Working memory and second language acquisition:
Theory and findings

Daniel O. Jackson

Abstract

This article presents a narrative review of research on working memory (WM) in order to offer a foundation for understanding issues in second language acquisition (SLA). In the first section, current psychological views on WM are described in terms of three distinct theoretical models associated with Baddeley, Cowan, and Engle. Based on this, general issues for WM theories are summarized and practical implications for research are stated. In the second section, empirical findings from SLA research on WM in three domains are discussed. These domains include learning conditions, cognitive processes, and linguistic outcomes. The review concludes with some brief reflections on the future of WM research in SLA.

Introduction

How individuals process and store linguistic information, among other factors, shapes their learning of second and foreign languages. The construct of working memory (WM) has a decades-long history within cognitive psychology and a large number of empirical studies support its influence on second language acquisition (SLA). Several theoretical accounts, varying in detail, have been posited to explain WM. These accounts lend insight into the results obtained by second language (L2) researchers. Additionally, an understanding of the role of WM can inform future research on specific aspects of L2 processing and learning (e.g., the role of attention, age effects, and morphological acquisition). This narrative review, based on my dissertation research, describes theoretical perspectives on WM from psychology and empirical results from SLA to offer a foundation for understanding issues central to L2 learning.

Current psychological views on WM

This section introduces WM models, specifies general issues surrounding WM, describes changes in WM across the lifespan, discusses its neural underpinnings, and briefly explains WM measurement. SLA-related issues, discussed more fully in the next section, will only be briefly noted here.

Models of WM

Three models are widely discussed in the WM literature and in L2 research. These models have fueled a large body of empirical work and generated ongoing theoretical discussion. The three models reviewed here are depicted together in Figure 1 (based on Ricker, AuBuchon, & Cowan, 2010).

Baddeley's multicomponent model. Building on a series of experiments, Baddeley and Hitch (1974) were the first to elaborate on WM as a workspace of limited capacity to handle storage and manipulation of information. Their proposal extended the notion of short-term memory, which refers to storage only, in order to highlight the WM system's functional role (Baddeley, 2007). The specific details of the model have evolved over time, from the original three-component model to the addition of a fourth component (Baddeley, 2000).

Presently, the multicomponent model consists of a central executive assumed to carry out four processes: (a) focusing, (b) dividing, and (c) switching attention, as well as (d) linking WM and long-term memory. This executive component connects to three additional subcomponents. The phonological loop and the visuospatial sketchpad allow storage and rehearsal of information in STM that is phonological and articulatory, or visual and spatial, respectively. The latest component, the episodic buffer, serves to bind information from various sources into episodes. Recent accounts of the model retain this structure, while adding detail related to the information processed by each component (Baddeley, 2012; see also Baddeley & Logie, 1999).

Cowan's embedded-processes model. By comparison with the multicomponent model, one difference in Cowan's model of information processing (1988, 1995, 2005) is the emphasis placed on the intersection between memory and attention. In order to understand the mutual influence of these two areas on each other, this model distinguishes between two phases of memory activation: an initial, brief sensory store and a longer, short-term store, which is the activated subset of long-term memory, rather than a separate component. A central executive guides attention toward either external stimuli, or internal, long-term memories. Memory activation is limited in time, while attentional focus is limited in capacity to about four chunks in normal adults (Cowan, 2000).

Together, these features have been described as an embedded processes model, in which mechanisms of memory activation, attention, and long-term memory jointly comprise working memory (Cowan, 1999). Accordingly, this model emphasizes the interaction of various processes at a general level (Cowan, 2005).

