Bilingualism: Language and Cognition 17 (1), 2014, 56–72 © Cambridge University Press 2013 doi:10.1017/S1366728912000715

Declarative and procedural memory as individual differences in second language acquisition*

KARA MORGAN-SHORT
University of Illinois at Chicago
MANDY FARETTA-STUTENBERG
University of Illinois at Chicago
KATHERINE A. BRILL-SCHUETZ
University of Illinois at Chicago
HELEN CARPENTER
Upper-Story Consulting
PATRICK C. M. WONG
The Chinese University of Hong Kong & Northwestern University

(Received: October 11, 2011; final revision received: August 26, 2012; accepted: October 23, 2012; first published online 1 March 2013)

This study examined how individual differences in cognitive abilities account for variance in the attainment level of adult second language (L2) syntactic development. Participants completed assessments of declarative and procedural learning abilities. They subsequently learned an artificial L2 under implicit training conditions and received extended comprehension and production practice using the L2. Syntactic development was assessed at both early and late stages of acquisition. Results indicated positive relationships between declarative learning ability and syntactic development at early stages of acquisition and between procedural learning ability and development at later stages of acquisition. Individual differences in these memory abilities accounted for a large amount of variance at both stages of development. The findings are consistent with theoretical perspectives of L2 that posit different roles for these memory systems at different stages of development, and suggest that declarative and procedural memory learning abilities may predict L2 grammatical development, at least for implicitly trained learners.

Keywords: second language, individual differences, declarative memory, procedural memory

First language (L1) acquisition in normally developing children invariably results in the ability to speak and understand one's native language. The acquisition of a second language (L2) among adults, however, does not seem to follow the same pattern. The level of L2 attainment reached by adult learners varies widely (Dörnyei, 2005, p. 6), such that any two adult L2 learners may attain different levels of proficiency in an L2 even when they share commonalities such as native language, educational level, and experience with the L2. This variation is apparent both within studies of L2 acquisition, which find that different learners make different levels of gains, as well as across studies, where learners of similar linguistic structures show different amounts of development. Identifying individual difference factors that contribute to this variation in L2 attainment is crucial for arriving at a theoretical understanding of second language acquisition (SLA) and should also prove informative for applied perspectives, including approaches to teaching adult-learned L2.

Previous research on individual differences in SLA has examined several factors that may account for differential levels of L2 attainment including, but not limited to, age, working memory, intelligence, aptitude, developmental disorders, affective states, and memory systems, as well as auditory perceptual, neurophysiological, neuroanatomical, and potential genetic factors (e.g., Bialystok & Frölich, 1978; Carpenter, 2008; Carpenter, Morgan-Short & Ullman, 2009; Carroll, 1958, 1962, 1981; Chandrasekaran, Kraus & Wong, 2012; DeKeyser, 2000; Dörnyei, 2005; Ehrman & Oxford, 1995; Ettlinger, Bradlow & Wong, in press; Harley & Hart, 1997, 2002; Horwitz, Horwitz & Cope, 1986; Mackey, Adams, Stafford & Winke, 2010; Miyake & Friedman, 1998; Perrachione, Lee, Ha & Wong, 2001; Robinson, 2002, 2003; Ross, Yoshinaga & Sasaki, 2002; Skehan, 1986, 1991; Sparks, Ganschow & Pohlman, 1989; Wong et al., 2008; Wong, Morgan-Short, Ettlinger & Zheng, 2012; Wong & Perrachione, 2007; Wong, Perrachione & Parrish, 2007). Previous findings have been invaluable, but as

* We would like to acknowledge members of the Cognition of Second Language Laboratory at the University of Illinois at Chicago for assistance in collecting data for this study. We thank Alice H. D. Chan, Francis C. K. Wong and Michael T. Ullman for thoughtful discussions about this work, and we appreciate the insightful comments on an earlier version of the manuscript from anonymous reviewers. This work was supported by a Language Learning Small Research Grant to K.M.S. and by grants from the National Institutes of Health (R01DC008333 & K02AG035382) awarded to P.C.M.W.

Address for correspondence:
Kara Morgan-Short, 601 S. Morgan St., 1706 University Hall, M/C 315, University of Illinois at Chicago, Chicago, IL 60607, USA
karams@uic.edu

research in cognitive (neuro)science has increased and provided new insights into adult memory and cognition, examination of individual differences as informed by extant knowledge of neural-based memory systems can potentially provide further insight into variation in L2 development (Carpenter, 2008; Carpenter et al., 2009). Among the many cognitive abilities that potentially subserve L2 development, we focus specifically on the dynamic contributions of two long-term memory systems that have been well-studied in non-linguistic contexts (Eichenbaum & Cohen, 2001; Sherry & Schacter, 1987; Squire & Knowlton, 2000) and that have been posited to play an important role in both L1 and L2 acquisition (Paradis, 1994, 2004; Ullman, 2001, 2004, 2005): declarative and procedural memory. The longitudinal L2 study of adult learners reported here explores how individuals, who receive implicit training (i.e., exposure to meaningful L2 examples without provision of metalinguistic information), may rely on these two memory systems differentially at early and late stages of development of an artificial L2.

Declarative and procedural memory as individual differences in L2 development

There are multiple distinct, long-term memory systems in the brain, many of which are commonly categorized as being either DECLARATIVE or NONDECLARATIVE (Eichenbaum & Cohen, 2001; Sherry & Schacter, 1987; Squire & Knowlton, 2000). Declarative memory comprises knowledge about facts and events related to the world or the self. It has been characterized as knowledge "that" and encompasses representations of both semantic memory, which is knowledge of facts about the world, and episodic memory, which is knowledge of events that one has experienced (Tulving, 1993). Learning with the declarative memory system (a) can occur quickly and after only one exposure to the information to be learned, (b) requires attentional resources, such as working memory, and (c) benefits from the intention to learn (Knowlton & Moody, 2008). Specific brain structures subserve declarative memory, primarily including bilateral medial temporal lobe structures, including the hippocampus, and associated neocortical regions (for more information, see Ullman, 2001, 2004). Declarative memory may be assessed using recall and recognition tasks, both verbal, such as paired-associates word recognition tasks (Knowlton & Squire, 1995), and nonverbal, such as the Continuous Visual Memory Task (CVMT; Trahan & Larrabee, 1988), among other measures.

Although the term declarative memory refers to a unitary memory system, nondeclarative memory refers to a number of memory systems that are dissociable from declarative memory and from each other. In the field of cognitive neuroscience, the procedural memory system is considered to be the nondeclarative system that underlies both motor and cognitive skill and habit learning (Eichenbaum, 2002; Eichenbaum & Cohen, 2001; Knowlton & Moody, 2008). Often, procedural memory is described as knowledge of "how" (Ohlsson, 2008) and is at play when individuals display behaviors involving the integration of coordinated actions or sequences, such as driving a car or typing fluently – one's hands appear to know how to accomplish the task while one attends to other thoughts or actions. Procedural memory is not available to conscious awareness; it is considered to be a type of nondeclarative, implicit memory (Eichenbaum & Cohen, 2001; Squire & Zola, 1996). Note that although all procedural memory is implicit, not all nondeclarative, implicit memory is procedural, as there are several different types that have been identified, e.g., priming, simple forms of conditioning, perceptual learning, among others (Squire & Wixted, 2011). Learning that relies on the procedural memory system (a) occurs gradually with repeated exposure or experience with the task to be learned, (b) appears to require fewer attentional resources than declarative learning, and (c) can occur even without the intention of learning (Knowlton & Moody, 2008). The procedural memory system is supported by particular brain structures, including frontal lobe/basal ganglia circuits as well as certain portions of the cortex and the cerebellum (for more information, see Ullman, 2001, 2004). The skills associated with procedural memory have been assessed using a variety of tasks, such as the Weather Prediction Task (WPT; Knowlton, Squire & Gluck, 1994), specifically the dual-task version of the WPT (Foerde, Knowlton & Poldrack, 2006) for probabilistic learning, and the Tower of London task (TOL; Kaller, Rahm, Köstering & Unterrainer, 2011; Kaller, Unterrainer & Stahl, 2012; Ouellet, Beauchamp, Owen & Doyon, 2004; Unterrainer, Rahm, Leonhart, Ruff & Halsband, 2003) for cognitive skill acquisition.[1]

The declarative and procedural memory systems are posited to play particular roles in the language domain (Paradis, 1994, 2004, 2009; Ullman, 2001, 2004, 2005). A recent model, Ullman's (2001, 2004, 2005) declarative/procedural (DP) model, ties different components of language to the declarative and procedural memory systems. For L1, Ullman posits that the declarative memory system subserves the acquisition, representation, and use of the mental lexicon, including the sounds and meanings of words,

