Background

The term “working memory” (WM) is generally believed to have made its debut in a study by Miller et al. (1960), and it was interpreted as the cognitive capacity that allows us to make plans and keep goals in mind. Since then, enthusiasm over the WM concept has been boosted significantly by the seminal model of Baddeley and Hitch (1974), which gradually developed into the classic, canonical model widely used in major fields of psychology and cognitive science. The popularity of the notion of working memory is partly attributable to the much older synonym of short-term memory (STM), though it is distinguished from it with an additional “working” (“processing” or “attention”) component. WM can be defined as the cognitive capacity that allows us to maintain a limited amount of information in our mind (e.g., be it linguistic or visual), temporarily available for use in thought or action (cf. Cowan, 2017; Oberauer et al., 2018).

A major focus of more modern conceptions of WM has been its limited capacity (Carruthers, 2013). WM limitations, or the consequences brought by them, are tangible and ubiquitous, permeating most essential facets of our cognitive life (e.g., in the guise of an overflowing brain; Klingberg, 2008). The constraint effects brought by WM limitations are normally manifested in two ways (Wen & Li, 2019). On the one hand, cognitive psychologists have long conjectured that our normal capacity to maintain information in focal attention for later recall is estimated to be just around “four to seven” single units or chunks (Miller, 1956; Cowan, 2001). On the other hand, information that can indeed be held in our mind available for use or further manipulation is rather short-lived, normally lasting no more than “12 to 20” seconds unless rehearsed repeatedly (Waugh & Norman, 1965). This is akin to the phenomenon of the “now-or-never” bottleneck effects (cf. Christiansen & Chater, 2016).

Despite its limited capacity, WM (or STM to that effect) has long been held as the primary memory (system) as early as William James’ time (1890). It is assumed that WM may play a larger role than long-term memory (LTM) in subserving human cognition and action (e.g., Baddeley et al., 1988; Miyake & Shah, 1999; also see Lieder & Griffiths, 2020). Research endeavors from all major fields and subfields of cognitive science (e.g., Miller, 2003; also see Roberts & Kreuz, 2015) are contributing a clearer and deeper understanding of the nature, structure, and functions of WM. They have also led to the propagation of a dozen or so theoretical models of WM as a result of epistemological stances and perspectives (Miyake & Shah, 1999; Baddeley, 2012; Cowan, 2017; Logie et al., 2021b; Schwieter & Wen, 2022). Notwithstanding, this proliferation of diverse theories and models can sometimes backfire and create confusion among scholars at the consumer end of
application (cf. Logie et al., 2021a). This situation can become even worse when scholars who need to rely on WM theories and assessment techniques are not so well versed in the real distinctions among these multiple perspectives and models, thus requiring scholars working in this domain to integrate theories of WM (Logie et al., 2021b).

Having long been baffled by deep-rooted controversies and debates over WM theories and models (Baddeley, 2012; Cowan, 2017; Logie et al., 2021a), not to mention the dazzling number of WM span tasks and assessment procedures from psychology and neuroscience (Conway et al., 2005), many scholars at the applied end of WM have repeatedly called in recent years to re-conceptualize the construct to bring it to bear on more specific domains of human cognition and action. Some newly formulated frameworks or “micro” models of WM have sprouted out in recent years in more specific applied domains in such areas as general academic learning and intervention (e.g., Dehn, 2008), educational research on multimedia learning (Fenesi et al., 2015), and classroom-based instructional design of tasks in tandem with the cognitive load theory (Sepp et al., 2019), among many others (Logie et al., 2021b). Following these recent trends, Wen (2016) has also proposed an integrated framework of WM set out to predict and explain individual differences in the process and product of second language acquisition (SLA).

As shown in Figure 3.1, the integrated framework of WM and SLA (Wen, 2016) is constructed based on the unified understanding of the WM concept together with its putative components and functions in relation to their potential effects on specific aspects of L2 acquisition and processing. In essence, the framework is premised on four interrelated theoretical assumptions of the WM construct being conceptualized either as 1) a single pool of cognitive resources with limited capacity subserving human thought and action (Oberauer et al., 2018; Lieder & Griffiths, 2020); 2) an interactive gateway to long-term memory; 3) a memory (sub)system comprising multiple components (following Baddeley’s 2012 structural model; also see Baddeley, 2022) a human cognitive system consisting of embedded executive processes (following Cowan’s 1999 model; also see Cowan, 2022) or attentional control functions (following Engle & Kane’s 2004 model).