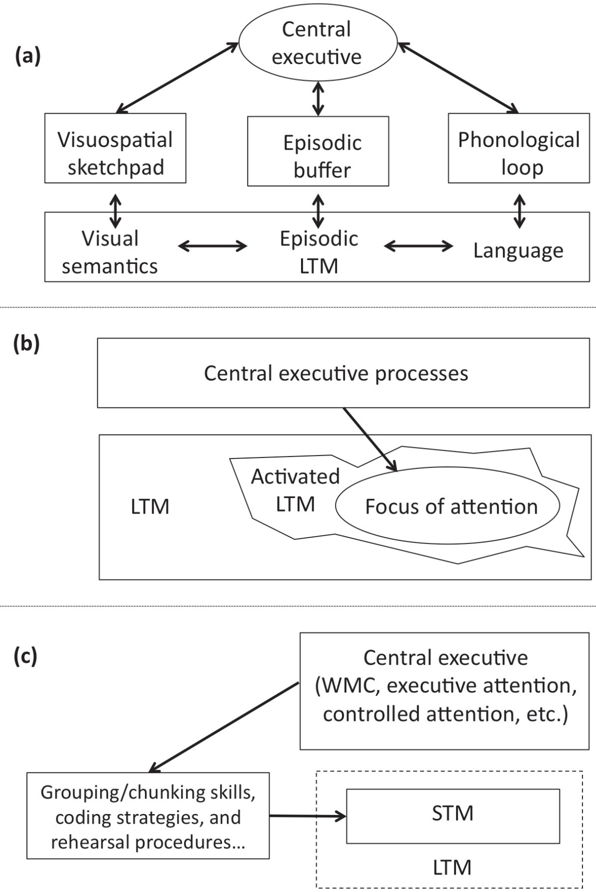

Figure 1. Three WM model visualizations, adapted from primary sources: (a) Baddeley's multicomponent model, (b) Cowan's embedded-processes model; and (c) Engle's model. LTM = long-term memory; STM = short-term memory; WMC = working memory capacity.

Engle and colleagues' resource-dependent inhibition model. Research by Engle and associates has contributed the theoretical insight that inhibition, or the ability to disregard irrelevant information during tasks, is crucial to working memory (Conway & Engle, 1994; Engle, Conway, Tuholski, & Shisler, 1995). In an overview of this approach, Engle, Kane, and Tuholski (1999) illustrated the connections between the central executive, long- and short-term memory, and procedures for maintaining activation, arguing that, "individual differences on measures of working memory capacity primarily reflect differences in capability for controlled processing" (p. 104). This model has also shed light on the relationships between WM capacity, general fluid intelligence, and executive attention (see Hambrick, Kane, & Engle, 2005). Given its emphasis on domain-general processes of maintenance and retrieval in the face of interference, this model has also been called, "an executive attention theory" of WM capacity (Kane, Conway, Hambrick, & Engle, 2007, p. 22; see also Engle, 2002).

Other perspectives on WM. Many other influential views on WM have been developed. For extensive reviews and comparison of multiple WM theories, see the volumes edited by Miyake and Shah (1999) and Conway, Jarrold, Kane, Miyake, and Towse (2008).

General issues

To summarize, WM is a system of temporary storage and attentional control. Several general issues arise when considering the main three models reviewed. These pertain to the non-unitary nature of WM, the role of the central executive, and the determination of capacity limits.

First, the three models align with slightly different perspectives on the division of WM into domain-general versus domain-specific components. Generally, researchers have tended to accept a non-unitary view of WM (Miyake & Shah, 1999). This is obviously true in the case of the multicomponent model, which from its early days has suggested roles for visual and auditory subsystems (Baddeley & Hitch, 1974). In contrast, Cowan (1999) takes a more unitary view, by arguing that, “different codes may be processed according to the same principles” (p. 79). The group led by Engle has concluded that WM capacity is primarily domain-general and secondarily domain-specific (Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004).

Second, related to the above, the central executive seems to personify WM in spite of differences in theoretical emphasis. Baddeley (2012) has recently elaborated on this component, by delineating attentional processes belonging to it (i.e., focusing, dividing, switching attention). From the beginning, Cowan’s model has included a central executive to direct attention and control processing (1988, 1995, 2005). In Engle’s work, attention takes center stage, particularly for the reason that differences in WM capacity are attributed to differences in the control of attention. Furthermore, each author gives precedence to the central executive by placing it *above* all other components in visual representations (see Fig. 1).