[1] Artifical grammar learning is a task that might be viewed as reflecting procedural memory (Knowlton & Moody, 2008). We do not discuss it here or below because (a) patient work suggests that Parkinson's and Huntington patients who are impaired on other procedural tasks are not impaired on artificial grammar learning (e.g., Witt, Nuhsman & Deuschi, 2002), and (b) its relevance to L2 development has been challenged both conceptually and experimentally (Robinson, 2005; VanPatten, 1994).

irregular morphological forms, and idioms. Declarative memory may also underlie the memorization of rule-based grammar as "chunks". For example, regular past tense forms, such as *looked* may be memorized rather than computed. The procedural memory system, in contrast, is not expected to underlie any aspect of the mental lexicon or memorization of chunks. Instead, it subserves aspects of the mental grammar, including rule-based (computational) morphology, syntax, and possibly phonology, as well as other complex forms and rules, such as long distance dependencies and subordination routines. Given an increase in declarative memory and possible attenuation of procedural memory throughout childhood until early adulthood (Dorfberger, Adi-Japha & Karni, 2007; see Ullman, 2001, 2004, for further discussion), the DP model makes somewhat different predictions for adult-learned L2s than for L1. For L2, Ullman's DP model posits that the L2 lexicon will always depend on declarative memory, as in L1, because there is relatively limited attenuation of this system with age. The L2 grammar, unlike the L1 grammar, is expected to also rely on the declarative memory system, at least INITIALLY, AT LOW LEVELS OF EXPOSURE AND CORRESPONDING PROFICIENCY. Importantly however, Ullman (2005) notes that complete attenuation of procedural memory does not occur in adults and that the development of procedural memory can occur gradually with repeated experience. Therefore, the DP model predicts that WITH INCREASING EXPOSURE AND PROFICIENCY, components of the L2 grammar may come to rely on the procedural memory system. More specifically, this switch in reliance from declarative to procedural memory should be mediated by practice and experience with the L2. In sum, Ullman's DP model predicts that, at higher levels of L2 proficiency and/or exposure, learners may display native-like representations and processing of both the mental lexicon and grammar. This model has received empirical support from various lines of research, including behavioral and neuroimaging work (e.g., Bowden, Gelfand, Sanz & Ullman, 2010; Brovetto & Ullman, 2005; Ferman, Olshtain, Schechtman & Karni, 2009; Opitz & Friederici, 2003).

Other cognitive perspectives of L2 development also posit that declarative and procedural memory or knowledge account for L2 development (DeKeyser, 1995, 2007; Paradis, 1994, 2004, 2009). Perhaps most closely related to the DP model, Paradis also argues that declarative and procedural memory support language acquisition. In this model, L1 acquisition of both the lexicon and the grammar is implicit and relies on procedural memory. For L2, however, acquisition of the lexicon and the grammar is expected to be learned explicitly and to depend on declarative memory (Paradis, 1994, 2004, 2009). Only in rare cases and under implicit, communicative learning conditions are L2 learners expected to come to rely on procedural memory (Paradis, 1994, 2009). Declarative and procedural notions have also been evoked in skill acquisition perspectives of L2 acquisition (DeKeyser, 2007). According to this perspective, there are three stages of development: declarative, procedural, and automatization. In the declarative stage, learners must develop explicit knowledge about a L2 form. With practice, this declarative knowledge becomes proceduralized, such that declarative information is compiled into a set of productions that can be executed without resorting to declarative knowledge. Finally, with repeated use, access to procedural knowledge is automatized. Note that there are important differences between these three perspectives, including which linguistic forms can come to rely on procedural memory, whether L2 development necessarily passes through an explicit stage or not, and what the underlying neurocognitive mechanisms are, but all perspectives elicit notions of declarative and procedural memory or knowledge and have some overlapping predictions, including the expectation of some kind of qualitative shift in the representation and processing of L2. Although the present study's research questions and design are directly motivated by the DP model, which is based on extant knowledge of the declarative and procedural memory systems as informed by a multiple memory systems perspective (Eichenbaum & Cohen, 2001; Sherry & Schacter, 1987; Squire & Knowlton, 2000; Squire & Wixted, 2011), the design does not include any factor that would necessarily distinguish the DP model from other neurocognitive or skill acquisition models that posit a role for declarative and procedural memory or knowledge. Thus, findings from this study that are consistent with the DP model should also be largely consistent with other related models (DeKeyser, 1995, 2007; Paradis, 1994, 2004, 2009).

Given the potential roles of declarative and procedural memory in L2 learning, we hypothesize that an individual's ability to learn within these memory systems may be predictive of L2 attainment. To the authors' knowledge, this hypothesis has only been examined directly in one previous dissertation study (Carpenter, 2008; Carpenter et al., 2009), which examined the contributions of individual differences in declarative and procedural memory to L2 at different stages of proficiency. The research was conducted as part of a larger study where adult learners of an artificial L2 were trained under either explicit conditions, where grammatical rules were explained, or implicit conditions, where learners were exposed to the language without any grammatical rule explanation (see Morgan-Short, 2007; Morgan-Short, Sanz, Steinhauer & Ullman, 2010; Morgan-Short, Steinhauer, Sanz, 2012). In Carpenter's (2008) study, declarative memory was assessed using the Modern Language Aptitude Test, Part V (MLAT-V (verbal)) and the CVMT, and procedural memory was assessed using a dual task version of the

https://doi.org/10.1017/S1366728912000715 Published online by Cambridge University Press

WPT. Second language grammatical development was assessed after participants demonstrated above chance performance during comprehension practice and then again after all practice items ($n = 880$) were completed. Carpenter's results showed that declarative and procedural memory differentially predicted L2 proficiency depending on the training condition. More specifically, for syntactic structures, regression analyses indicated that explicitly trained learners relied on declarative memory during initial learning, but that this relationship did not hold for implicitly trained learners. For the same syntactic structures, procedural memory predicted accuracy for implicitly trained learners but only as proficiency approached the most advanced stages. The relationship between syntactic development and procedural memory, however, was parabolic rather than linear, implying that procedural memory ability only benefited the portion of learners who had the highest levels of procedural learning ability.

Carpenter's (2008; Carpenter et al., 2009) findings suggested that declarative and procedural memory did play a role in L2 grammatical development, as predicted by the DP model. However, several questions remained open, including whether additional practice would lead to consistent relationships between development and the memory systems across all learners, and whether these relationships might also be detected, and thus further validated, through additional measures of declarative and procedural memory. The affirmation of these questions would provide more robust evidence for the claims that the declarative and procedural memory systems play distinct roles at different stages of L2 development and that individuals' varying abilities to learn within these two systems may account for the variation in L2 attainment among learners.

Motivation and research questions

The present study investigated the role of individual differences in declarative and procedural learning ability in L2 learning under implicit training conditions. Following Carpenter (2008; Carpenter et al., 2009), participants' declarative and procedural learning abilities were assessed prior to language instruction to examine whether these cognitive abilities predicted development at early and later stages of L2 acquisition. In the current study, declarative and procedural learning abilities were assessed across sub-domains of these memory systems. We investigated two types of declarative memory as indexed by the CVMT (visual) and the MLAT-V (verbal) as well as two types of procedural memory as indexed by the WPT (probabilistic) and the TOL (cognitive skill learning). Participants were trained under implicit conditions and received extended practice using the language for comprehension and production purposes.

With this design, we first wanted to establish whether a relationship between these memory systems and L2 development would be evidenced and thus posed the following research question:

RQ1a: Is there a relationship between declarative learning ability and L2 grammatical development at early and/or late stages of acquisition?

RQ1b: Is there a relationship between procedural learning ability and L2 grammatical development at early and/or late stages of acquisition?

Next, because the mere existence of a relationship between two variables is not informative as to how performance in regard to one factor accounts for performance of another, we wanted to provide a more explanatory account by also asking:

RQ2: Will declarative and procedural learning ability predict L2 grammatical development at early and late stages of acquisition?

These research questions are examined under implicit training conditions, where learners are provided with meaningful examples of the language but are not provided metalinguistic explanations about the language.

Based on Ullman's (2001, 2004, 2005) DP model for L2, we hypothesize that there will be a positive relationship between declarative learning ability and L2 grammatical development at early, but not late, stages of acquisition, whereas there will be a positive relationship between procedural memory ability and L2 grammatical development at late, but not early, stages of acquisition. Likewise, we expect that declarative and procedural memory will predict L2 grammatical development.

Methods

In order to explore the role of individual differences in declarative and procedural learning ability in L2 grammatical development, the study was designed such that participants completed one cognitive test session, four language training and practice sessions, and two assessment sessions, for a total of seven experimental test sessions.[2] The cognitive test session (Session 1) included the four measures of declarative and procedural learning ability, as well as a measure of overall intelligence. The language training and practice sessions (Sessions 2, 4, 5, and 6) consisted of exposure to and practice with an artificial L2 in a meaningful context. Finally, during the assessment sessions (Sessions 3 and 7), participants

[2] The current study was part of a larger study that aimed to examine the neural substrates involved in L2 grammatical development, as assessed by functional magnetic resonance imaging. The imaging results, however, are beyond the scope of the research questions posed by the current study, and will not be reported here.

completed judgment tasks in which they indicated the acceptability of grammatical and ungrammatical sentences from the artificial language. Each element of the study is described in detail below.