More importantly, the corollary of the integrated framework lies in its hierarchical nature that can be operationalized from four distinctive albeit interconnected levels of analysis (cf. Lieder & Griffiths, 2020). First of all, WM can be operationalized from the perspective of LTM knowledge and contents, which refers to the activated section of focused attention within the broader store of L1 competence and L2 knowledge (Cowan, 1999). Second, WM can also be operationalized as

---

**Figure 3.1** An Integrated Conceptual Framework for WM in SLA Research (Wen, 2016).
Zhisheng (Edward) Wen and Daniel O. Jackson

consisting of multiple components, in which the phonological buffer (aka, the phonological loop or phonological working memory, or phonological memory for short) and the central executive (executive WM, or “executive attention”; cf. Bialystok, 2018) are the two key components in language learning and processing (cf., Baddeley, 2003, 2015; Baddeley et al., 1998; Gathercole & Baddeley, 1993). Third, WM can also be considered from the viewpoint of finer-grained mechanisms and functions, including the phonological store and rehearsal mechanism (as subsumed by phonological working memory), and the executive functions of information updating, task switching, and inhibitory control (cf. the “unity and diversity” model of executive functions by Miyake & Friedman, 2012). Finally, the field of cognitive psychology and neuroscience also offers a wide array of memory span tasks and assessment procedures to measure the construct of WM (Conway et al., 2005). Among these, two major categories of memory span tasks can be identified: simple (storage-only) memory span tasks such as the digit span, the letter span, and the nonword span tasks, and complex (storage-plus-processing) memory span tasks such as the reading span and operation span tasks.

The aforementioned integrated framework of WM for SLA and categories of WM measures can offer the necessary background to understand existing research or to design new studies probing WM effects on L2 learning and processing. Alternatively, these four levels of analysis can provide guidelines and criteria to triangulate the results and findings of existing empirical studies of WM in SLA research, with a view toward unveiling the emerging patterns regarding putative links posited for the “WM-SLA nexus” (Wen, 2012, 2016).

Research Evidence

In this section, we review the research on working memory by reporting its overall effects, the differential effects of phonological working memory (PWM) and executive working memory (EWM), and the interaction of WM with treatment effects.

Overall Effects

The number of published studies conducted by SLA scholars probing WM effects has grown considerably over the last three decades, particularly after the publication of the seminal study by Harrington and Sawyer (1992). For example, a recent book-length survey (Wen, 2016) manually tabulated and retrieved 80 empirical studies published between 1992 and 2015 from major international journals (including PhD theses). Chen et al. (2016) found over 81 WM-SLA studies published in key linguistics journals in China alone. In the latest position paper by Jackson (2020), the author was able to retrieve over 45 empirical studies from the ERIC database just in the three-year interval between 2016 and 2019, a period that we will focus on in this chapter (given the current chapter’s limited space and scope), though we will extend slightly by incorporating studies in 2019 and 2020 as well. With the number of empirical studies probing WM effects in SLA growing at such an exponential rate, it has become a daunting, if not impossible, task to manually keep track of them all. Fortunately, the SLA research field has benefited substantially in the last two decades or so by the introduction of systematic methods of research synthesis and meta-analysis (Norris & Ortega, 2000). Therefore, we discuss the empirical results from both book-length surveys and meta-analytic studies.