Third, in a seminar paper, Miller (1956) identified limits on memory and attention on the order of seven plus or minus two, which can be expanded by organizing input into chunks. WM models offer insight, but little consensus, on this matter. Baddeley (2007) believes that processing speed and storage impose separate limits, noting that the role of inhibition in determining limits appears unresolved. Cowan (2000, 2005) claims that the focus of attention within active memory is limited, summarizing multiple, alternative accounts. Engle considers attentional control important, attributing differences in WM capacity to maintenance and inhibition. Of the three, Baddeley seems the most sympathetic toward the notion of capacity limits in both storage and processing, while others appear to agree more with Craik and Lockhart’s (1990) suggestion, that “limited capacity turns out not to be an invariant structural property of a memory system, but a limitation imposed by processing” (p. 103).

Practical implications

There are many practical applications of WM theories for research on language learning, in general, and SLA, specifically. A few such implications are described here.

First, the WM capacity of children, young adults, and older adults varies. This is invaluable for understanding SLA, particularly considering Newport's (1990) 'less is more' hypothesis, which attributed variation in language learning ability across the lifespan to "differences between adults and children in the way linguistic input is perceived and stored" (p. 23). Newport pointed out that children's more limited STM capacity might give them an advantage over adults in learning morphology (see Hitch, 2006 and Park & Payer 2006 for reviews of age-related issues).

Second, a spate of research in the early 1990s demonstrated that the neural activity evoked by WM tasks is distributed throughout the brain. These basic findings can provide the foundation for a cognitive neuroscience of SLA, as illustrated by recent work investigating the connectivity of language pathways in the brain in relation to aptitude and WM (Xiang et al., 2012).

Third, researchers have developed a wide array of related measures in various knowledge domains (see Ackerman, Beier, & Boyle, 2005, for a review). These can be classified as measures of (a) *short-term memory* (e.g., the nonword repetition task, see Gathercole & Baddeley, 1990; Service, 1992), (b) *complex WM* (e.g., the reading span task, see Daneman & Carpenter, 1980; Kane et al., 2004) and (c) *dynamic WM* (e.g., the *n*-back task, see Conway et al., 2005; Jaeggi et al., 2010). In brief, short-term measures involve recall of stored material, whereas complex WM measures draw upon both recall and processing. Dynamic measures additionally require rapid matching and inhibition. To the extent that these processes are relevant to comprehending and producing language, each measure type is important for understanding how cognitive abilities relate to SLA.

WM and SLA: Conditions, processes, and outcomes

This section presents a selective overview of WM research within SLA. Rather than attempting a comprehensive review, I consider relationships between WM and three domains of SLA research. For additional reviews, see Ellis (2001), Juffs and Harrington (2011), Linck, Osthus, Koeth, and Bunting (2014), Robinson (2003), Wen, Borges Mota, and McNeill (2015), and Williams (2012).

Learning conditions

The role of WM in learning and performance under different conditions is an active area of research. Learning conditions may include input delivery, or how input is presented and processed, its modality (i.e., spoken versus written forms), and, modification, or changes to the shape of input based on simplification, elaboration, or recasts. Because teachers and sometimes learners may control these factors, it is important understand how they interact with WM abilities.

Input delivery. Regarding input delivery, Erlam (2005) reported a classroom study based on learning French in three treatment groups, each containing roughly 20 students: deductive instruction, inductive instruction, and structured input. WM ability, measured using a test of phonological capacity, correlated significantly with immediate and delayed tests of written production in the structured input group (r = .49 and .57, respectively). In the other groups, correlations between WM and outcomes were mostly positive, though not significant. Thus, it seems that production may be influenced to some degree by the learner-internal processes invoked by input delivery.

