Background

A substantial body of work has investigated the role of cognitive abilities in learning a second or subsequent language (L2), including how individual differences in such abilities affect individual differences in L2. This research has increasingly considered contributions of domain-independent (neuro)cognitive mechanisms that have independently been shown to underlie learning and long-term memory. Most of this work has examined declarative memory (DM) and procedural memory (PM). Although various conceptualizations of DM and PM have been proposed in L2-related research (DeKeyser, 2020; Paradis, 2009; Ullman, 2020), with correspondingly different claims about their roles in L2 (for comparisons and discussion, see Morgan-Short & Ullman, in press; Ullman, 2020), here we focus on one perspective: the declarative/procedural (DP) model of language.

This theoretical account of L2 and L1 (first language) is based on current neurocognitive research. In brief, the DP model proposes that two evolutionarily ancient learning and memory systems (circuits) in the brain—declarative and procedural memory—have been co-opted, that is, hijacked, for language (as well as for other higher cognitive domains such as reading and math). For more information on the model, see Ullman (2004, 2016, 2020) and Ullman et al. (2020). In this framework, the two memory systems are neuroanatomically grounded. Specifically, DM is defined as the learning and memory that rely on the medial temporal lobe (including the hippocampus) and its associated circuitry, while PM is defined as the learning and memory that rely on the basal ganglia and its associated circuitry (Ullman et al., 2020; Ullman, 2020). Crucially, the two systems have been extensively studied in both humans and non-human animals, leading to a substantial independent understanding of the systems that can generate a wide range of specific and often novel predictions for language (see especially Ullman, 2016 and Ullman et al., 2020; for L2, see Ullman, 2020). The following predictions are of particular relevance to the present chapter.

First of all, because DM seems to be critical for learning arbitrary pieces of information and their associations, this memory system should also be critical for learning such idiosyncratic information in language, including lexical knowledge. In contrast, lexical knowledge is expected to depend much less on PM (Ullman et al., 2020).
Grammar is a different story. Given the commonalities between PM and this linguistic domain (e.g., both are involved in implicit knowledge, rules, sequences, categories, and prediction, and acquired knowledge can be automatized in both), PM is expected to play an important role in grammar, across phonology, morphophonology, morphosyntax, and syntax (Ullman et al., 2020). However, since DM is extremely flexible, and thus can learn just about anything, DM is predicted to also play roles in learning grammar. Yet the two systems should support grammar in different ways (Ullman et al., 2020). PM may subserve the learning of rule-governed (generalizable) grammatical knowledge that underlies real-time combination across language domains (e.g., in affixed forms such as “kick” + “-ed”). In contrast, due to its flexible nature DM can memorize individual chunks (e.g., “kicked”), generalize analogically across such forms to new forms (e.g., “blicked”), and/or learn explicit rules.

Because the two memory systems have different characteristics, grammar (as well as other aspects of language that can depend on both systems, such as speech–sound representations and perhaps pragmatics) should depend more on one or the other system and in different ways as a function of various factors. First, DM learns more quickly than PM. Therefore, grammar should depend more on DM during early stages of L2 (and L1) learning, but more on PM at later stages. Second, learning in PM seems well-established early in life and then may decline, whereas learning in DM is poor early on, but then improves during childhood. Therefore, grammar should tend to rely more heavily on PM for early-learned languages (L1 or L2), and more on DM for later-learned languages. Note that this last point refers to the developmental trajectory (related to chronological age), whereas the previous point refers to the learning trajectory (independent of chronological age). Third, although both memory systems underlie implicit knowledge, only DM underlies explicit knowledge. Therefore, explicit instructed grammatical information (e.g., in classroom settings) should lead to a greater dependence of grammar on DM, whereas an absence of such input may shift reliance to PM—especially when instead the learner is exposed to an immersion environment. This is because the rapid input of natural speech may be particularly well-suited for procedural learning, given that this system learns by predicting subsequent elements and receiving rapid feedback (e.g., subsequent linguistic elements can be predicted, and the rapid appearance of such elements in the input in natural speech can constitute rapid feedback) (Ullman, 2020).

Additionally, evidence suggests that both systems may have greater difficulty learning non-adjacent than adjacent relations, though for somewhat different reasons: adjacent elements seem to be learned faster than non-adjacent elements in PM, and adjacent elements can be more easily chunked in DM than non-adjacent elements (Ullman et al., 2020). Various predictions follow. For example, specific phonological strings should be reasonably easy to learn as chunks in DM (e.g., “cat” and “kicked”). Morphologically complex forms that involve simple adjacencies (i.e., morphophonological relations such as in the regular English past-tense) should also be learned with relative facility, either based on rules in PM or as chunks in DM. In contrast, morphosyntactic relations, such as agreement, may be harder to learn, either as rules in PM or as chunks in DM, in particular when these involve non-adjacent dependencies (e.g., The man who … walks …).