Regarding meta-analytic studies of WM and SLA (or bilingualism to that effect), several studies can be retrieved (from search engines such as Google Scholar and the ERIC database). Watanabe and Bergsleithner’s (2006) initial synthesis of WM measures was followed by a major and more widely cited study by Linck et al. (2014) targeting the relationship between WM and SLA. The mean correlations between WM and learning outcomes found by major meta-analytic studies are
Working Memory

Table 3.1 Summary of Results from Recent Meta-Analytic Studies of WM and SLA

<table>
<thead>
<tr>
<th>Studies</th>
<th>No. of Studies</th>
<th>Overall</th>
<th>PWM</th>
<th>EWM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watanabe &amp; Bergsleithner (2006)</td>
<td>16</td>
<td>r = 0.18</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Linck et al. (2014)</td>
<td>79</td>
<td>r = 0.26</td>
<td>0.18</td>
<td>0.27</td>
</tr>
<tr>
<td>Li (2017b)</td>
<td>24</td>
<td>r = 0.23</td>
<td>0.22</td>
<td>0.28</td>
</tr>
</tbody>
</table>

displayed in Table 3.1. These studies showed that WM is a significant but weak predictor of L2 learning, with mean correlations ranging from 0.18 to 0.28. Taken together, these meta-analytical studies have provided a broad picture for confirming the overall main effects of WM, as a whole, on SLA. But, to get a clearer picture of some of these relationships, we still need to summarize the more specific results and emerging patterns regarding the main effects of the two types of working memory examined in the research, namely PWM and EWM.

**Phonological Working Memory (PWM) and Executive Working Memory (EWM)**

As indicated in Table 3.1, PWM effects, as indexed by the “simple” versions of memory span tasks such as the nonword recognition or repetition span task, are generally weaker than EWM effects (e.g., \( r = 0.18 \) for PWM and 0.27 for EWM, respectively, as reported by Linck et al., 2014, p. 870). Furthermore, when major effects are indeed detected, they are generally more related to the developmental (as opposed to real-time, online processing, or performance) domains of SLA, including L2 phonology and speech, L2 lexical or vocabulary development, L2 formulaic chunks, L2 grammatical development, and L2 oral abilities development (see Wen, 2015, 2016 for a detailed review). Finally, most of these studies seem to suggest that PWM effects tend to be more prominent among children or younger L2 learners, especially when they are at early or lower stages of L2 proficiency (e.g., Cheung, 1996; also see Serafini & Sanz, 2016 who reported this effect among adult learners).

In contrast, EWM measured by means of complex working memory tasks, such as the reading span task and the operation span task, has been found to have stronger predictive power than PWM in most cases of L2 learning and processing. These results seem to be related to the attentional and executive control (including information updating, task switching, and inhibitory control) aspects of WM, subserving such resource-demanding cognitive processes as the noticing of L2 grammatical cues, corrective feedback in L2 interactions (to be elaborated on in the next section), as well as some (online) performance areas, such as L2 listening performance, transfer skills in L2 reading, L2 speech performance (under certain conditions, such as pressured within-task or online planning) (see studies in L2 task literature; also see Ellis et al., 2020), and bilingual interpreting (also see Wen & Dong, 2019 for results of a meta-analytical study of 14 empirical studies). EWM effects (expressed as a composite Z-score!) were also detected among selective L2 written performance areas, such as an academic editing task (Michel et al., 2019; Revesz et al., 2019). Furthermore, similar to PWM, overall EWM effects are more obvious during the early stages of L2 learning, and their effects become less stable and more variable as participants’ L2 proficiency develops (Serafini & Sanz, 2016).

There is insufficient research that aims to investigate PWM and EWM effects in the longer term; however, several notable exceptions exist. A study by Linck and Weiss (2011) found that EWM (measured by the operation span task) predicted gains in L2 proficiency during a semester of instruction, after controlling for other ID variables. Also, Serafini and Sanz (2016) reported that both PWM and EWM effects seem to diminish gradually as participants’ L2 proficiency grows. This study was conducted over one academic term, immediately before and after instruction, as well as four weeks after instruction. Finally, the two longitudinal studies reported in Sagarra (2017).
highlight the key issue of selecting appropriate measures of WM. In Experiment 1 no effect was found using a less taxing WM measure, whereas Experiment 2 replicated this result, also finding an effect on L2 proficiency using a more taxing WM measure. It should be noted that these two studies also differed in length (one year and 11 weeks, respectively). Additional longitudinal studies are needed to fill gaps in our understanding of which WM components best predict L2 development.