Controlled laboratory studies may also offer insight into this issue. For example, a laboratory study by Tagarelli, Borges Mota, and Rebuschat (2015) reported an artificial language learning experiment employing incidental and rule-search conditions. Two working memory tasks (letter-number ordering and operation span) were administered.

Both groups scored above chance, with the rule search group significantly outperforming the incidental group. Rule-search group participants who could state the rules showed a significant correlation of $r = 0.47$ between the grammaticality judgment task and the letter-number ordering task. No correlations, however, were found for the incidental group. Based on these two rather different (classroom versus laboratory) studies, WM may play a stronger role under certain input processing conditions than others, although more work on this topic is needed.

Input modality. Considering the aforementioned issue of domain-specificity, one may wonder whether associations between WM and L2 performance depend on testing these relationships in the same modality. This is an important question, to which the answer appears to be that relationships between WM and L2 performance in spoken and written tasks, interestingly, are modality-independent.

Regarding WM and L2 reading, an early study by Harrington and Sawyer (1992), showed that L2 reading span correlated significantly with L2 reading comprehension in a group of 34 Japanese learners of English (this study also revealed the language-independent nature of WM as L1 and L2 span measures were significantly correlated). In Geva and Ryan's (1993) investigation of L1 English-speaking children learning L2 Hebrew, a model fit to the data included Hebrew opposites span (HWM-O) and word span, both involving listening, among significant predictors of L2 reading, as measured by a cloze test.

Concerning WM and L2 listening, Robinson (2005) conducted a laboratory study on the incidental learning of Samoan by 37 Japanese college students in which WM ability was assessed using an L1 reading span task. WM significantly correlated with scores on a listening grammaticality judgment task (GJT), $r = .42$. Kormos and Sáfár's (2008) study of Hungarian EFL learners used two verbal WM tasks: nonword repetition and backward digit span task. Based on a sample of 45 beginning learners, the latter task showed a significant, moderate correlation with listening ability, $r = .37$.

Thus, significant relationships have been reported irrespective of the spoken versus written formats of the WM and L2 tests. This is not to deny that, as argued by Wen (2012), it may be preferable for researchers to employ measures relying on the same modality. Rather, these findings support the view that WM is, as argued above, at least partly domain-general (see Table 1).

Table 1. *Four Studies Reporting Significant Relationships between WM Capacity and L2 Abilities*

		L2 measure	
		Written	Spoken
WM measure	Written	Harrington & Sawyer (1992)	Robinson (2005)
	Spoken	Geva & Ryan (1993)	Kormos & Sáfár (2008)

Input modification. It has been suggested that WM may be among the variables related to learners' use of modified input. An important, preliminary investigation of this by Mackey, Philp, Egi, Fujii, and Tatsumi (2002) found a marginally significant relationship between a composite measure of WM capacity and noticing of recasts during a stimulated recall task. Also, there was stronger evidence of sustained language development in the high rather than low WM group, although this was based on a comparison of only seven learners.

Other studies suggest that WM may influence L2 outcomes when input is partially reformulated for the sake of elucidating form-meaning connections (e.g., Mackey, Adams, Stafford, & Winke, 2010; Mackey & Sachs, 2012). For instance, Goo (2012) reported that reading and operation span tests predicted learning from recasts, but did not predict learning from metalinguistic feedback, arguing that learning from metalinguistic feedback may demand less cognitive control than recasts. This line of research has also been extended recently to computer-delivered input modification (see, e.g., Sagarra & Abbuhl, 2013).

Cognitive processes

WM is closely linked to noticing, or the subjective awareness of surface features of the input (Schmidt, 1990, 1995, 2001, 2012). Schmidt has specified that, "noticing is related to rehearsal within working memory and the transfer of information to long-term memory, to intake, and to item learning" (1993, p. 213). Robinson has expounded on this relationship by characterizing noticing as detection plus rehearsal and awareness in WM (2003, see also 1995). Studies seeking to operationalize such factors are key to understanding the role of WM in basic SLA processes.