Finally, the DP model makes predictions about receptive versus expressive language. For example, recall is harder than recognition in DM (Eichenbaum et al., 2007), whereas this contrast might not hold for PM. Therefore, grammar may depend relatively less on DM and more on PM in expressive language (e.g., in production tasks) as compared to receptive language (e.g., in comprehension or judgment tasks). For more details and explanations of the DP model’s predictions, see especially Ullman (2020) and Ullman et al. (2020).

Research

Data Elicitation

Prior research examining the predictions of the DP model in L2 and L1 has employed a wide range of methodological approaches, with the evidence broadly supporting the model (Morgan-Short
Declarative and Procedural Memory in L2

& Ullman, in press; Ullman, 2004, 2016, 2020; Ullman et al., 2020). Here we focus on a specific psycholinguistic approach, both because it is especially effective at revealing links between language and the learning and memory systems (Hamrick et al., 2018; Ullman, 2020) and because it is based on and can reveal individual differences.

This “correlational approach” simply examines predicted correlations between how well participants learn or process language and how well they learn in the memory systems. For example, one can test the predicted reliance of lexical abilities on DM by examining whether individuals who are better at learning in DM are also better at word learning or have a larger vocabulary. Thus, the correlational approach leverages individual differences both in language abilities and in learning abilities in the memory systems to test for links between them. The correlational approach may be a particularly valid test of the predictions of the DP model since, as we have seen, declarative and procedural memory are operationalized as the learning and memory abilities that depend on particular neural circuitry. Therefore, testing associations between language abilities and these learning abilities—that is, learning performance in tasks that have been independently linked to the neural circuitry of one or the other memory system—can directly test the model’s predictions (Hamrick et al., 2018; Ullman, 2020).

Nevertheless, this approach must be used cautiously, since correlation does not imply causation. For example, a correlation between word learning and DM could be explained not only by words being learned in DM but instead or in addition by some other cognitive process that underlies both word learning and DM. There are various ways that one can address this problem, such as minimizing the involvement of these processes in one’s task (e.g., minimizing working memory involvement in the task probing DM), or attempting to hold these processes constant if one can identify them (e.g., covarying out working memory in statistical analyses). Another approach is to demonstrate the specificity of the correlation. Importantly, given that the DP model generates multiple specific predictions regarding which aspects of language depend on which memory systems in which circumstances (see above), it is quite unlikely that the full pattern of predictions will be explained by other accounts, including other processes, flipped causation (better memory due to better language), or chance (Hamrick et al., 2018; Ullman, 2020).

A recent paper by Hamrick et al. (2018) provided an in-depth examination of studies using this correlational approach. It presented multiple meta-analyses that synthesized findings from 16 L1 and L2 studies of such language/memory system correlations, with a total of 665 participants. The meta-analyses tied lexical abilities to learning only in DM, while grammar was linked to learning in both systems in both child L1 and adult L2, in specific ways. Of interest here, in adult L2 learners, grammar was associated with only DM at lower levels of language experience, but with only PM at higher levels of experience. The findings yielded large effect sizes and were observed across languages, language families, linguistic structures, and tasks, underscoring their reliability and validity. The results provide compelling evidence that—as predicted by the DP model—language is indeed linked closely to the two learning and memory systems in typical development, both in children acquiring their native language and in adults learning an additional language.

Nevertheless, this seminal study has a number of gaps, which we attempt to address in the present chapter. First, although the meta-analyses captured general patterns, in particular regarding lexicon and grammar, and grammar at lower versus higher L2 experience, it did not delve into further predictions of the DP model. In fact, it generalized across various conditions in suggesting that the same pattern held not only across different languages and language families, but also across different aspects of grammar (including regular morphophonology, morphosyntax, and syntax), across both receptive and expressive language, and across various measures of both declarative and procedural memory. Although this powerfully captures the generality of the findings, it does not reveal more fine-grained distinctions of the sort discussed in some of the predictions laid out above. Second, the meta-analyses and included studies were restricted to lexical and grammatical abilities, with the latter excluding phonology. Thus, the meta-analyses did
not include correlational studies examining associations between the memory systems and other aspects of language, including not only phonology but also other functions such as speech–sound representations, pragmatics, or more general measures such as overall L2 proficiency. Third, given that an increasing number of studies using this approach have examined links between language and the memory systems, additional studies have been published even in the short span of time since Hamrick et al. (2018).

