006.01.003

Zakariás, L., Kelly, H., Salis, C., & Code, C. (2018). The methodological quality of short-term/working memory treatments in post-stroke aphasia: a systematic review. *Aphasiology*, *32*, 251-254. https://doi.org/10.1080/02687038.2018.1490004 Zotero tab.