Noticing. Establishing empirical links between WM and noticing is no simple matter. One challenge lies with finding the appropriate methodological tools to observe noticing behaviors. Retrospective measures of noticing have been used for this purpose. The above cited paper by Mackey et al. (2002) utilized stimulated recall protocols to demonstrate that those with higher composite scores on nonword recall, L1 listening span, and L2 listening span displayed more instances of noticing related to target forms during videotaped interactions.

Understanding. Awareness can give rise to metalinguistic knowledge and WM may also be important to such understanding. Here, the evidence suggests that any such relationship is sensitive to the measures adopted. An early study by Ellis and Sinclair (1996) tested explicit rule learning of novel Welsh constructions, finding that participants who repeated input aloud during the learning phase outperformed those whose WM function was suppressed by having them recite an unrelated number sequence. Thus, the quality of input processing in WM may also be crucial when learning is based on awareness at the level of understanding.

However, three other studies point to the conclusion that WM and metalinguistic awareness may be not be closely related. Robinson (2002), finding no association between a reading span task and an unspeeded GJT, noted that these measures do not exhibit process continuity (p. 256). In another study, Bell (2009) showed that inductive language learning ability, but not reading span, predicted metalinguistic awareness of French gender rules. Roehr and Gánem-Gutiérrez (2009) also found no significant relationship between span tests and untimed measures of metalinguistic knowledge. They also suggested this result was due to the incompatibility of measures.

Aptitude. The idea that WM is closely related to language aptitude has received support from cognitive psychologists (Miyaki & Friedman, 1998). Yet, there is quite a range of positions on their precise relationship. Miyake and Friedman's (1998) WM-as-language-aptitude proposal suggests that WM may be the primary construct behind aptitude for language learning. Others hypothesize that WM plays a recurring role in various aptitude constructs corresponding to stages of L2 processing such as segmentation, noticing, and pattern identification (Dörnyei & Skehan, 2003). Still others see WM as one of a number of cognitive resources underlying aptitude complexes which apply in a given task and setting, such as learning from recasts (Robinson, 2007). Thus, as the traditional notion of aptitude derived from the use of tests such as the MLAT—which included phonetic coding ability, grammatical sensitivity, memory abilities, and inductive language learning ability (Carroll, 1990)—gives way to newer aptitude theories, the status of WM within them varies considerably.

Two recent studies embody these theoretical tensions. Sáfár and Kormos (2008) established significant, moderate correlations between the Hungarian version of the

MLAT and a backward digit span task in a study of IDs among EFL learners. Based on a regression analysis including both as independent measures, they argued, “working memory is a better predictor of language learning success than the traditional construct of language aptitude” (p. 129). Building on Carroll’s work, Doughty et al. (2010) conducted a study on the reliability of the High-level Language Aptitude Battery (Hi-Lab). In a factor analysis of the test battery, WM was one of several factors, which also included tolerance of ambiguity, task switching, perceptual acuity, rote memory, and speech perception in noise. As such, WM seems unlikely to supersede the construct of aptitude entirely, though it stands out as a significant component in newer conceptualizations of language aptitude (see also Granena & Long, 2013).

Linguistic outcomes

As already implied, the potential role of working memory or short-term memory has been investigated in relation to numerous L2 outcomes. In this section, I focus on the specific outcomes addressed in Jackson (2014; under review). Given the artificial L2 constructions used in that study, this includes studies in which participants learned word meanings and bound morphology with no prior knowledge of the specific target forms.