Participants

The 14 participants (six female) included in the analysis were right-handed, healthy adults between the ages of 18 and 30 ($M = 22.21$, $SD = 2.72$) and were either currently enrolled in college or held a Bachelor's degree in a non-language-related field. All participants had only one native language, which was English. Given that the artificial language had been designed to be similar to Romance languages, potential participants were prescreened and excluded from the study if they reported (a) having studied a Romance language for more than two semesters in college, three years total or within three years prior to the study, and/or (b) having been immersed in a Romance language environment for more than two weeks at any time. On average, participants had exposure to 1.21 non-native languages ($SD = 0.58$; Spanish, French, German and/or Polish) and had received an average of 3.14 years of total formal L2 instruction ($SD = 1.97$). The average age of first exposure to a language other than English was 14.15 ($SD = 2.12$). Six additional participants began the study but were excluded from analysis: Three participants failed to complete all seven study sessions; one participant was eliminated due to inattention (falling asleep) during the final language assessment; and a final two participants were eliminated because their IQ scores fell below a score of 95. This cutoff score was established based on the finding from Carpenter (2008) that no learner with an IQ under 95 was able to reach even a low level of proficiency in the artificial language Brocanto2.

Materials and procedures

Cognitive tests

During the cognitive test session (Session 1), participants' declarative and procedural memory learning ability was assessed using four separate cognitive measures. A description of each measure is provided below.

Measures of declarative learning ability

Following Carpenter (2008; Carpenter et al., 2009), participants completed verbal and nonverbal measures of declarative learning ability. The Modern Language Aptitude Test, Part V (MLAT-V; Carroll & Sapon, 1959) was used as a verbal measure of declarative learning ability. Following the standard MLAT-V procedure, the researchers directed participants to learn 24 word association pairs that consisted of pseudo-Kurdish words and their English translation equivalents. Participants studied the written word pairs for two minutes, and then completed a two-minute practice session, during which time they were able to reference the written association list while they wrote the English meanings of the pseudo-Kurdish words. After a 30-second delay, participants completed a 24-item, timed multiple-choice test (four minutes) where they chose the English equivalent for each pseudo-Kurdish word from five options from the original word list. MLAT-V scores are based on the number of correct responses selected on the multiple-choice test.

The nonverbal measure of declarative learning ability was the Continuous Visual Memory Task (CVMT; Trahan & Larrabee, 1988), a test of visual declarative memory designed to minimize reliance on verbal strategies or knowledge. Participants viewed a series of complex, abstract designs on a computer screen and then indicated whether each design was novel ("new") or had appeared previously ("old"). The "old" items consisted of seven target designs, presented seven times interspersed among 63 "new" distractor items. All items were presented in a randomized order, which was constant for all participants, with each item appearing for two seconds. During this time, participants stated aloud whether the item was "new" or "old", although they could also respond after the item had disappeared. After the participant responded, the presentation advanced to the next slide. Participants' responses were used to calculate a CVMT d' score.

Measures of procedural learning ability

Participants also completed two independent measures of procedural learning ability. The first measure of procedural memory was a computerized version of the Tower of London task (TOL; Kaller et al., 2011; Kaller et al., 2012; Unterrainer et al., 2003). In this task, participants clicked and dragged ball-like shapes on pegs, or "towers", from an initial configuration to a goal configuration. For each trial, participants viewed a new initial and goal configuration and were told to match the goal configuration in a specified number of moves ranging from three to six moves. Participants completed a practice trial followed by a set of four 3-move trials, a set of eight 4-move trials, a set of eight 5-move trials, and finally a set of eight 6-move trials. The measure of learning for the TOL task was based on participants' "initial think time", that is, the time between the presentation of a trial and the first move made by the participant on that trial, regardless of whether the trial was solved in the specified number of moves. Specifically, the proportion change in the initial think time was calculated as follows: The reaction time of the first trial was subtracted from the reaction time of the final trial and then divided by the reaction time of the first trial. The direction of the subtraction in this formula resulted in more positive values representing improved performance, which was actually a DECREASE in initial think time.

The second measure of procedural learning ability was a dual-task version of the Weather Prediction Task (WPT; Foerde et al., 2006), also used in Carpenter (2008; Carpenter et al., 2009). The WPT is an implicit, probability-based task where participants predict the weather ("sunshine" or "rain") based on patterns of four different "tarot cards" presented on the computer. Each combination of cards represents a different probability for "sunshine" or "rain". For example, a screen showing a card of squares, a card of circles, and a card of pentagons may represent a 75% chance of rain. The goal of the task is for participants to learn the probabilities represented by particular combinations of cards. Participants are typically unable to articulate the probabilities, but nevertheless tend to predict weather outcomes more accurately as they gain more experience with the task (Foerde et al., 2006). A total of 320 trials (card configurations) were divided into eight pseudo-randomized blocks, with "sunshine" and "rain" responses never occurring for more than four consecutive trials. Each trial was presented for three seconds, and participants indicated their predictions, "sunshine" or "rain", on the keyboard. After the participant's response, the correct answer was displayed on the screen. While making sunshine/rain judgments, participants performed a secondary distractor task. During each trial, one to three tones, which were either high (1000 Hz) or low (500 Hz), were pseudo-randomly presented. Participants' secondary task was to count the number of high tones that occurred during each block. This distractor task was designed to increase reliance on procedural memory and reduce the development of explicit knowledge (Foerde et al., 2006). Following Carpenter (2008; Carpenter et al., 2009), the WPT score used for analysis was based on the percent change in accuracy from block one to block eight, with responses to trials associated with a 50 percent chance of rain or sunshine excluded from the accuracy calculation.

Artificial language

In the training and practice sessions (Sessions 2, 4, 5, and 6) and in the linguistic assessment sessions (Sessions 3 and 7), participants engaged with an artificial language. The artificial language, Brocanto2 (Morgan-Short, 2007; Morgan-Short et al., 2010; Morgan-Short, Finger, Grey & Ullman, 2012; Morgan-Short, Steinhauer et al., 2012), which was modeled after Brocanto (Friederici, Steinhauer & Pfeifer, 2002), was used in order to allow participants to reach relatively high levels of proficiency in a brief period of time. The use of an artificial language also allows for strict control over factors such as type and amount of exposure to the L2 and the (dis)similarity between the L2 and learners' native language. Brocanto2 is an artificial language with a productive structure that is consistent with natural languages: Importantly and unlike what is common for artificial grammars, in Brocanto2 novel sentences can be generated, spoken and understood within a meaningful context. Previous research with Brocanto and Brocanto2 has been shown to produce brain activity characteristic of natural language processing in ERP and fMRI studies (Friederici et al., 2002; Opitz & Friederici, 2003; Morgan-Short et al., 2010; Morgan-Short, Finger et al., 2012; Morgan-Short, Steinhauer et al., 2012), suggesting that the results from the current study can be generalized to natural language learning (see Morgan-Short, Steinhauer et al., 2012 for a detailed discussion regarding the motivation and validity of using an artificial language research paradigm).

As described in Morgan-Short, Steinhauer et al. (2012), the grammar of Brocanto2 was designed to be dissimilar to English in order to avoid the possibility of L1 transfer. The lexical items, however, were designed to follow the rules of English phonotactics and to be easily pronounceable for English native speakers, in order to avoid phonological L1/L2 confounds. Brocanto2 has 13 lexical items, which are only presented auditorily during the study: four nouns (*pleck*, *neep*, *blom*, *vode*), two adjectives (*troise*/*o*, *neime*/*o*), one article (*li*/*u*), four verbs (*klin*, *nim*, *yab*, *praz*) and two adverbs (*noyka*, *zayma*). Unlike English, each noun in Brocanto2 has a formal grammatical gender designation, either masculine or feminine. Also unlike English, adjectives and articles, which both appear post-nominally, are morphologically marked to agree with the grammatical gender of the noun. Brocanto2 employs a fixed subject–object–verb word order, which differs from the subject–verb–object word order found English. Adverbs, when used, appear at the end of the sentence, immediately following the verb. All Brocanto2 words were recorded in isolation, rather than as part of a phrase or sentence, and were always presented auditorily one by one, prosodic information at the phrasal or sentential level was not available. See Table 1 for an example of a Brocanto2 sentence.