Crucially, the methodological issue of the selection of appropriate tasks for probing one or the other memory system was also outside the scope of Hamrick et al. (2018). Here we discuss four criteria that are imperative for achieving high task validity. We emphasize validity rather than reliability, since high reliability in no way implies high task validity; if the task does not measure what it’s supposed to measure, its reliability does not matter. As we lay out the four issues in the following paragraphs, we focus on how they relate to tasks employed in the studies presented below. For more information on these tasks, see Supporting Table 1, which has been uploaded to the Open Science Framework at https://osf.io/h8e9k/ (also see Buffington et al., 2021; Janacsek et al., 2020; Lum et al., 2015; Ullman, 2020; Ullman et al., 2020).

First, it is critical that the tasks that are selected to probe learning in one or the other system have been independently shown (e.g., in lesion or neuroimaging studies) to reflect the learning that depends on the key brain structures of that system—in particular, the medial temporal lobe for DM and the basal ganglia for PM. Some types of tasks or even specific tasks have been reasonably clearly tied to DM; that is, learning in the tasks has been linked to DM neural substrates, especially the medial temporal lobe. These include classes of tasks such as recognition memory tasks (e.g., the Continuous Verbal Memory Test, or [CVMT]), paired-associates tasks (e.g., the Modern Language Aptitude Test Part V [MLAT-V], and the LLAMA-B), and list-learning tasks—even if learning in specific such tasks might not yet have been tied to the medial temporal lobe. In contrast, other tasks have been tied quite clearly to PM, such as sequence learning in classic Serial Reaction Time (SRT) tasks. Additional tasks have also been linked to PM, though perhaps not as clearly, including sequence learning in Alternating SRT (ASRT) tasks, category learning in some versions of classification tasks (e.g., the Weather Prediction Task [WPT] with rapid feedback), and skill learning in the Tower of London (TOL) task. Still, other tasks have not yet been obviously linked to this memory system (e.g., many “statistical learning” tasks, which cover a wide range of task types); note that in some cases this is due to a lack of evidence, rather than evidence suggesting that the task is not tied to the memory system. We emphasize that even if learning in a given task (e.g., the SRT) has been clearly shown to rely on one or the other brain circuit, changes to that task—sometimes even relatively minor changes—can change this dependence (Conway et al., 2019); thus, modified tasks must be treated with caution.

Importantly, the presence of explicit knowledge in a learning task can be taken to suggest a dependence on DM, since such knowledge must have been learned in this system. However, if a task involves implicit learning, this in itself tells us nothing about which system it depends on, since both systems—as well as others—underlie implicit knowledge. Relatedly, tasks deemed to probe “non-declarative memory” (which is not equivalent to implicit memory, since even DM underlies implicit knowledge) are also non-specific in their targets, since the term is generally taken to reflect all learning and memory other than that which occurs in DM. This is akin to trying to pinpoint functions to either the liver (DM) or the non-liver (non-declarative memory), which does not seem highly informative.

Second, tasks probing learning in one or the other memory system should be as process-pure as possible, in that they should reflect only learning in that system. More realistically, since it is difficult if not impossible to create process-pure tasks, at the very least such tasks should minimize the involvement of other processes. Unfortunately, many tasks do not do so. For example, various tasks designed to probe DM in fact also involve working memory, such as list-learning tasks (which involve the immediate repetition of lists), and possibly also tasks with testing that immediately fol-
Declarative and Procedural Memory in L2

lows study (MLAT-V, LLAMA-B) rather than testing after a delay, which should decrease reliance
on working memory. As mentioned above, although the influence of confounding factors such as
working memory can be reduced by covarying out independent measures of these factors, this has
been rare in correlational studies of the sort discussed here. Finally, along the same lines, composite
or other over-arching measures (e.g., latent variables) that include other measures in addition to
those probing one of the memory systems (e.g., a score from the full LLAMA) are problematic
because any correlations may be due to these other measures.

Third, and related to the issue of process purity, many tasks designed to test learning in PM can
also involve learning in DM (Ullman et al., 2020). This dependence on declarative versus proce-
dural memory is a function of similar types of factors as are laid out above for grammar. For ex-
ample, tasks that may be learned in PM, such as sequence or category learning, may rely particularly
on DM in circumstances involving slow (vs. fast) feedback, less (vs. more) experience, and explicit
instructed (vs. implicit uninstructed) input (Foerde & Shohamy, 2011; Poldrack et al., 2001; Ullman
et al., 2020). Thus, it is critical that these factors be taken into account when selecting or designing
tasks to be used to probe PM.

Fourth (and again related to process purity), the learning tasks should be nonverbal. Specifically,
although correlations that are found between language measures and verbal tasks that probe learn-
ing in one or the other memory system (e.g., the MLAT-V or LLAMA-B) could indeed reflect
a predicted reliance of language on the memory system, such correlations could also simply be
due to shared verbalness. Thus, these correlations would not reliably link language to the memory
systems. In contrast, this alternative account of shared verbalness cannot be invoked for nonverbal
DM or PM tasks. Moreover, some tasks that have been used to test DM with verbal material may
actually be construed as word learning tasks (again, the MLAT-V and LLAMA-B). In such cir-
cumstances, one is essentially testing links between one language measure and another language
measure, which is unlikely to reliably reflect links between language and learning in one of the
memory systems.