Word meanings. Several studies have offered empirical evidence linking WM and the acquisition of word meanings. For example, Atkins and Baddeley (1998) used a training method in which learning trials presenting 32 adult participants with English-Finnish sentence pairs were interspersed with test trials requiring them to translate from English to Finnish. Verbal WM span (but not visuo-spatial span) was significantly related to their rate of learning, in that, as span scores rose, the number of test trials until participants provided correct translations fell, r = -.52. Next, Speciale, Ellis, and Bywater (2004) reported two studies on the acquisition of foreign language vocabulary by undergraduates. In these studies, correlational analyses showed that nonword repetition was significantly related to oral production of English-to-German translation equivalents (Study 1, *N* = 38) and to lexical comprehension on a Spanish course examination (Study 2, *N* = 44). Also, a recent laboratory training study by Martin and Ellis (2012) reported correlations between a composite measure of artificial vocabulary learning and three independent measures: nonword repetition, nonword recognition, and listening span, all of which showed significant correlations (ranging from *r* = .34 to .40), based on 40 participants. In sum, WM capacity clearly supports the ability to translate to and from a new lexicon (for reviews, see Baddeley, Gathercole, Papagno, 1998; Ellis, 1996).

Morphology. Studies have also explored the role of WM abilities in the development of morphological knowledge. Williams and Lovatt (2003) targeted the role of phonological short-term memory (PSTM) in acquiring determiner-noun agreement rules in novel input. In the first experiment, involving Italian stimuli, no significant relationship between nonword repetition and generalization was found based on a partial correlation controlling for language background. However, in their second experiment, Williams and Lovatt used artificial stimuli to reduce the influence of language background with 21 participants. This time, the mediating influence of language background on PSTM disappeared, and PSTM and language background independently contributed to generalization (see Williams, 2012 for discussion). In the last test cycle, this correlation reached *r* = .53, which was significant. Memory for morpheme combinations (among trained items) was also a significant predictor of generalization (*r* = .64). Here, free morphemes encoded grammatical gender, definiteness, and number.

Brooks and Kempe have conducted a series of investigations to assess the contribution of WM to the learning of bound morphology in Russian, particularly nominal inflections marking gender and case. An overview of their work suggests that it is not only WM, but measures of intelligence, similarly involving control of attention, such as Cattell's Culture Fair Intelligence Test, that predict outcomes. For instance, they have argued that “a capacity for memorizing unfamiliar items along with an ability to effectively allocate attention to the analysis of distributional characteristics of the input is crucial for learning complex inflectional paradigms” (Kempe & Brooks, 2008, p. 742). These researchers have demonstrated how learning is shaped by the interaction of these cognitive factors and (a) type variation in the input (Brooks, Kempe, & Sionev, 2006), (b) linguistic characteristics of the input (Kempe, Brooks, & Kharkhurin, 2010), and (c) metalinguistic awareness emerging from training (Brooks & Kempe, 2013).

Specifically regarding WM, Brooks, Kempe and Sionev (2006) reported significant relationships between reading span and learning of a case marking paradigm on old items (r = .69) and on new items (r = .53) with 20 participants, when the target vocabulary size was 24 tokens, but not in conditions where vocabulary size was only 6 or 12 items. To explain their findings, the researchers turned to the critical mass hypothesis, which assumes that morphosyntactic development is driven by vocabulary size (Marchman & Bates, 1994). Martin and Ellis’ (2012) study also demonstrated that while vocabulary accounted for the most variance in knowledge of inflectional morphology, WM made a further independent contribution. Hence, among researchers who place lexical and grammatical knowledge along a continuum, it appears that not only vocabulary growth, but also WM, may impact the learning of inflectional morphology.

The future of WM research in SLA

This paper has described theoretical perspectives on WM and considered its influence on SLA, in terms of conditions, processes, and outcomes. In closing, language teachers and researchers may find it valuable to reflect on several developing research trends in WM. First, it is important to remember that WM is not only complex, but interacts with other individual differences in complex ways. Thus, studies adopting an aptitude complex approach (Robinson, 2007) to explore how WM and other cognitive abilities (e.g., analogical reasoning) facilitate SLA under specific conditions can lead to deeper understanding. Second, closely related to this, aptitude-treatment interaction research (Vatz, Tare, Jackson, & Doughty, 2013) assumes that instruction can be designed to match learner abilities, such as WM. The results of such research may eventually help to identify more psycholinguistically valid approaches to language teaching. A final emerging line of research investigates the role of WM training in improved language processing. For instance, Novick and colleagues (2014) have achieved promising results with L1 English participants who, after training on complex and dynamic WM tasks, showed gains in resolving syntactic ambiguities in garden-path sentences (e.g., *While the thief hid the jewelry that was elegant and expensive sparkled brightly*). Theoretical and practical advances are likely to follow from future studies based on these developments. Such empirical studies are the focus of a volume on cognitive individual differences being edited by Granena, Jackson, and Yilmaz (in preparation).