Computer-based game board moves provide a communicative context for the artificial language such that participants receive and convey information about moves on the game board using Brocanto2 sentences (Friederici et al., 2002; Morgan-Short et al., 2010; Morgan-Short, Finger et al., 2012; Morgan-Short, Steinhauer et al., 2012; see Figure 1 for an example of a game board configuration). Each Brocanto2 sentence describes a move on a computer-based board game. The four game tokens are represented by distinct symbols, which correspond to the four nouns in Brocanto2. Each token is presented within a circle or a square background (described using the adjectives). Players can move, swap, capture, and release tokens, with each of these actions corresponding to a Brocanto2 verb, as well as move them either horizontally or vertically (corresponding to Brocanto2 adverbs). Importantly, the rules of the game and the rules of the artificial language are completely independent of each other. To ensure that learning the game was not

Table 1. *Example grammatical and ungrammatical Brocanto2 sentences.*

Sentence type	Brocanto2 stimuli					
Grammatical sentence	*Blom*	*neimo*	*lu*	*neep*	*li*	*praz*
	Blom-piece	square	the	neep-piece	the	switch
	"The square blom-piece switches with the neep-piece."					
Ungrammatical sentence with word order violation	*Blom*	**nim*	*lu*	*neep*	*li*	*praz*
	Blom-piece	capture	the	neep-piece	the	switch
	"The *capture blom-piece switches with the neep-piece."					

* = violation

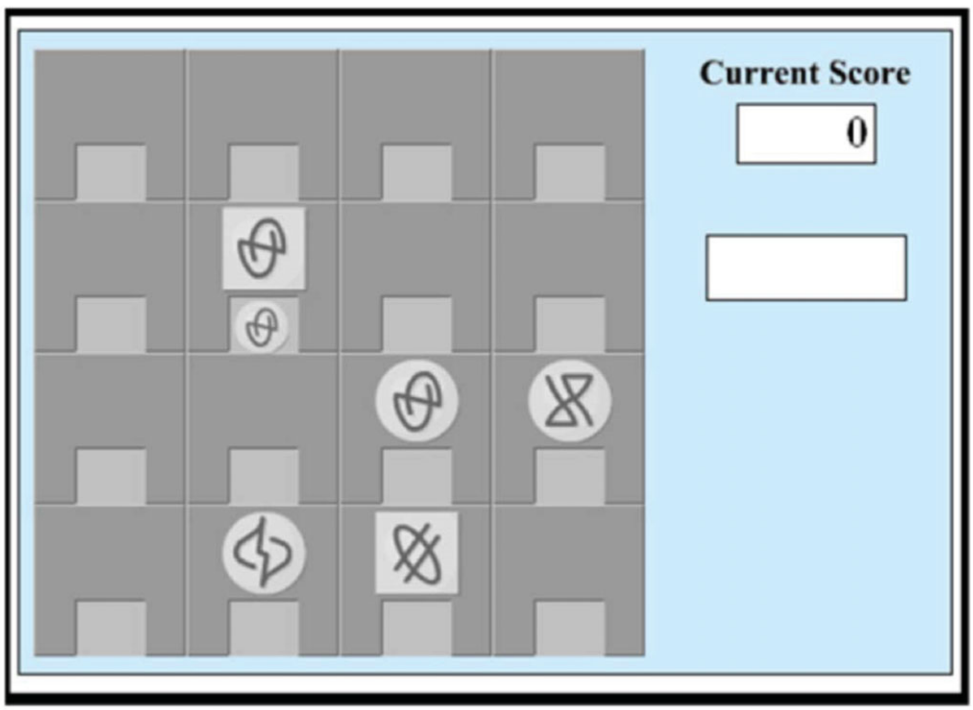

Figure 1. (Colour online) Computer-based game board.

confounded with learning the language, all participants were trained to play the game before they engaged in learning the language.

Vocabulary training

In each training and practice session (Sessions 2, 4, 5, and 6), participants were required to demonstrate mastery of all Brocanto2 lexical items before proceeding with grammar training and practice. To this end, at the beginning of each of these sessions, participants completed a computer-based vocabulary training program. During vocabulary training, each Brocanto2 lexical item was presented auditorily, accompanied by a visual symbol that represented its meaning. Participants listened to each item and viewed the corresponding symbol at their own pace. When they believed that they had learned all the lexical items, a vocabulary test was administered that asked participants to state out loud the lexical item that corresponded to a particular symbol. During the vocabulary test, each symbol representing a lexical item was presented twice at maximally distant points in the test. If participants did not achieve a score of 100% accuracy on this test, they listened to and viewed the presentation of the lexical items and took the test again until they evidenced complete mastery of all lexical items.

Language training

In each training and practice session (Sessions 2, 4, 5, and 6), after having successfully demonstrated mastery of the lexical items, participants received training with the full language. Learners were informed that they would "receive training on an artificial language" and were asked to "listen carefully to the information that is given". Full language training, which lasted approximately 13.5

https://doi.org/10.1017/S1366728912000715 Published online by Cambridge University Press

minutes, occurred under an implicit condition, in which no explicit grammar rules or explanations regarding any aspect of the language were provided. Instead, learners were exposed to 129 auditorily presented meaningful phrases and sentences of the language. As each phrase or sentence was presented, participants also viewed the corresponding game token or move. This implicit training condition had been designed to approximate learning a language under a natural (immersion) setting, where learners are exposed to phrases and sentences but do not generally receive grammatical explanation about the language (Morgan-Short et al., 2010; Morgan-Short, Steinhauer et al., 2012). Note that the training condition was not designed to be reflection of an incidental training condition, where learning occurred in absence of the intention to learn, as participants were told that they would be learning an artificial language. See Appendix A for examples of stimuli included in language training.

Language practice

The final element of each training and practice session (Sessions 2, 4, 5, and 6) was practice with the artificial language, which occurred in the context of the computer-based game. Learners were informed that they would "use the artificial language to play the computer-based game". Practice consisted of 72 modules (20 stimuli per module) of comprehension and production practice. During comprehension modules, participants heard sentences in the language and were instructed to "make the move on the game board that corresponds to the statement you heard". During production modules, participants saw a move and were instructed to describe it by using a Brocanto2 sentence to "state the move out loud". For both comprehension and production practice, participants received feedback on whether their response was correct or incorrect. No additional information or opportunity to modify the response was provided. None of the sentences presented in practice were presented during language training or other practice sessions. Participants completed 12 practice modules during Session 2 and 20 practice modules in each of the subsequent training and practice sessions, Sessions 4, 5, and 6.

Linguistic assessment

After the initial and final training and practice sessions (e.g., Sessions 2 and 6), participants completed linguistic assessment sessions (Sessions 3 and 7). Linguistic development was assessed using a computer-administered grammaticality judgment task (GJT). Participants judged the acceptability of 120 novel Brocanto2 sentences, consisting of 60 grammatical and 60 ungrammatical sentences. The ungrammatical sentences, each derived from a grammatical sentence, contained a word order violation that had been created by replacing a word from one of the five word categories (e.g., noun) with a word from a different category (e.g., adjective, article, verb, adverb). Violations never occurred on the first word of the sentence and were distributed among each word category as equally as possible. However, because replacing an adjective by an article would not have created a word order violation, adjectives were replaced by the other categories slightly more often. The stimuli were divided into two experimental blocks, with ungrammatical and their matched grammatical sentences occurring in different blocks, and were pseudo-randomly intermixed so that there were never more than three consecutive grammatical or ungrammatical sentences.

During the administration of the GJT, participants were instructed to judge the acceptability of each sentence ("good" or "bad") using buttons on a response box. Stimuli presentation began with a fixation cross "+" that was displayed for 500 ms prior to sentence presentation and remained on the screen for the duration of the sentence. Sentence stimuli were presented auditorily at a rate of 800 ms per word. A prompt asking for an acceptability judgment ("Good?") appeared for 2000 ms after the final word of each sentence, during which time participants provided a response. Participants had up to five seconds to respond. Responses were recorded and accuracy and d' scores, which account for potential response bias, were calculated for each participant.

Summary of general procedure

Participants were scheduled for seven experimental sessions over a two-week period with one to three nights between each session. During Session 1, the cognitive assessment session, participants first provided informed consent, filled out background questionnaires, and completed an IQ assessment (The Kaufman Brief Intelligence Test, 2nd edition; Kaufman & Kaufman, 2004). Participants subsequently completed the declarative and procedural learning tasks. This session lasted approximately three hours, and the order of the cognitive tests was counterbalanced across participants. The remaining sessions were devoted to language training and practice (Sessions 2, 4, 5, and 6; average duration per session was 2.6 hours) and assessment (Sessions 3 and 7; average duration per session was 1 hour), as described above. Participants were compensated for time spent completing the study.