Thus, valid tasks probing learning in one or the other system should: 1) have been tied to the
relevant neural circuit, 2) be process-pure, 3) rely on only one of the two systems, and 4) be non-
verbal. In practice, however, few tasks satisfy all of these criteria. This includes many of the tasks in
the studies presented below (in part because some of these studies were not designed to test the
memory systems as operationalized here). Nevertheless, we emphasize that results from such tasks
can still be informative, especially if patterns are found in common across different tasks with dif-
ferent weaknesses.

Evidence

In order to systematically review correlational studies related to the predictions laid out above
regarding the role of declarative and procedural memory in L2, we first conducted a literature
search to identify relevant primary research. We aimed to identify studies that included any measure
of L2 and at least one task that potentially probed declarative or procedural memory learning abili-
ties, regardless of whether the study was actually designed to examine research questions related
to declarative and procedural memory as operationalized above. For example, if a study examined
the association between a measure of L2 and learning in an SRT task, we included the study in
our review because learning on the SRT likely reflects learning in PM—even if the SRT was
administered for a different purpose, such as assessing statistical learning. We emphasize that we only
searched for and include in our review studies that examine L2; thus, those that probe associations
between declarative or procedural memory and aspects of only L1 (e.g., Arthur et al., 2021; Lum
et al., 2012) are not included.

We conducted our search for articles in Linguistics and Language Behavioral Abstracts, PsycInfo,
and ERIC in February 2021. We searched the title and abstract fields for articles that were pub-
lished (in English) after April 2016, which was the end date for the systematic search in Hamrick et al. (2018). Our search syntax was based on that used by Hamrick et al., in that we searched for language terms (e.g., “lexi*”, “synta*”) and learning terms (e.g., “procedural mem*”, “declarative mem*”) using “AND” to connect the search terms. However, unlike in Hamrick et al., we also added a second “AND” with two L2 terms (“second language”, “foreign language”) in order to constrain the search to L2-relevant studies.

The search yielded 611 articles, from which 227 duplicate articles were identified and removed. The titles and abstracts of the remaining 384 articles were screened to assess whether they were likely to include at least one task that potentially probed declarative or procedural learning abilities and at least one L2 task. Based on these screening criteria, we identified 18 articles. We added to this list any published or in-press articles known to the authors (as of February 2021) that met our screening criteria. This included articles that were part of the L2 meta-analyses in Hamrick et al. (2018). Additionally, we included appropriate articles that were identified by Hamrick et al. but were excluded from their meta-analyses because these focused on lexicon and grammar, while our review covers all language domains. Altogether, 19 articles known to the researchers that were not already identified through the literature search were added to the reference list, for a grand total of 37 articles. A final full-paper screening was then conducted to ensure that all articles in fact examined associations between at least one task that potentially probed declarative or procedural learning and at least one L2 task, resulting in four articles being excluded. Thus, in all, 33 studies are included in our final review pool. These constituted 31 peer-reviewed articles and two peer-reviewed papers from conference proceedings. See Supporting Table 1 on the Open Science Framework at https://osf.io/h8e9k/ for details of the studies (the Table is not in the chapter itself for reasons of space).

To gain insight into just how individual differences in learning abilities in the two memory systems are related to individual differences in L2, here we discuss general patterns that emerge across the studies identified in the Table. We approach the literature by linguistic domain: lexicon, phonology, morphophonology, morphosyntax, syntax, and then other areas of L2. Within each of these subsections, we first briefly describe what types of L2 studies have been conducted. Then we consider how L2 outcomes from these studies relate to learning measures from DM and PM tasks, that is, tasks that possibly probe DM or PM. We emphasize that for each study in the Table, we report all DM and PM measures and all L2 outcomes for which associations were examined. For simplicity in regard to the interpretation of the results, we use terms like “positive relationship” to indicate that better DM or PM is linked to better L2 performance (irrespective of whether the actual finding showed a positive association or a negative association, for example in a response time study), and likewise for “negative relationship” (in contrast, for a veridical representation of the findings, in the Table we report the direction of the actual association, such as a negative correlation between better learning and faster L2 response times). Finally, we lay out any emerging pattern and consider that pattern in light of the qualities of the tasks used to probe DM and PM.