Interaction with Instruction Type

An issue that is not covered adequately by the meta-analytic studies or book-length surveys is the distinction between main effects and interaction effects with regard to WM, that is, between the predictive studies and those investigating the possibility of differential WM effects under different instructional treatments (Goo, 2012, 2016; Li, 2017a, 2017b; also see Wen & Li, 2019). Partly this is because there has not been much work probing the interaction effects of WM until recently (Li, 2017a; Ruiz et al., 2019). Li (2017b) conducted a small-scale meta-analytic study (involving 24 studies) targeting the associations between WM and language aptitude, the effects of noticing on corrective feedback, and modified output after feedback. The meta-analysis revealed significant, albeit weak, correlations between all measures of WM and the immediate effects of L2 corrective feedback: \( r = 0.23 \); the mean correlations were 0.22 for PWM and 0.28 for EWM. The correlations with the delayed effects of feedback were near zero. For noticing and modified output, there was insufficient research for a meta-analysis, and a narrative review showed inconsistent associations between working memory and L2 learners’ noticing of corrective feedback. For modified output, a general pattern seemed to be that working memory was only associated with learners’ responses after feedback that pushed learners to produce output, not feedback that provided the correct form such as recasts. Compared with working memory, language aptitude was found to have a stronger effect on corrective feedback in terms of both immediate and delayed effects. Finally, it was found that both WM and language aptitude seemed to play a more important role in explicit rather than implicit treatments, pointing to the need for more research in implicit learning abilities.

We now turn to the results of several recent empirical studies that have explored the interactions between WM and L2 instruction that have been published since Li’s meta-analytical study (2017b). Four such studies are discussed and summarized below. The first study is by Indrarathne and Kormos (2018), who explored the interaction between WM and input enhancement. The participants were 80 L2 English learners exposed to causative constructions in reading passages under four experimental conditions (enhanced input plus instructions, enhanced input plus instructions and an explanation, enhanced input only, or unenhanced input). Results indicated that WM (as indexed by a composite score based on four memory span tasks, i.e., the forward digit-span test, the keep track task, the plus–minus task, and the Stroop task) was associated with learning in all four conditions, with a significantly stronger association in the more explicit conditions.

Malone (2018) also investigated the interactions between WM and L2 instruction, with respect to L2 vocabulary learning. Intermediate L2 English learners were administered memory span tasks tapping PWM (i.e., the nonword repetition span task) and EWM (i.e., the operation span task and shapebuilder memory task), scores on which were then combined into a composite WM capacity score. Students were all asked to read stories containing low-frequency words. One group received aural support while reading and the other group did not receive aural support. Both groups were later assessed on their knowledge of word form and word meaning. The results indicated that WM capacity (i.e., the composite score) was predictive of word form recognition, but only for the group that received aural support.

In another recent study by Ruiz et al. (2019), the authors investigated the relationship between WM and LTM (i.e., declarative memory) and different types of vocabulary instruction. In the context of web-based intelligent computer-assisted language learning (ICALL), 127 advanced German native learners of L2 English were assigned to two treatment conditions: form-focused and mean-
Working Memory

ing-focused instruction. Mixed-effects regression analyses indicated that WM was associated with L2 vocabulary learning only in the form-focused condition. Declarative memory was not associated with either instructional condition.

Li et al. (2019) examined the effects of WM and language analytic ability (LAA) on L2 narrative task performance under five different instructional conditions by 150 young EFL learners who underwent a two-hour treatment session. The five groups were based on whether and when they received form-focused instruction: 1) pretask instruction on the linguistic target (English passive voice) before the performance of tasks; 2) within-task feedback but no pretask instruction; 3) both pretask instruction and within-task feedback; 4) post-task feedback; and 5) a control group who just performed the tasks. Statistical analyses indicated that WM was associated with learning outcomes in the two groups receiving within-task feedback, but not in the group receiving pretask instruction or post-task feedback, suggesting that WM effects may emerge when the learning task imposes a heavy processing burden. The results also showed that LAA was predictive of the posttest scores of the control group and the group that received post-task feedback, indicating its effects may weigh in when external assistance is not provided.