Results

Before examining the relationship between L2 grammatical development and cognitive factors, we first quantified the level of attainment of L2 grammatical development that occurred during the study as measured by participants' abilities to distinguish between grammatical and ungrammatical sentences on the judgment tasks

https://doi.org/10.1017/S1366728912000715 Published online by Cambridge University Press

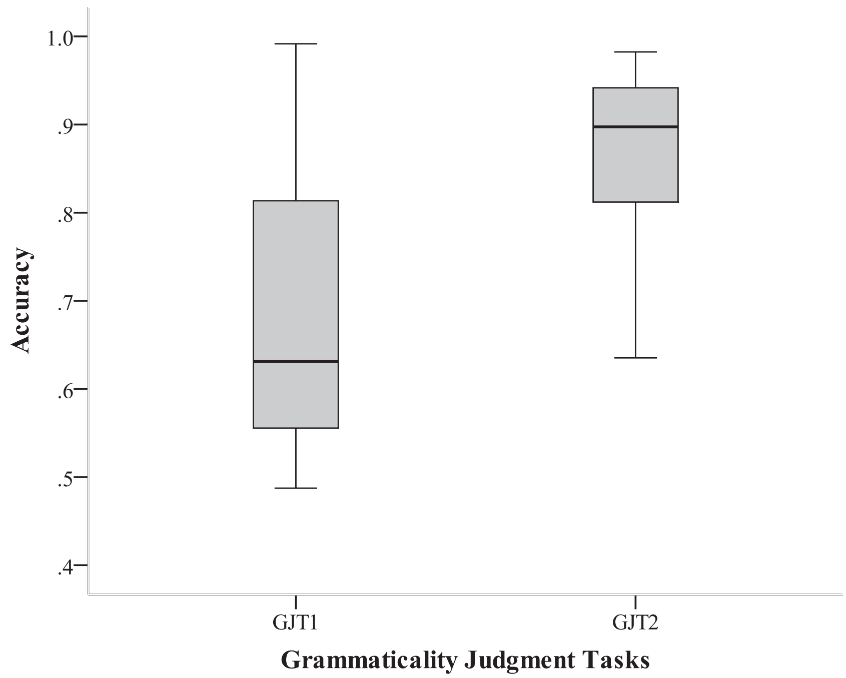

Figure 2. Box plot representation of variation in performance on the GJT1 and GJT2. Median scores and interquartile and overall ranges for accuracy are displayed.

Table 2. *Means (Ms) and standard deviations (SDs) for GJT1 and GJT2 scores.*

	Accuracy		d'		d'log10	
Assessment	*M*	*SD*	*M*	*SD*	*M*	*SD*
GJT1	.690	.161	1.255	1.367	.292	.235
GJT2	.863	.111	2.619	1.072	.537	.148

GJT1 = Grammaticality Judgment Task at Session 3; GJT2 = Grammaticality Judgment Task at Session 7

administered during the assessment sessions (GJT1 and GJT2). Prior to performing analyses, we examined the d' scores for normal distribution by inspecting the skewness and kurtosis ratios for both GJT1 and GJT2, and found that the skewness ratio on GJT1 was greater than an accepted ratio of 2.0, suggesting a non-normal distribution (Weinberg & Abramowitz, 2002). Therefore we performed a log transformation on the GJT d' scores, which resulted in data that exhibited a normal distribution (skewness and kurtosis ratios < 2.0). We then examined the log transformed data for outliers and found that no participant score fell 3 *SD*s above or below the group means. The complete set of log transformed d' data was used for all subsequent data analyses.

Examination of the GJT1 and GJT2 scores suggested that participants did indeed learn the artificial language over the course of the study (see Table 2 for accuracy, d', and log transformed d' scores). A paired-samples *t*-test revealed that performance significantly differed between the first and second GJT ($t(13) = -4.658$, $p < .001$). Participants performed significantly better on GJT2 than on GJT1, demonstrating development in the artificial language over time. Note, however, that performance accuracy ranged from around chance to near perfect scores on GJT1 and from .64 to near perfect scores on GJT2, indicating that performance varied widely among participants at both assessment sessions (see Figure 2).[3]

To address the first research question, WHETHER THERE IS A RELATIONSHIP BETWEEN DECLARATIVE AND PROCEDURAL LEARNING ABILITIES AND L2 GRAMMATICAL (WORD ORDER) DEVELOPMENT AT EARLY AND/OR LATE STAGES OF ACQUISITION, we examined L2 grammatical development as evidenced on GJT1 and GJT2 in relation to individual performance on measures of declarative and procedural memory. As with the GJT scores, before performing analyses we checked whether the data from these measures fell along a normal distribution by examining the skewness and kurtosis ratios, and found that these ratios were all less than 2.0. Furthermore, there were no apparent outliers as all

[3] Although participants had relatively limited experience with previous L2 learning, especially with Romance L2s, it is still possible that participants with more language learning experience would reach higher levels of development in the current study. However, correlational analyses between measures of language learning experience, i.e., number of L2s, age of exposure to first L2, and total number of years of formal exposure, and performance on GJT1 and GJT2 revealed no significant relationships (all *p*s > .05).

https://doi.org/10.1017/S1366728912000715 Published online by Cambridge University Press

Table 3. *Means (Ms) and standard deviations (SDs) for assessments of declarative and procedural learning abilities.*

Assessment	*M*	*SD*	Range or unit of measurement
MLAT	18.360	4.877	0–24
CVMT	2.119	0.329	*d′*
TOL	−0.164	0.351	ms
WPT	0.574	0.178	0–1

MLAT-V = Modern Language Aptitude Test, Part V; CVMT = Continuous Visual Memory Task; TOL = Tower of London task; WPT = Weather Prediction Task

Table 4. *Correlations between GJT1 and GJT2 and assessments of declarative and procedural learning ability.*

Variable	MLAT-V	CVMT	TOL	WPT
GJT1	.620*	.497ˆ	.187	.281
GJT2	−.053	.420	.588*	.557*
MLAT-V	–	.149	−.172	.015
CVMT		–	.270	.481ˆ
TOL			–	.557*
WPT				–

GJT1 = Grammaticality Judgment Task at Session 3; GJT2 = Grammaticality Judgment Task at Session 7; MLAT-V = Modern Language Aptitude Test, Part V; CVMT = Continuous Visual Memory Task; TOL = Tower of London task; WPT = Weather Prediction Task
ˆ $p < .10$; * $p < .05$

participant means fell within 3 *SD*s of the group means. See Table 3 for descriptive statistics for the cognitive tests (MLAT-V, CVMT, TOL, WPT).

Pearson's correlations (two-tailed) were run in order to reveal any relationships between performance on GJT1 and GJT2 and assessments of declarative and procedural learning abilities (see Table 4). For GJT1, the analysis revealed a positive relationship with assessments of declarative learning ability, which was statistically significant for the MLAT-V and trended towards significance for the CVMT. In contrast, no significant relationship was found for assessments of procedural learning ability, either for the TOL or the WPT. For GJT2, the analysis yielded the opposite pattern. There was a positive relationship with assessments of procedural learning ability, which was statistically significant for both the TOL and the WPT, but a lack of a significant relationship with assessments of declarative learning ability, for both the MLAT-V and the CVMT. These results provide evidence for a relationship between L2 development and declarative, but not procedural, learning ability at an early stage of L2 acquisition and between L2 development and procedural, but not declarative, learning ability at a later stage.

Given these significant relationships, we performed regression analyses in order to address the second research question: Will declarative and procedural memory predict L2 grammatical development at early and late stages of acquisition? Before running regression analyses, we calculated standardized composite scores for declarative and procedural learning ability in order to resolve issues of collinearity between variables (e.g., between the TOL and WPT) and to simplify the regression given the relatively low number of participants. For declarative learning ability, participants' scores on the MLAT-V and CMVT were first transformed into standard scores. Standard composite scores were then calculated by dividing the sum of the standard scores on the MLAT-V and CVMT by the standard deviation of the combined variances and covariances (Crocker & Algina, 1986, p. 95). For procedural learning ability, the same procedure was followed using scores from the TOL and WPT. The standard composite scores for declarative and procedural learning ability, heretofore referred to as the declarative and procedural composite scores, followed the same correlational pattern as the scores on the individual measures: The declarative composite scores showed a significant positive correlation with GJT1 but not with GJT2 ($r = .737$, $p = .001$ and $r = .272$, $p = .184$, respectively), and the procedural composite scores showed a significant positive correlation with GJT2 but not with GJT1 ($r = .675$, $p = .006$ and $r = .208$, $p = .248$, respectively).

In order to determine the amount of variance in L2 grammatical development explained by declarative and procedural learning ability, the *d′* scores from GJT1 and GJT2 (dependent variables) were regressed onto the composite scores representing both types of learning abilities (predictor variables; see Table 5). For GJT1, the regression model accounted for approximately 56% of the variance in the *d′* scores, which is a large effect ($f^2 = 1.252$; Cohen, 1992). The declarative composite score, but not the procedural composite score, served as a significant predictor of the ability to distinguish between grammatical and ungrammatical sentences. For GJT2, the regression model accounted for approximately 46% of the variance in the *d′* scores, which is also considered to be large effect ($f^2 = 0.848$; Cohen, 1992). The procedural composite score, but not the declarative composite score, served as a significant predictor of the ability to distinguish between grammatical and ungrammatical sentences. These results indicate that declarative and procedural learning abilities contributed to L2 grammatical development differentially at early and late stages of L2 acquisition.