We underscore that in our review—both in the Table and in the discussion below—we report all studies that we identified as examining associations between at least one task that potentially probed declarative or procedural learning and at least one L2 task, regardless of whether or not the study reported statistically significant associations. For the sake of clarity, however, we explicitly present and discuss the specifics only of significant findings. Crucially, because virtually all the studies tested associations between all DM and PM measures and all L2 measures included in the given study, any association not reported in the Table, and thus not discussed below, can be reasonably assumed to represent null findings; see the specific studies for more details. Therefore, the patterns presented below can be interpreted as reflecting the significant effects that emerged from the full set of associations reported in the studies identified in our systematic review. For a graphical summary of the patterns, see Figure 4.1.
Ten studies examined relationships between the L2 lexicon and tasks probing learning in declarative and/or procedural memory (Brooks & Kempe, 2013; Brooks et al., 2017; Granena, 2019; Hamrick et al., 2020; Mor & Prior, 2020; Murphy et al., 2021; Ruiz et al., 2018; Ruiz et al., 2021; Saito, 2017; Walker et al., 2020). Of these, five were conducted in laboratory settings, four in classroom settings, and one in a web-based setting. Three studies examined learning of artificial L2s, while the remainder probed natural L2s. The lexical measures came from various tasks, including word-picture matching, oral production tasks, and multiple-choice tasks.

Of the ten studies, seven examined relationships between the L2 lexicon and DM tasks. Four of these evidenced significant positive relationships or related interactions (Hamrick et al., 2020; Murphy et al., 2021; Ruiz et al., 2021; Walker et al., 2020), across different learning contexts and L2 tasks, with both natural and artificial L2 lexicons. Additionally, five of the studies examined associations between the L2 lexicon and PM tasks, with only one finding a significant result (Walker et al., 2020). That study found a relationship between the PM task and an L2 factor (latent variable) that included noun and adjective lexical information, but also case-marker information, which makes it unclear whether the association was driven by PM influence on the lexicon, morphosyntax, or both.

Thus, in the majority of studies, the L2 lexicon was associated with DM tasks, while no clear association with PM tasks was evinced. Importantly, most of the positive associations between the lexicon and DM used a nonverbal DM task (CVMT). This underscores the validity of the
DM-lexicon association in L2, since these findings cannot simply be due to shared verbalness (see the Background section). The pattern also fills an important gap left by Hamrick et al. (2018), since the meta-analyses in that paper did not include any studies examining the L2 lexicon.

**Phonology**

Three studies examined associations between L2 phonology and DM or PM tasks (Bowles et al., 2016; Saito, 2017, 2019). Two of these tested advanced L2 classroom learners’ performance on various oral production tasks probing aspects of phonology (Saito, 2017, 2019). One study trained naïve learners on a new phonological system and then tested them on a receptive task (Bowles et al., 2016).

All three studies found positive associations between L2 phonological measures and DM tasks. Only one study (Bowles et al., 2016) included a PM task, which yielded no association with phonology. This may be taken to suggest that L2 phonology is positively associated with DM, with no evidence for a relationship with PM. The finding of DM-phonology links even in advanced learners (Saito, 2017, 2019), in which a role for PM might rather be expected, is perhaps less surprising given that both studies examined classroom learners, in which a greater reliance on DM may be expected. Note also that possible links in these two classroom studies with PM are unknown, since these studies did not report any PM tasks. Moreover, the absence of a link with PM in the third study might have been due to insufficient training (Bowles et al., 2016). Thus overall, although the evidence to date suggests a relationship between L2 phonology and only DM, caution is warranted for various reasons, including the small number of studies, the fact that all of the DM tasks were in fact verbal, and the particular dearth of PM tasks, especially at higher levels of training.

**Morphophonology**

Three studies tested DM and PM associations with L2 morphophonology (Antoniou et al., 2016; Buffington & Morgan-Short, 2018; Ettlinger et al., 2014). All three of these lab-based studies trained naïve participants on an artificial linguistic system, specifically on both “simple” affixation-based and “complex” analogical morphophonological transformations, with a receptive task following training.

All three studies found positive associations between L2 morphophonology and DM tasks. These associations were found both for the analogical (Antoniou et al., 2016; Ettlinger et al., 2014) and the affixed forms (Antoniou et al., 2016; Buffington & Morgan-Short, 2018). In contrast, associations with PM tasks were found only for the affixed forms, across all three studies.

The evidence to date thus suggests that whereas DM can support both analogical and affixed forms, PM may only support the latter. This pattern is consistent with the DP model, which predicts that analogical forms can be generalized in DM (but not PM), while affixed forms can either be combined in PM or stored as chunks and generalized in DM (Ullman, 2016). The PM tasks were nonverbal, underscoring the validity of the PM-L2 morphophonology associations. However, thus far there have only been a handful of studies. Moreover, the PM tasks may not be particularly process-pure (TOL, dual task-WPT; Buffington et al., 2021), and thus it is possible that the correlations with these tasks might have reflected some DM (or other) involvement. Finally, two of the three DM tasks were verbal, and thus DM–morphophonology associations could be explained by shared verbalness for such tasks. Therefore, some caution is warranted in drawing conclusions at this point.