Data Elicitation

The methods used in the WM studies described above were based on well-established approaches to the study of individual differences. Such studies typically feature within-participant designs wherein each learner contributes scores on the WM and L2 measure, and these respective scores for the predictor and criterion (or outcome) variables are compared using correlational techniques, including regression analyses. In the case of interaction effects, WM would be regarded as a moderator that influences the strength of the relationship between an independent variable (e.g., an instructional treatment) and a given learning outcome.

Over the years, as the construct of WM has evolved (Nee & D’Esposito, 2018), its assessment measures have evolved as well, making way for a wide range of memory span tasks and scoring procedures (Conway et al., 2005). Thus far, SLA studies have mostly relied on the well-established research paradigms of WM span tasks and procedures from psychology and cognitive science (Wen et al., 2021). On the one hand, the simple (i.e., storage-focused) versions of WM span tasks have been widely adopted by SLA scholars in favor of the British/European tradition (e.g., led by Baddeley and Gathercole) as proxies for PWM. This group of measures includes the digit span, the letter span, and most representative of all, the nonword recognition or repetition span task (Gathercole, 2006). On the other hand, many more SLA researchers subscribe to North American functional views of WM in opting to implement the complex versions of the memory span tasks that purportedly tap into the dual storage-plus-processing functions of WM (i.e., EWM). A classic representative in this category is the reading span task devised by Daneman and Carpenter (1980). As time goes on, recent years have witnessed more and more SLA researchers adopting the domain-general, language-neutral operation span task (Turner & Engle, 1989; Unsworth et al., 2005). These tasks arguably help to mitigate against the confounding of language and cognitive ability, which is inherently an issue when using WM measures that rely heavily on the use of either the L1 or L2 (Wen et al., 2021). Even more recently, studies have expanded to include dynamic measures of WM, requiring finer-grained subprocesses of monitoring and updating as participants respond to a subset of typically visual or auditory stimuli (Conway et al., 2005), such as the N-back task (see Jackson, 2016, for a study using this measure alongside the reading span task). Many new formats of WM tasks are being launched and put to use, as indicated in some of the studies discussed above.

As pointed out succinctly, yet circularly, by Dehn (2008, p. 58), “[u]ntil research and measurement tools allow us to further delineate WM processes, it might be safest to define WM as what simple and complex WM span tasks measure”. This highlights the ongoing nature of WM con-
struct and test development in the cognitive sciences, yet the resultant broad range of measures can sometimes backfire and create confusion among secondary researchers and meta-analysts. To further improve the situation, several tips have recently been proposed regarding measuring WM (Wen et al., 2021). These include, for example, the adoption of a hierarchical order of linguistic and nonlinguistic factors potentially impacting upon the outcomes of WM span task performance (following Cai & Dong, 2012), in order of their weights, the information types (verbal vs. nonverbal), the encoding modalities (listening vs. reading span task vs. speaking span task, etc.), and the encoding languages (i.e., whether the test is conducted in participants’ L1 or L2).

Once the aforementioned issue is addressed, namely that of choosing WM tasks, the data elicited then needs to be scored and analyzed appropriately. Given their prevalence, span tasks are an appropriate example. In a reading span task, learners may be asked to read a sentence, remember its final word, and judge whether the sentence is plausible. This task would generate data regarding the number of items (i.e., words) recalled, the accuracy of the plausibility judgments, and the response time (Waters & Caplan, 1996). Sentences are presented in sets, which may range from about three to seven. Thus, a set size score can additionally be determined based on the maximum number of words that can be recalled perfectly. As noted earlier, a composite Z-score may be used to combine separate performance indicators from one or multiple tasks. Leeser and Sunderman (2016) found significant, positive correlations among five distinct scoring methods, for both the reading and operation span tasks, yet these methods did not all significantly correlate with L2 processing. This underscores the need to plan carefully in order to determine scoring procedures in WM research. Conway and colleagues (2005) offered an excellent discussion for readers interested in the many decisions involved in computing WM tasks scores.