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30–60.

Atkins, P. W. B., & Baddeley, A. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19*, 537–552.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.

Baddeley, A. D. (2007). *Working memory, thought, and action*. Oxford, UK: Oxford University Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation, Vol. 8* (pp. 47–89). New York, NY: Academic Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). New York, NY: Cambridge University Press.

Bell, P. (2009). *Le cadeau* or *la cadeau*? The role of aptitude in learner awareness of gender distinctions in French. *The Canadian Modern Language Review / La revue canadienne des langues vivantes, 65*, 615–643.

Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: the mediating effect of metalinguistic awareness. *Memory & Cognition, 41*, 281–296.

Brooks, P. J., Kempe, V., & Sionev, A. (2006). The role of learner and input variables in learning inflectional morphology. *Applied Psycholinguistics, 27*, 185–209.

Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. S. Parry & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 11–29). Englewood Cliffs, NJ: Prentice Hall.

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*, 354–373.

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. (Eds.). (2008). *Variation in working memory*. New York, NY: Oxford University Press.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engel,

R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769–786.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*, 163–191.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York, NY: University Press.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). New York, NY: Cambridge University Press.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–185.

Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, UK: Psychology Press.

Craik, F. I. M., & Lockhart, R. S. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology, 44*, 87–112.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Malden, MA: Blackwell.

Doughty, C. J., Campbell, S. G., Mislevy, M. A., Bunting, M. F., Bowles, A. R., & Koeth, J. (2010). Predicting near-native ability: The factor structure and reliability of Hi-LAB. In M. T. Prior, Y. Watanabe & S.-K. Lee (Eds.), *Selected proceedings of the 2008 Second Language Research Forum* (pp. 10–31). Somerville, MA:

Cascadilla Proceedings Project.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18*, 91–126.

Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33–68). Cambridge, UK: Cambridge University Press.

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology, 49*, 234–250.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19–23.

Engle, R. W., Conway, A. R. A., Tuholski, S. W., & Shisler, R. J. (1995). A resource account of inhibition. *Psychological Science, 6*, 122–125.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). New York, NY: Cambridge University Press.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research, 9*, 147–171.

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology, 81*, 439–454.

Geva, E., & Ryan, E. B. (1993). Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning, 43*, 5–42.

Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition, 34*, 445–474.

Granena, G., & Long, M. (Eds.). (2013). *Senitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam, The Netherlands/Philadelphia, PA: John Benjamins.

Granena, G., Jackson, D. O., & Yilmaz, Y. (Eds.). (in preparation). *Cognitive individual differences in L2 processing and acquisition.*

Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2005). The role of working memory in higher-level cognition: Domain-specific versus domain-general perspectives. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence: Identifying the mechanisms of the mind* (pp. 104–121). Cambridge, UK: Cambridge University Press.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*, 25–38.

Hitch, G. J. (2006). Working memory in children: A cognitive approach. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 112–127). New York, NY: Oxford University Press.

Jackson, D. O. (2014). *The relative contribution of input modification, learner awareness, and individual differences to second language construction learning (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3582920).* Unpublished doctoral dissertation.

Jackson, D. O. (under review). Variation sets, explicit knowledge, and individual differences in second language construction learning.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory, 18*, 394–412.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching, 44*, 137–166.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold & M. J. Kane (Eds.), *Variation in working memory* (pp. 21–48). Cary, NC: Oxford University Press.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189–217.