Discussion

We examined whether individual differences in declarative and procedural memory could account for

https://doi.org/10.1017/S1366728912000715 Published online by Cambridge University Press

Table 5. *Regression models examining declarative and procedural learning ability and performance on GJT1 and GJT2.*

	GJT1			GJT2		
Variable	*B*	*SEB*	β	*B*	*SEB*	β
Constant	.292	.047		.529	.934	
Declarative	.198	.060	.711**	.011	.042	.065
Procedural	.028	.051	.118	.099	.036	.661*
R^2		.556			.459	
F		6.273*			4.247*	

Declarative = standardized composite score for declarative learning ability; Procedural = standardized composite score for procedural learning ability; GJT1 = Grammaticality Judgment Task at Session 3; GJT2 = Grammaticality Judgment Task at Session 7; *B* = unstandardized regression coefficient; *SEB* = standard error of *B*; β = standardized regression coefficient
* $p < .05$; ** $p < .01$

variation in adult L2 learning under implicit training conditions. First, the study addressed whether there were relationships between declarative and procedural learning ability and L2 grammatical development, as measured by participants' ability to distinguish between ungrammatical sentences containing a syntactic word order violation and matched grammatical sentences, at early and late stages of acquisition. At early stages of acquisition, the results revealed a positive correlation between L2 grammatical development with declarative, but not procedural, learning ability. This positive relationship was found to be significant for the MLAT-V, was trending for the CVMT, and was significant for a composite declarative learning ability score. In contrast, at late stages of acquisition, declarative learning ability no longer correlated with L2 grammatical development, but procedural learning ability did, as evidenced by significant positive correlations for the WPT and the TOL, as well as for the composite procedural learning ability score.

Second, the study explored how well declarative and procedural learning abilities predict L2 grammatical development at early and late stages of acquisition. The results from the regression analyses indicated that individual abilities in these memory systems explained variation in L2 grammatical development. Effect size calculations showed that the effect of these memory systems on L2 grammatical development was large. In addition, the analyses indicated that declarative and procedural learning abilities predicted L2 grammatical development differentially at early and late stages of acquisition: At an early stage of acquisition, only declarative memory served as a unique predictor of participants' ability to distinguish between grammatical and ungrammatical sentences, whereas at a late stage of acquisition, only procedural memory uniquely predicted this ability.

Thus, the results from the current study indicate that declarative and procedural memory can account for a relatively large amount of variance in L2 grammatical development at early and late stages of acquisition, at least when learning occurs under implicit training conditions. However, the role of these two memory systems on L2 grammatical development appears to differ at different stages of acquisition. What might explain the differential role of the memory systems at different stages of acquisition? Recall that learning with the declarative memory system can occur quickly and after only one instance of exposure but that learning supported by the procedural memory system proceeds gradually with repeated exposure (Knowlton & Moody, 2008). Also recall that the first judgment task occurred after learners had received 13.5 minutes of implicit training and had completed 12 modules of comprehension and production practice, whereas the second and final judgment task occurred after learners had received implicit training four times and had completed 72 practice modules. For learners with strong declarative learning ability, the relatively minimal exposure before the first judgment task may have been sufficient to allow them to form declarative memory about the language, such as hypotheses about the way the language worked or knowledge of patterns of co-occurrence among the words. This knowledge may have then supported their performance on the first judgment task leading to the finding that declarative learning ability predicted performance at the early stage of development. During this early stage, procedural memory may have also been engaged, but because procedural learning occurs gradually with repeated exposure, more training and exposure may have been needed before memory based on this system was strong enough to exert a detectable role, at least for the assessment used in this study. At the second judgment task, learners had six times the amount of training and practice as at the first judgment task. This increased experience may have been sufficient for gradual but substantial development of procedural memory. The strength of procedural memory may have then interfered with the application of declarative knowledge leading to the finding that procedural learning ability predicted performance later in development. This see-saw effect between declarative memory and procedural memory is one of the ways that declarative and procedural memory are expected to interact (Knowlton & Moody, 2008; Ullman, 2001, 2004).

These findings are largely consistent with the only known study that previously addressed this question, Carpenter (2008; Carpenter et al., 2009), although there are some differences among the results. First, our results largely confirm and extend the findings in Carpenter's dissertation study in regard to procedural memory. Both studies found a relationship between procedural memory and performance at higher levels of proficiency for implicitly trained learners. Recall, however, that Carpenter found a parabolic relationship with procedural memory and L2 grammatical development, with only a subset of learners (those with either higher or lower procedural memory scores as assessed by the WPT) showing increased L2 development. Carpenter suggested that learners with high procedural memory scores may have been able to rely on procedural memory to promote successful language development and that learners with low procedural memory scores many have been able to recruit other mechanisms to successfully acquire L2 grammar. For learners with mid-range procedural memory scores, procedural memory may have been strong enough to interfere with the recruitment of other mechanisms or strategies but not strong enough to lead to successful development.

Although the current study also found a significant relationship between procedural memory and development, the relationship was a (positive) linear relationship, rather than a parabolic relationship. This difference between the two studies might be explained by the fact that at the late stage of learning, learners in the current study had more experience with the artificial language than those in Carpenter (2008; Carpenter et al., 2009). Note that the second judgment task in Carpenter was administered after two training sessions (with 129 meaningful examples provided during the first session and repeated in the second session) and 44 practice modules (880 unique practice items in total) whereas the second judgment task in the current study was administered after four training sessions (with 129 meaningful examples provided during the first session and repeated in the third, fourth and fifth sessions) and 72 practice modules (1440 unique practice items in total). Learners' increased experience with the L2 may have led to greater engagement of procedural memory, with the effect being that those with lower procedural memory scores could not avoid an augmented effect of procedural memory, which was not as effective for them as other strategies might have been. It is generally understood that performance on procedural tasks may rely on non-procedural mechanisms when the opportunity to develop procedural knowledge has not been sufficient, but that after extensive practice, procedural knowledge becomes stronger and may interfere with the ability to apply other types of knowledge (Knowlton & Moody, 2008). For participants with mid-to-high-range procedural learning ability, procedural memory would have been strong enough to support development, especially in learners with high procedural learning ability. This explanation may reasonably account for the positive linear relationship between procedural learning ability and L2 grammatical development across all learners in the current study.

The current study also found that the relationship between procedural learning ability and L2 grammatical development was evidenced across two different measures of procedural learning ability: Not only was development correlated with performance on a probabilistic classification task (the WPT), but it was also found for a task reflecting cognitive skill acquisition (the TOL). More specifically, the correlation for the TOL was based on the proportion change in the initial think time as participants progressed through the problem solving task. Note that it was the DECREASE in initial think time (here reflected by a positive score) on the TOL that was correlated with higher performance on the GJT2. The fact that two independent measures of procedural memory are related to L2 grammatical development at late stages of L2 acquisition suggests a fairly robust relationship between procedural memory and performance at late stages of acquisition.

In regard to declarative memory, the findings from the current study were not consistent with the findings from Carpenter (2008; Carpenter et al., 2009). In the current study, greater declarative learning ability was correlated with greater L2 grammatical development on the first judgment task, but this was not the case for the implicitly trained learners in Carpenter's study. Again, the difference in the amount of experience with the L2 at the time of the linguistic assessment may explain the inconsistent results in the two studies. In Carpenter, learners completed GJT1 only after reaching a low-proficiency benchmark during practice. Thus, the amount of practice completed before GJT1 varied among participants, with some learners having more experience with the language than others prior to completing the first linguistic assessment. This may have led to the recruitment of both declarative and nondeclarative memory systems and the lack of a detectable relationship between declarative memory and L2 grammatical development in that study. In the current study, however, all learners completed exactly twelve practice modules before the first linguistic assessment. In other words, learners did not have the opportunity for extended practice, and thus we may have better captured their performance while they still relied more heavily on declarative memory. Consequently, individuals with weaker declarative learning abilities would not have been able to perform as well on the first linguistic assessment as learners with stronger declarative learning abilities, as was the case for GJT1 in the current study. If this interpretation of the two studies is valid, it suggests that declarative memory may play a larger role when the amount of experience with the L2 is limited rather than when L2 proficiency is low (but the amount of experience is larger), at least for implicitly trained learners.

The findings from the current study provide evidence consistent with predictions made by theoretical perspectives of adult L2 acquisition that posit a role for declarative and procedural memory in development (e.g., DeKeyser, 1995, 2007; Paradis, 1994, 2004, 2009; Ullman, 2001, 2004, 2005). In particular, the results appear to support predictions that may be made by the DP model (Ullman, 2001, 2004, 2005). The finding that declarative, but not procedural, memory is a significant predictor of development at early stages of acquisition is in line with the claim that L2 grammar initially relies on declarative memory. In addition, the finding that procedural, but not declarative, memory is a significant predictor of development at high proficiency supports the claim that grammatical aspects of L2 may come to depend on procedural memory at later stages of development. Therefore, the results of the present study are highly consistent with DP model predictions, at least in regard to implicitly trained adult L2 learners' ability to distinguish between grammatical and ungrammatical word order stimuli.