**Morphosyntax**

Twelve studies were identified that examined associations between L2 morphosyntax and DM and/or PM tasks (Brooks & Kempe, 2013; Brooks et al., 2017; Curcic et al., 2019; Granena, 2013; Granena & Yilmaz, 2019; Saito, 2017; Suzuki & DeKeyser, 2015, 2017; Suzuki, 2018; Walker et al.,
Declarative and Procedural Memory in L2

They comprised two classroom studies, three studies conducted in immersion contexts, and seven studies in lab settings. The majority of the lab studies examined uninstructed conditions with natural language, with two of these also including an instructed condition. L2 measures of morphosyntax varied and included GJTs, production tasks, and visual-world eye-tracking measures.

Of the 12 studies, four included DM tasks. Of these four, two found positive relationships between DM and L2 measures: one with L2 morphosyntax in classroom learners (Yalçın & Spada, 2016) and the other specifically with explicit knowledge of L2 morphosyntax (Curcic et al., 2019). Note that another of the four studies, Walker et al. (2020), found a positive association between a DM task and an L2 factor (latent variable) that included morphosyntactic as well as lexical information, precluding conclusions about morphosyntax specifically; also see above under the Lexicon subsection.

Nine of the 12 studies included PM tasks. Eight of these reported positive associations between these tasks and L2 morphosyntax. These associations were found across various L2 forms (e.g., case marking and grammatical gender agreement) and for different contexts (e.g., lab-based and immersion). However, the associations were often specific to particular conditions or groups. For example, positive associations with PM tasks were found for old but not new items after training in two studies (Brooks & Kempe, 2013; Brooks et al., 2017); for learners who received implicit feedback but not for those who received explicit feedback in two studies (Granena & Yilmaz, 2019; Yilmaz & Granena, 2019); and for learners whose practice was spaced apart three days rather than seven days (Suzuki, 2018). Moreover, for immersed learners, associations were found specifically for late but not early learners (Granena, 2013) and for advanced speakers with longer lengths of residence (Suzuki & DeKeyser, 2015; although no association was found for immersed advanced speakers in Suzuki & DeKeyser, 2017).

Overall, the evidence does not support a strong relation between L2 morphosyntax and DM. Where such a relation was found, it may have been due specifically to the L2 learning context (classroom) or the fact that the knowledge was explicit. Both of these outcomes are consistent with the DP model. Additionally, the validity of these DM-L2 morphosyntax relations is weakened by the fact that the DM tasks were verbal, and thus the associations might have been due to shared verbalness. In contrast, there seems to be growing evidence for a relationship with L2 morphosyntax and PM. Interestingly, this relationship might be modulated by a variety of factors. Although some of the situations in which PM-L2 morphosyntax associations have been observed are consistent with the DP model (e.g., implicit feedback, immersion and advanced speakers), others may require more investigation (e.g., old vs. new items and less vs. more spacing). Importantly, almost all of the PM tasks were nonverbal, and moreover the majority were SRT tasks (which have been clearly tied to PM; see the Background section), underscoring the validity of associations between PM and L2 morphosyntax. Finally, note that all but one of the L2 morphosyntax studies discussed here were not actually designed to examine DM or PM, and thus were presumably not biased to find associations with these memory systems. This further strengthens the validity of the observed relationships.

Syntax

Ten studies examined associations between L2 syntax and DM and/or PM tasks (Brill-Schuetz & Morgan-Short, 2014; Faretta-Stutenberg & Morgan-Short, 2018; Granena, 2019; Hamrick, 2015; Morgan-Short et al., 2014; Pili-Moss et al., 2020; Ruiz et al., 2018; Saito, 2017; Tagarelli et al., 2016; Walker et al., 2020). Of the ten studies, three were conducted with classroom learners, with one of these also including study-abroad learners. The remaining seven were conducted in the lab with uninstructed groups, though three of these studies also included instructed groups. Various measures probing L2 syntax, such as oral production, GJT, and event-related potentials, were employed across the studies.
Eight of the studies included DM tasks. Positive associations were found in five of them, all of which were artificial language studies. Associations with DM were observed with syntax especially at earlier stages of L2 learning (Hamrick, 2015; Morgan-Short et al., 2014; Ruiz et al., 2021; Walker et al., 2020), and in one study with a receptive (but not expressive) measure of practice that tapped syntax, across both earlier and later stages of learning (Pili-Moss et al., 2020).