Kempe, V., & Brooks, P. J. (2008). Second language learning of complex inflectional systems. *Language Learning, 58*, 703–746.

Kempe, V., Brooks, P. J., & Kharkhurin, A. (2010). Cognitive predictors of generalization of Russian grammatical gender categories. *Language Learning, 60*, 127–153.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11*, 261–271.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*, 861–883.

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning, 60*, 501–533.

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed second language learning* (pp. 181–209). Amsterdam, The Netherlands/Philadelphia, PA: John Benjamins.

Mackey, A., & Sachs, R. (2012). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning, 62*, 704–740.

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*, 339–366.

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34*, 379–413.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 101*, 343–352.

Miyake, A., & Friedman, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Erlbaum.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY: Cambridge University Press.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11–28.

Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, I., & Bunting, M. F. (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language, Cognition and Neuroscience, 29*, 186–217.

Park, D. C., & Payer, D. (2006). Working memory in children: A cognitive approach. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 128–142). New York, NY: Oxford University Press.

Ricker, T. J., AuBuchon, A. M., & Cowan, N. (2010). Working memory. *WIREs Cognitive Science, 1*, 573–585.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*, 283–331.

Robinson, P. (2002). Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfield and Hernstadt (1991). In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 211–266). Amsterdam, The Netherlands/Philadelphia, PA: John Benjamins.

Robinson, P. (2003). Attention and memory during SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631–678). Malden, MA: Blackwell.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition, 27*, 235–268.

Robinson, P. (2007). Aptitudes, abilities, contexts, and practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and psychology* (pp. 256–286). Cambridge, UK: Cambridge University Press.

Roehr, K., & Gánem-Gutiérrez, G. A. (2009). The status of metalinguistic knowledge in instructed adult L2 learning. *Language Awareness, 18*, 165–181.

Sáfár, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *IRAL, 46*, 113–136.

Sagarra, N., & Abbuhl, R. (2013). Optimizing the noticing of recasts via computer-delivered feedback: Evidence that oral input enhancement and working memory help second language learning. *The Modern Language Journal, 97*, 196–216.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York, NY: Cambridge University Press.

Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. In W. M. Chan, K. N. Chin, S. Bhatt & I. Walker (Eds.), *Perspectives on individual characteristics and foreign language education* (pp. 27–50). Boston, MA: De Gruyter.

Schmidt, R. (Ed.). (1995). *Attention and awareness in foreign language learning*. Honolulu, HI: University of Hawai'i Press.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology, 45*, 21–50.

Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics, 25*, 293–321

Tagarelli, K. M., Borges Mota, M., & Rebuschat, P. (2015). Working memory, learning conditions, and the acquisition of L2 syntax. In Wen, Z., Borges Mota, M., & McNeill, A. (Eds.). (2015). *Working memory in second language acquisition and processing* (pp. 224–247). Bristol, UK: Multilingual Matters.

Vatz, K., Tare, M., Jackson, S. R., & Doughty, C. J. (2013). Aptitude-treatement interaction studies in second language acquisition: Findings and methodology. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 273–292). Amsterdam, The Netherlands/Philadelphia, PA: John Benjamins.

Wen, Z. (2012). Working memory and second language learning. *International Journal of Applied Linguistics, 22*, 1–22.

Wen, Z., Borges Mota, M., & McNeill, A. (Eds.). (2015). *Working memory in second language acquisition and processing*. Bristol, UK: Multilingual Matters.

Williams, J. N. (2012). Working memory and SLA. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 427–441). New York, NY: Routledge.

Williams, J. N., & Lovatt, P. J. (2003). Phonological Memory and Rule Learning. *Language Learning, 53*, 167–121.

Xiang, H., Dediu, D., Roberts, L., van Oort, E., Norris, D. G., & Hagoort, P. (2012). The structural connectivity underpinning language aptitude, working memory, and IQ in the perisylvian language network. *Language Learning, 62(S2)*, 110–130.