The results of the study might also be viewed as largely consistent with Paradis' neurocognitive perspective of L2 acquisition (Paradis, 1994, 2004, 2009) and with DeKeyser's skill acquisition perspective (DeKeyser, 1995, 2007). At the very least the study does not produce any evidence that is highly inconsistent with these perspectives. However, the finding that procedural memory predicts L2 attainment at late stages of development might be somewhat surprising under Paradis' perspective given that Paradis claims that L2 development comes to rely on procedural memory only in rare cases. In regard to skill acquisition theory, it would be interesting to delve into the question about how procedural learning ability contributes to the purportedly independent stages of proceduralization and automization, or even if procedural learning ability, as assessed by the current study, maps onto the mechanisms posited to underlie proceduralization of declarative knowledge. These issues, which are more fine-grained than those examined in the current study, should be considered in future research.

More broadly, the findings from this study suggest that domain-general cognitive abilities, particularly declarative and procedural memory, play important roles in adult L2 grammatical development and that such abilities serve as individual difference factors that account for a substantial amount of variation in adult-learned L2. This conclusion is consistent with domain-general accounts of L2 acquisition but does not negate a possible contribution by domain-specific mechanisms, as a portion of variance in L2 grammatical development remains to be explained. Indeed, it may be reasonable to expect that specific linguistic processes remain at play in adult L2 acquisition, but that the engagement of these processes is mediated by domain-general processes. On a similar note, non-cognitive and non-linguistic factors, such as motivation and sociocultural factors, certainly also play a role. It is well-known that a diverse range of factors including cognitive, linguistic, affective, and sociocultural factors must be understood before our knowledge of L2 grammatical development in adults is more complete (Sanz, 2005). Regardless of the potential role of other factors, the results from the current study provide evidence that cognitive factors, specifically declarative and procedural memory, contribute to L2 grammatical development, but do so differentially at different stages of development.

Limitations and directions for future research

The findings from the current study have provided new evidence and insight into the issue of individual differences in L2 acquisition. As is the case for any study, however, its findings should be considered in light of its limitations. First, perhaps the most significant consideration is that the study examined the relationship between declarative and procedural memory and L2 grammatical development only in implicitly trained learners, who were not provided with any metalinguistic information about the L2. It may be the case that the relationship between these memory systems and L2 grammatical development plays out differently when learners are exposed to an L2 under different conditions, such as more explicit conditions, where metalinguistic information or direction to look for grammatical rules is provided. Although many adults do learn L2 in immersion settings, where grammatical information about the language may not be available, clearly a great number of adults learn L2s in classroom-based settings, where grammatical rule explanation is often standard. Thus, it is important to further examine the contributions of declarative and procedural memory, and other individual difference factors, under different training conditions (Carpenter, 2008; Carpenter et al., 2009; Robinson, 2002; Snow, 1991) that are representative of the different contexts under which adult learners are exposed to L2s.

Indeed, differences would most likely be expected. Note that Carpenter (2008; Carpenter et al., 2009) found a different set of relationships between declarative and procedural memory and development in learners who had received explicit training. These same learners also showed different patterns of neurocognitive processing at early and late stages of acquisition and at retention testing (Morgan-Short, Finger et al., 2012; Morgan-Short, Steinhauer et al., 2012). Furthermore, research on procedural learning more generally has shown that performance comes to rely on different neural substrates depending on the learning condition (Foerde et al., 2006). If different conditions lead learners to rely on different neural processes and substrates, then it is likely that individual differences in declarative and procedural memory systems exert an influence differentially under these different conditions. Thus it will be important for future research to examine the role that different types of memory may play under different conditions. Finally, it is important to point out that although the training condition in this study was an implicit CONDITION, this does not imply that the LEARNING itself was implicit. It would be interesting to delve into the role of declarative and procedural memory for implicit L2 learning.

Second, although the present study has focused on two particular long-term memory systems that have been posited to play a role in L2 development, it would be interesting to also examine the role that working memory may have in relation to declarative and procedural memory and L2 development. Working memory has received a significant amount of attention in recent L2 research, which has found evidence of a positive relationship between working memory and development, at least when conscious, intentional, explicit learning processes are involved (see Williams, 2011, for a recent review). Interestingly, working memory may interact differentially with declarative and procedural memory, being more beneficial for declarative memory (Knowlton & Moody, 2008). Thus one might expect that, at least under implicit training conditions, L2 learners with a smaller working memory might be less capable of utilizing declarative memory as a basis for L2 development and might rely more heavily on procedural memory. The prediction could be made that, at late stages of acquisition, learners with lower working memory might outperform those with higher working memory. Of course, the prediction might be quite different for explicit conditions. Thus, it might prove especially fruitful for future research to examine a potential mediating role of working memory in regard to the declarative and procedural memory systems and L2 development.

Third, it is important to point out that this study examined performance on just one type of linguistic structure, syntactic word order, and on only one type of linguistic assessment, a judgment task. The role of individual differences in declarative and procedural learning ability, however, may be moderated by the linguistic structures being tested, as well as by the assessment of L2 grammatical development being used. Ideally, future research will examine the relationship among individual differences in declarative and procedural memory, training condition, and linguistic structure using diverse measures of L2 grammatical development. Fourth, replication of the pattern of results described here will further validate the findings. In particular, replication in the context of longitudinal development of a natural language would substantiate the ecological validity of investigating this issue under an artificial language paradigm.

Fifth, it will be important for future studies to include larger samples of participants in order to gain further insight into how different cognitive measures account for L2 development and the nature of the interrelationship between the cognitive measures. Note that, in the current study, the correlation between the two measures of declarative memory, the MLAT-V and the CVMT, is low. This is not expected to be problematic because it is understood that different aspects of declarative memory can be dissociated from each other, e.g., remembering and knowing (Knowlton & Squire, 1995; Squire & Knowlton, 2000). It is somewhat surprising that the CVMT and the WPT seem to pattern together, although the statistical relationship between them does not reach significance. This potential relationship between these tasks may be due to their shared processing of complex visual information. These potential interrelationships do not lessen the validity of the tasks used in the current study as measures of declarative and procedural memory, but understanding any interrelationship between them should be informative to any account of the contribution of different types of memory to L2 acquisition.

Finally, the use of neuroimaging techniques, such as those that measure the electrophysiological or hemodynamic responses of the brain to linguistic stimuli, may provide further evidence about L2 learners' reliance on different neural-based memory systems and whether this reliance is subject to individual differences (Carpenter, 2008; Morgan-Short et al., 2010; Morgan-Short, Finger et al., 2012; Morgan-Short, Steinhauer et al., 2012).

Conclusion

The current study examined individual differences in declarative and procedural learning ability as potential factors that could account for the wide range of L2 attainment evidenced by adult learners. The results from the study showed that, in a group of implicitly trained learners of an artificial language, declarative memory was associated with syntactic word order development at early stages of acquisition and that procedural memory was associated with development at later stages of acquisition. Moreover, these cognitive abilities accounted for a large amount of variance in L2 grammatical development at both early and late stages, with declarative memory serving as a unique predictor at an early stage and procedural memory serving as a unique predictor at a late stage. The findings are consistent with theoretical perspectives of L2 grammatical development that posit a role for these memory systems and confirm the hypothesis that differences in declarative and procedural memory learning abilities are predictive of individual L2 grammatical development, at least for implicitly trained learners.

Appendix A. Example section from implicit training condition

The text below provides a sample of meaningful sentences related to aspects of Brocanto2 word order that were presented to participants. All examples were auditorily presented together with visually presented corresponding

https://doi.org/10.1017/S1366728912000715 Published online by Cambridge University Press

game constellations. Note: " . . . " indicates that additional examples were provided.

pleck li vode lu praz
vode lu neep li praz
blom lu pleck li praz
neep li blom lu praz
. . .
vode lu nim
vode lu neep li nim
pleck li neime li nim
pleck li blom lu nim
. . .
neep li yab
blom lu pleck li yab
blom lu pleck li yab
blom lu yab
. . .

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning, 59,* 249–306.

Bialystok, E., & Frölich, M. (1978). Variables of classroom achievement in second language learning. *The Modern Language Journal, 62,* 327–336.

Bowden, H. W., Gelfand, M. P., Sanz, C., & Ullman, M. T. (2010). Verbal inflectional morphology in L1 and L2 Spanish: A frequency effects study examining storage versus composition. *Language Learning, 60,* 44–87.

Brovetto, C., & Ullman, M. T. (2005). The mental representation and processing of Spanish verbal morphology. In D. Eddington (ed.), *Selected proceedings of the Seventh Hispanic Linguistics Symposium*, pp. 98–105. Somerville, MA: Cascadilla Proceedings Press.