Additionally, eight studies included PM tasks. Five of these showed positive associations with L2 syntax. These positive associations were found 1) at later but not earlier stages of L2 learning (Brill-Schuetz & Morgan-Short, 2014; Hamrick, 2015; Morgan-Short et al., 2014); 2) in un instructed or immersed rather than instructed or classroom L2 learners (Brill-Schuetz & Morgan-Short, 2014; Faretta-Stutenberg & Morgan-Short, 2018; Hamrick, 2015; Morgan-Short et al., 2014); 3) for an expressive (but not receptive) measure of practice that tapped syntax, especially during earlier stages of learning (Pili-Moss et al., 2020); and 4) for automaticity on a receptive measure of practice, although this interacted with DM such that the positive relationship with PM became stronger over practice for learners with higher DM (Pili-Moss et al., 2020). Interestingly, one study reported a negative association, which was found after a relatively short exposure period (Tagarelli et al., 2016); the reasons for this finding remain unclear.

Overall for L2 syntax, the pattern reported by Hamrick et al. (2018) across syntax and other aspects of grammar—with DM playing a role at earlier stages of L2 learning and PM playing a role at later stages—seems to hold true specifically for syntax, even for this larger set of L2 syntax studies than those that were examined by Hamrick et al. The evidence presented here reveals further patterns as well. It links syntax following un instructed and immersed L2 learning, particularly to PM. Moreover, receptive syntax showed associations mainly with DM while expressive syntax was linked to PM. All of these patterns are expected by the DP model, though surprisingly, syntax in instructed and classroom contexts showed relations with neither memory system. It also remains unclear why PM was linked to an expressive measure tapping syntax during earlier rather than later stages of practice. No discernible pattern emerged based on the particular tasks used to measure DM or PM, suggesting the generalizability and validity of the associations.

Other Areas of L2

Several studies have also examined relations between DM and/or PM tasks and other areas of L2. Three studies examined and indeed found positive associations between DM tasks and L2 proficiency, at different stages of learning across the studies (Artieda & Muñoz, 2016; Kaufman et al., 2010; Linck et al., 2013). Two of these studies also included a PM task (SRT); both showed a positive relation between the task and L2 proficiency, especially at more advanced L2 levels (Kaufman et al., 2010; Linck et al., 2013). Two studies have examined relations between DM tasks and L2 fluency measures (e.g., number of syllables or words per minute): one study found such an association in advanced L2 learners (Saito, 2017), while another study found no association in intermediate L2 learners (Granena, 2019). One of these studies also included a PM task but found no association with fluency (Granena, 2019). Finally, we are aware of one study that examined the relation between DM (but not PM) and pragmatics (Li, 2017). This study reported a positive association between a DM task and a performance measure of oral production for a pragmatic structure. Overall, these findings are intriguing and suggest that the relationship between the memory systems and language can be extended beyond lexicon and grammar. Nevertheless, the small number of these studies precludes any conclusions at this point.

Practical Applications

Inasmuch as L2 learning and processing are tied to declarative and procedural memory, what is known about these memory systems should have relevance (Ortega, 2012) for language learn-
Declarative and Procedural Memory in L2

In particular, the substantial independent knowledge of the two learning and memory systems—and how the learning, retention, and use of information in these systems can be enhanced using a variety of behavioral, pharmacological, and other approaches—should be applicable to L2 (Ullman & Pullman, 2015; Ullman & Lovelett, 2018; Ullman, 2020).

Depending on factors such as learning context, curricular aims, and learners’ goals, teachers may find it useful to employ pedagogical practices that promote successful learning in DM (e.g., explicit instruction, retrieval practice/the testing effect, and slow feedback) and/or PM (e.g., immersion contexts and rapid feedback) (Ullman & Lovelett, 2018; Ullman, 2020). For example, some learning outcomes, such as rapid early learning or better retention of vocabulary, may be best achieved via approaches that promote learning and retention in DM. Other desired outcomes, such as increased automaticity or better retention over time, may be better achieved by enhancing learning in PM—since learning in PM indeed seems to yield these characteristics more than learning in DM (Ullman et al., 2020). For instance, in computer-based L2 learning, rapid feedback could be provided to promote learning in PM (Foerde & Shohamy, 2011; Ullman, 2020). Using such instructional techniques that promote learning in PM may be especially useful in contexts where students are not exposed to immersion, which itself may enhance learning in this memory system (see the Background section).

Another potential point of relevance for language education concerns individual differences themselves, since teachers could target learning in DM and/or PM depending on students’ strengths or weaknesses. For example, teachers could promote learning in individuals’ stronger memory system to leverage these existing strengths. Alternatively or in addition, they could target learning and retention in the weaker system to level the neurocognitive playing field. It is important to keep in mind, however, that the use of behavioral (and other) methods to enhance L2 learning in these memory systems is still not well-understood, and substantial empirical research is still needed before the pedagogical validity of such approaches becomes clear.