With these basic considerations in mind, Wen et al. (2021) argue that SLA researchers should stay abreast of psychologists’ work or, even better, collaborate with them to devise and contribute new formats of WM span tasks that are reliable and ecologically valid for L2 research (which then could be contributed to the IRIS database website https://www.iris-database.org/). In addition, it can be expected that consistency in documentation and the adoption of Open Access will benefit subsequent research efforts such as replication or meta-analysis. Above all, systematic decisions are needed as to what, exactly, can and should be used to measure WM and for what types of L2 learners and tasks. To do this, as we shall further argue, also requires a paradigm shift in the conception of WM from a structural view to a more dynamic and functional view that takes into account the adaptive and interconnected nature of WM.

### Practical Applications

Having already touched on some practical considerations for researchers using WM tasks in the previous section, we here consider two further concerns: 1) specific practices that educators can use to support language learners whose WM is below average, and 2) the potential for WM training to improve not only WM but also skills related to L2 development. In each area, research is limited in comparison to that investigating the effects of WM on L2 processing and production; thus, they constitute important avenues for future pedagogically oriented research.

First, sound advice for teachers of children who need interventions to support WM can be found outside of the field of SLA (Gathercole & Alloway, 2008; Novick et al., 2019). Within SLA, Gregersen and MacIntyre (2014) have offered clear guidance to address issues with WM. Noting its significant contribution in various areas of SLA, as well as its dynamic role in shaping outcomes at different stages of L2 performance, these authors provided a threefold action plan, which included attempting to avoid WM failures, not equating slower processing with reduced learning ability, and recruiting multimodal technology in support of learning. They described several classroom activities based on “SRVS Working Memory Enhancement” (p. 80), which invokes four elements: “schemactivate”, rehearse/repeat, visualize, and strategize. Such materials could be provided to
teachers for use in developing their own lesson plans tailored to provide WM support in a wide range of L2 instructional settings.

Furthermore, the literature on TBLT has provided theoretical frameworks and empirical accounts that could be used to reduce task demands and thereby alleviate the burden on WM (Ellis et al., 2020). For example, the sequencing of pedagogical tasks can be arranged according to task complexity (e.g., from simple to complex, in terms of, for e.g., the number of elements in a task; Robinson, 2015) or task procedure (e.g., structured vs. unstructured; Foster & Skehan, 1996). In a similar vein, whether or when to provide strategic support or interventions, during task preparation or task completion stages, can be manipulated strategically to improve task performance (e.g., Skehan, 2015). Future research can be effectively extended to explore how carefully designed tasks may help circumvent WM limitations during task performance (R. Ellis, 2005; Skehan, 2015, 2021).

Turning to the second practical application of WM research, there has been a recent proliferation of apps and other technologies intended to promote cognitive performance and WM functioning. Though the effectiveness of such training is a hotly debated research area, work has been carried out to investigate whether or not enhanced performance extends to the ability to learn an L2. Tsai et al. (2016) put forth a case for the malleability of WM, describing research based on the hypothesis that transfer of training across WM and language tasks may occur when their processing components match and they engage similar brain mechanisms. As they cautiously noted, empirical evidence for this possibility in L2 studies remains sparse; however, at least one study has shown a link between visual (but not auditory) WM training, using the N-back task, and improved ability to learn Chinese characters by L1 German learners, as well as related changes in neural activation (Opitz et al., 2014). However, other studies, such as that by Hayashi’ (2019), failed to demonstrate far-transfer effects among L2 English learners, despite improved performance on WM measures.

**Future Directions**

Inspired by the pioneering work of such SLA researchers as Diane Larsen-Freeman, Nick Ellis, Zoltan Dörnyei, Kees de Bot, Marjolijn Verspoor, and others, the field of applied linguistics in the last two decades has seen blossoming research on, and applications of, the tenets and basic principles of the complex dynamic systems theory (CDST) approach as a meta-theory to track and explain L2 learning and development (e.g., Larsen-Freeman & Cameron, 2008; Hiver & Al-Hoorie, 2020). Thus far, encouraging progress has been reported in several domains of SLA research, including IDs such as L2 motivation, willingness to communicate, and L2 anxiety (Gurzynski-Weiss, 2020). Here, we further highlight several tenets of the CDST approach relevant to understanding the WM construct (Jackson, 2020). In doing so, unresolved issues raised previously are revisited in order to illuminate how they might be profitably dealt with under the CDST approach, which entails viewing phenomena as interconnected, emergent, and developmental.