Carpenter, H. S. (2008). A behavioral and electrophysiological investigation of different aptitudes for L2 grammar in learners equated for proficiency level. Ph.D. dissertation, Georgetown University.

Carpenter, H., Morgan-Short, K., & Ullman, M. T. (2009). Predicting L2 using declarative and procedural memory assessments: A behavioral and ERP investigation. Presented at the Georgetown University Round Table, Washington, DC.

Carroll, J. B. (1958). A factor analysis of two foreign language aptitude batteries. *Journal of General Psychology, 59,* 3–19.

Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (ed.), *Training research and education*, pp. 87–136. Pittsburgh, PA: University of Pittsburgh Press.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (ed.), *Individual differences and universals in foreign language aptitude*, pp. 83–118. Rowley, MA: Newbury House.

Carroll, J. B., & Sapon, S. M. (1959). *Modern Language Aptitude Test*. New York: The Psychological Corporation/Harcourt Brace Jovanovich.

Chandrasekaran, B., Kraus, N., & Wong, P. C. M. (2012). Human inferior colliculus activity relates to individual differences in spoken language learning. *Journal of Neurophysiology, 107,* 1325–1336.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern Test Theory*. Fort Worth, TX: Holt, Rinehart and Winston.

DeKeyser, R. M. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17 (3),* 379–410.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22,* 499–533.

DeKeyser, R. [M.] (2007). Skill acquisition theory. In B. VanPatten & J. Williams (eds.), *Theories in second language acquisition: An introduction*, pp. 97–113. Mahwah, NJ: Lawrence Erlbaum.

Dorfberger, S., Adi-Japha, E., & Karni, A. (2007). Reduced susceptibility to interference in the consolidation of motor memory before adolescence. *PLoS ONE, 2,* 1–6.

Dörnyei, Z. (2005). *The psychology of the language learner: individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Ehrman, M., & Oxford, R. (1995). Cognition plus: Correlates of language learning success. *Modern Language Journal, 79,* 67–89.

Eichenbaum, H. (2002). *The cognitive neuroscience of memory: An introduction*. New York: Oxford University Press.

Eichenbaum, H., & Cohen, N. J. (2001). *From conditioning to conscious recollection: Memory systems of the brain*. New York: Oxford University Press.

Ettlinger, M., Bradlow, A., & Wong, P. C. M. (in press). Variability in the learning of complex morphophonology. *Applied Psycholinguistics*.

Ferman, S., Olshtain, E., Schechtman, E., & Karni, A. (2009). The acquisition of a linguistic skill by adults: Procedural and declarative memory interact in the learning of artificial morphological rule. *Journal of Neurolinguistics, 22,* 384–412.

Foerde, K., Knowlton, B. J., & Poldrack, R. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences, 103,* 11778–11783.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences, 99,* 529–534.

Harley, B., & Hart, D. (1997). Language aptitude and language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition, 19,* 379–400.

Harley, B., & Hart, D. (2002). Age, aptitude, and second language learning on a bilingual exchange. In Robinson (ed.), pp. 301–330.

Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70,* 125–132.

Kaller, C. P., Rahm, B., Köstering, L., & Unterrainer, J. M. (2011). Reviewing the impact of problem structure on planning: A software tool for analyzing tower tasks. *Behavioural Brain Research, 216,* 1–8.

Kaller, C. P., Unterrainer, J. M., & Stahl, C. (2012). Assessing planning ability with the Tower of London task: Psychometric properties of a structurally balanced problem set. *Psychological Assessment, 24,* 46–53.

Kaufman, A. S., & Kaufman, N. L. (2004). *The Kaufman Brief Intelligence Test, Adult Version, Second Edition (K-BIT-2)* (2nd edn.). Circle Pines, MN: American Guidance Service.

Knowlton, B. J., & Moody, T. D. (2008). Procedural learning in humans. In. J. H. Bryne (ed.), *Learning and memory: A Comprehensive reference* (vol. 3): *Memory systems*, pp. 321–340. Oxford: Academic Press/Elsevier.

Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: Two different expressions of declarative memory. *Journal of Experimental Psychology, 21 (3),* 699–710.

Knowlton, B., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification in amnesia. *Learning and Memory, 1,* 106–120.

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

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

Morgan-Short, K. (2007). A neurolinguistic investigation of late-learned second language knowledge: The effects of explicit and implicit conditions. Ph.D. dissertation, Georgetown University.

Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. T. (2012). Second language processing shows increased native-like neural responses after months of no exposure. *PLoS ONE, 7*: e32974.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning, 60,* 154–193.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience, 24,* 933–947.

Ohlsson, S. (2008). Computational models of skill acquisition. In R. Sun (ed.), *The Cambridge handbook of computational psychology*, pp. 359–395. New York: Cambridge University Press.

Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language–like rules. *NeuroImage, 19,* 1730–1737.

Ouellet, M. C., Beauchamp, M. H., Owen, A. M., & Doyon, J. (2004). Acquiring a cognitive skill with a new repeating version of the Tower of London task. *Canadian Journal of Experimental Psychology, 58,* 272–288.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. C. Ellis (ed.), *Implicit and explicit learning of languages*, pp. 393–419. London: Academic Press.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.

Paradis, M. (2009). *Declarative and procedural determinants of second languages* (vol. 40). Amsterdam, Netherlands: John Benjamins.

Perrachione, T., Lee, J., Ha, L., & Wong, P. C. M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *Journal of the Acoustical Society of America, 130,* 461–472.

Robinson, P. (ed.). (2002). *Individual differences and instructed language learning*. Philadelphia, PA: John Benjamins.

Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M. 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 leaning: 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.

Ross, S., Yoshinaga, N., & Sasaki, M. (2002). Aptitude-exposure interaction effects on *wh*-movement violation detection by pre-and-post critical period Japanese bilinguals. In Robinson (ed.), pp. 266–299.

Sanz, C. (ed.). (2005). *Mind and context in adult second language acquisition: Methods, theory, and practice*. Washington, DC: Georgetown University Press.

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review, 94,* 439–454.

Skehan, P. (1986). Cluster analysis and the identification of learner types. In V. Cook (ed.), *Experimental approaches to second language acquisition*, pp. 81–95. Oxford: Pergamon.

Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition, 13,* 275–298.

Snow, R. E. (1991). Aptitude-Treatment Interaction as a framework for research on individual differences in psychotherapy. *Journal of Consulting and Clinical Psychology, 59,* 205–216.

Sparks, R., Ganschow, L., & Pohlman, J. (1989). Linguistic coding deficits in foreign language learners. *Annals of Dyslexia, 39,* 179–195.

Squire, L. R., & Knowlton, B. J. (2000). The medial temporal lobe, the hippocampus, and the memory systems of the brain. In M. Gazzaniga (ed.), *The new cognitive neurosciences*, pp. 765–780. Cambridge, MA: MIT Press.

Squire, L. R., & Wixted, J. T. (2011). The cognitive neuroscience of human memory since H. M. *Annual Review of Neuroscience, 34,* 259–288.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems.

Proceedings of the National Academy of Science, 93, 13515–13522.

Trahan, D. E., & Larrabee, G. J. (1988). *Continuous Visual Memory Test*. Odessa, FL: Assessment Resources.

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science, 2,* 67–70.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition, 4,* 105–122.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92,* 231–270.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (ed.), *Processing approaches to adult SLA: Theory and practice*, pp. 141–178. Washington, DC: Georgetown University Press.

Unterrainer, J. M., Rahm, B., Leonhart, R., Ruff, C. C., & Halsband, U. (2003). The Tower of London: The impact of instructions, cueing, and learning on planning abilities. *Cognitive Brain Research, 17,* 675–683.

VanPatten, B. (1994). Evaluating the role of consciousness in second language acquisition: Terms, linguistic features, and research methodology. *AILA Review, 11,* 27–36.

Weinberg, S. L., & Abramowitz, S. K. (eds.) (2002). *Data analysis for the behavioral sciences using SPSS*. New York: Cambridge University Press.

Williams, J. N. (2011). Working memory and SLA. In S. M. Gass & A. Mackey (eds.), *The handbook of second language acquisition*, pp. 427–441. New York: Routledge.

Witt, K., Nuhsman, A., & Deuschi, G. (2002). Intact artificial grammar learning in pateients with cerebellar degeneration and advanced Parkinson's disease. *Neuropsychologia, 40,* 1530–1540.

Wong, P. C. M., Morgan-Short, K., Ettlinger, M., & Zheng, J. (2012). Linking neurogenetics and individual differences in language learning: The dopamine hypothesis. *Cortex, 48,* 1091–1102

Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics, 28,* 565–585.

Wong, P. C. M., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping, 28,* 995–1006.

Wong, P. C. M., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., & Zatorre, R. J. (2008). Volume of left Heschl's gyrus and linguistic pitch learning. *Cerebral Cortex, 18,* 828–836.