Future Directions

In order to more fully understand the contributions of declarative and procedural memory to individual differences in second language abilities and any implications for pedagogy, further research is needed to more precisely flesh out the role of these systems in different L2 domains and learning contexts. Importantly, future research should adopt methodological approaches that increase the validity of the conclusions that can be drawn from the research. Below, we develop these points to help guide future research.

First, future studies should target L2 domains where the research has thus far been scarce, especially where predictions are reasonably clear. For example, only three studies have addressed the role of the memory systems in phonology, and it is still unclear if this domain can depend on PM at higher levels of experience. Further, we are not aware of any studies that have examined relations between learning in the memory systems and various other domains, such as speech–sound representations or adjacent versus non-adjacent relations in grammar either within, let alone across, linguistic domains (for predictions, see the Background section). Relatedly, it would be of interest to use a within-subjects approach to examine to what extent reliance on DM and PM for different domains patterns together as predicted. For example, at higher levels of L2 experience, the associations between declarative and procedural memory with phonology, morphophonology, morphosyntax, and syntax may be quite similar to each other, yet different from the association with the lexicon.

Second, future work needs to investigate the extent to which individual differences in declarative and procedural memory play out in different contexts. The majority of studies in this review examined DM and/or PM in uninstructed or immersion contexts, with 12 studies probing their roles in instructed, explicit, or classroom contexts. Notably, only six studies examined either...
instructed/uninstructed or classroom/immersion contexts within the same study. Thus, we do not yet understand possible aptitude–treatment interactions with DM and PM. Such insights will be critical to extend research on DM/PM roles in L2 learning to applied domains.

Third, more research employing within-subject longitudinal designs is needed to better understand how DM and PM may differentially contribute to L2, as well as how they interact over time. Intriguingly, as described above, one study revealed a potentially increasingly interactive relationship between the two memory systems over the course of L2 learning (Pili-Moss et al., 2020).

Fourth, moving forward, regardless of the specific research question, researchers should consider a few study design issues. For example, studies should carefully consider the amount of exposure to the L2 that is provided in relation to which questions are being asked. Given that PM is expected to play a role at later stages of L2 learning, a study with relatively little training or that focuses on low or intermediate proficiency learners might not generally be expected to detect a role for PM. Also, researchers should administer additional assessments of other cognitive factors, such as working memory and executive inhibitory function, in order to be able to either covary them out of analyses to examine the unique contributions of DM and PM and/or to examine how they may interact with learning in these memory systems. Such designs have the potential to yield valid, robust conclusions.

Fifth and finally, it is critical that future research also consider the tasks used to assess DM and PM. Here we give specific recommendations for which learning and memory tasks should ideally be used, that is, which tasks are likely to yield valid results. The validity of DM tasks should have been demonstrated with independent studies that link learning in these tasks to the hippocampus and/or other medial temporal lobe structures. Valid DM tasks should also minimize contributions of working memory, for example through the use of incidental encoding paradigms. Other potentially confounding factors, such as executive inhibitory function, should be minimized as well. The tasks should be nonverbal and should incorporate a delay between encoding and test (again, to reduce working memory contributions). For example, we suggest that a particular and increasingly used recognition memory task with incidental encoding involving real/novel object decision, and a delay between encoding and recognition, might be highly appropriate to use to test learning in DM (Arthur et al., 2021; Buffington et al., 2021; Hedenius et al., 2013; Reifegerste et al., 2021). Along similar lines, tasks probing learning in procedural memory should have independently been shown to rely on the basal ganglia and should avoid or minimize factors that can encourage learning in declarative memory. For example, classic versions of the SRT task (Clark et al., 2014; Janacsek et al., 2020) are highly appropriate. Overall, following these recommendations is likely to increase the validity of the findings and thus substantially advance our understanding of the role of the memory systems in individual differences in L2.

Conclusions

In this chapter, we presented a systematic review of L2 studies that provide evidence regarding links between individual differences in DM or PM and individual differences in L2 abilities. All of these studies employed a correlational approach, a technique that leverages individual differences both in the memory systems and in language to test for relations between them. Going beyond the meta-analyses of such correlations reported by Hamrick et al. (2018), we included studies probing a wide range of aspects of L2, including lexicon, phonology, morphophonology, morphosyntax, and syntax, as well as other areas of language. Overall, the evidence reviewed suggests not only that learning in the memory systems is at least somewhat linked with all of these aspects of L2, but moreover, these associations are by and large predicted by our independent understanding of the two memory systems—that is, they are predicted by the DP model. We discuss the practical implications of the findings for L2 learning and pedagogy. Finally, we suggest particular paths for future research. These are based not only on the gaps and mixed results that emerged from our
review but also on our specific methodological recommendations that should increase the validity of findings. Together, these should further our understanding of the contributions of declarative and procedural memory to L2 abilities.

References


Declarative and Procedural Memory in L2