To begin with, WM is multicomponential and multifunctional, and therefore it should be conceptualized and measured as such in SLA studies. Within this multifaceted and complex system, interactions can take place at different levels. For example, these interactions can exist either between the WM components (e.g., PWM vs. EWM; Wen, 2016) or among the embedded functions (e.g., updating, task switching, and inhibitory control; Miyake & Friedman, 2012). Meanwhile, WM can also interact with other internal or external factors dynamically, namely as a function of age, proficiency, or pedagogic tasks. The interactions between learner-internal and learner-external factors give rise to the complete interconnectedness of the complex WM-SLA nexus (Wen, 2012 & 2016). For instance, Gass and Lee (2012) used CDST to frame their examination of the interrelatedness of learner-internal factors, modeling the connections between WM and inhibitory control in both L1 and L2.

Second, WM is emergent. This means that SLA research should undertake a careful reexamination of its construct definition and measurement practices to reflect the idea that WM capacity is...
a dynamic process instead of a fixed state (Simmering & Perone, 2013). Many cognitive psychologists and neuroscientists (e.g., Postle, 2006) have proposed that WM is an emergent process of the evolving interactions of the human mind and brain. In a similar vein, some psycholinguists (e.g., Schwering & MacDonald, 2020; cf. Juffs, 2017) have thus argued for a reconceptualization of WM as emergent in language comprehension and production, thus necessitating the incorporation and adoption of emergent models of WM (such as those by Cowan and Engle) as viable alternative models to Baddeley’s structural view. We, therefore, call for future WM-SLA research to also incorporate alternative WM perspectives and models such as those belonging to the functional or emergent category (Schering & MacDonald, 2020). Studies in this vein have the potential for investigating interactions between WM and pedagogic task design, as proposed, for instance, by Robinson (2015).

Third, there is abundant evidence from psychology and cognitive development literature that WM capacity increases throughout childhood and declines later in life, thus rendering it a developmental concept. WM research adopting this perspective has enormous potential to address key issues in the expanding field of SLA. These include age effects, cognitive training, and the complex relationship between WM and the bilingual advantage (Grundy & Timmer, 2017). Regarding age effects, competing accounts of the advantages possessed by child language learners, namely the “less is more” versus “more is more” hypotheses (Brooks & Kempe, 2019), view the role of WM rather differently. In addition to age effects, CDST could offer valuable insight into the issue of malleability (Tsai, et al., 2016) because it may turn out that findings in this area are affected by differing trajectories of improvement across the domains of language and cognition. Moreover, CDST seems ideally suited to investigating the issue of the bi-directional growth of WM and language proficiency (Calvo et al., 2016). Couched within the CDST approach, both WM (Simmering & Perone, 2013) and bilingualism (Li, 2015) can be considered dynamic processes exhibiting, at times, stability or fluctuation (Jackson, 2020; Serafini, 2017).

To conclude, though the WM-SLA enterprise has undergone some 30 years of intensive research, many outstanding issues remain to be solved. In the years to come, it is likely that bilinguals or multilinguals will increasingly outnumber monolinguals, so it is high time to further explore how multidisciplinary insights from psychology and neuroscience can shed light on core issues in WM and SLA, or bilingualism, as essential aspects of human cognition and action. The ultimate success may hinge upon the extent to which WM with its limited capacity and multiple functions may emerge as central to language aptitude (Wen, 2019, 2021; Wen & Skehan, 2021), gradually being recognized as part and parcel of the abilities that optimally shape and modulate essential domains of human language design and evolution, bilingual acquisition, processing, and development (O’Grady, 2017; Schwieter & Wen, 2021).

Note

1 Many SLA studies combining multiple WM measures transform individual scores into a Z-score and use the average of these scores as one composite score, as described by Leeser and Sunderman (2016).

References


Brooks, P. J., & Kempe, V. (2019). More is more in language learning: Reconsidering the less-is-more hypothesis. Language Learning, 69(S1), 13–41.


Zhisheng (Edward) Wen and Daniel O. Jackson


Working Memory


