The Role of Working Memory in Implicit and Explicit Language Learning

Kaitlyn M. Tagarelli (kmt49@georgetown.edu)
Department of Linguistics, Georgetown University
37th and O Streets, NW, Washington, DC 20057

Mailce Borges-Mota (mailce@cce.ufsc.br)
Department of Foreign Languages and Literature, Federal University of Santa Catarina
CCE – Building B, Florianopolis, Santa Catarina, 88040900, Brazil

Patrick Rebuschat (per6@georgetown.edu)
Department of Linguistics, Georgetown University
37th and O Streets, NW, Washington, DC 20057

Abstract

Working memory capacity (WMC) has been shown to correlate with performance on complex cognitive tasks, including language comprehension and production. However, some scholars have suggested that performance outcomes result from an interaction between individual differences (IDs), such as WMC, and learning conditions (Robinson, 2005a). Reber, Walkenfeld, and Hernstadt (1991) specifically claimed that IDs influence performance on explicit, but not implicit, processes. In this study, English native-speakers were exposed to a semi-artificial language under incidental or rule-search conditions, and their WMC was measured by two complex-span tasks. Both conditions produced a clear learning effect, with an advantage for the rule-search group. No significant correlations between overall performance on a grammaticality judgment task and WM scores were found for either group. However, WMC predicted performance on grammatical items for the rule-search group. These results support Reber et al.'s (1991) claim that aptitude measures may only be predictive of learning in explicit conditions.

Keywords: implicit learning; explicit learning; working memory

Introduction

Working memory has been shown to play a role in many aspects of first language (L1) processing and performance. Yet because of certain limitations on second language (L2) learners, such as maturational constraints, WM may be more important in L2 than in L1 (Miyake & Friedman, 1998). Many studies have found relationships between WM and L2 proficiency and development. WM capacity (WMC) has been shown to positively correlate with performance on sections of the TOEFL, reading comprehension abilities (for a review, see Miyake & Friedman, 1998), gender and number agreement processing (Sagarra & Herschensohn, 2010), the ability to make use of interactional feedback in classroom settings (Mackey, Philp, Egi, Fujii, et al., 2002), and L2 proficiency (van den Noort, Bosch, and Hugdahl, 2006). Overall, the evidence suggests that WMC plays a role in L2 acquisition. However, the predictive power of WMC may be mediated by learning conditions.

The Role of Context

Research into individual differences (IDs) and pedagogical approaches to L2 acquisition have recently merged because of the inextricable link between the two. Skehan (2002) has suggested that aptitude is more predictive of L2 acquisition success in formal, structured learning contexts. Support for combining these lines of research has also come from Robinson (2002) and Erlam (2005), who expressed that "a particular method of instruction may not (...) benefit all learners uniformly" (p. 147). However, Robinson and Erlam make no claims about whether IDs should have a greater effect in one learning condition than in others.

Research in cognitive psychology suggests that there should be differences in how aptitude influences learning in implicit and explicit conditions because these two systems are fundamentally distinct. Reber et al. (1991) explain this distinction by evoking the idea of the "primacy of the implicit" (p. 888). The basis of this primacy is that implicit processes have older biological substrates, which vary little among "corticated species" (Reber, 1989, p. 232), and even less from human to human. Thus, implicit processes should be more robust than explicit processes, unconscious functions should operate relatively uniformly within the population as compared to conscious functions, and IDs should not contribute to variance in implicit processes, but should in explicit processes.

Studies that investigate implicit and explicit learning adhere to several fundamental criteria when operationalizing these conditions. In explicit conditions, subjects are made aware that they are supposed to learn something, and they usually know that they will be tested. There are two main options for explicit conditions. In rule-search conditions, subjects are exposed to the system to be learned and instructed to find rules. Alternatively, subjects can be presented with rules (rule-instruction) and exposed to the system. In some cases, explicit conditions consist of simply drawing the subjects' attention to the target structure without indicating that there is a system to be learned (awareness raising). In implicit conditions, subjects are not informed about the true purpose of the experiment; they do not know that they should be learning something or that they will be tested. Sufficient exposure to the system is an important component of implicit conditions. Exposure tasks, such as Reber et al.'s (1991) memorization task, are generally used to draw subjects' attention away from the true objective of the experimenter. Learning in such conditions in often therefore referred to as incidental.

Reber et al. (1991) investigated the relationship between IQ and learning in implicit and explicit (awareness raising) conditions and found greater variance in accuracy on grammaticality judgment tasks (GJTs) between subjects in the explicit condition than in the implicit condition, and a significant correlation between accuracy and IQ for the former group, but not for the latter. Replicating Reber et al.'s (1991) implicit and explicit learning conditions, Robinson (2005b) also found less variation in implicit than explicit learning outcomes. However, he found that learning in the implicit condition was significantly negatively correlated with the verbal abilities component of IQ, whereas there were no correlations between learning in the explicit condition and IQ. Robinson (2005b) added an additional condition, incidental learning of Samoan, in which subjects were presented with sentences and asked to try to understand the meaning, but were not told about grammar rules. Performance on a GJT was not related to any aptitude measures.

In a study on eXperanto, an artificial language, deGraff (1997) found that the explicit group (rule-search) outperformed the implicit group (sentence rehearsal), but performance in *both* conditions correlated positively with aptitude. Erlam (2005) investigated the effects of aptitude on learning in three conditions: deductive (rule-instruction + practice), inductive (practice only), and structured input (rule-instruction only). There were few correlations for the deductive group, suggesting that "instruction that provides students with explicit rule explanation and then gives them opportunities to engage in language production tends to benefit al language learners" (p. 163). The relationships between aptitude measures and learning differed for each condition. A measure of WMC correlated with performance in the structured input group, leading Erlam to conclude that students with higher WMC are better at processing explicit input. However, it must be noted that Erlam's "WM task" was actually a simple-span task which measures phonological short-term memory, not WM (Miyake & Friedman, 1998), so these findings cannot be generalized to WM, but only a component of this system.

In fact, few studies have specifically looked at the interaction between WM and instructional contexts. Considering the importance placed on the central executive in WM (Engle, 2001), one might expect WM to be a good predictor of language success in explicit rather than implicit learning conditions because explicit processes are closely related to attention, whereas implicit processes are not (DeKeyser, 2003). Robinson (2005b) included a measure of WMC and found that it positively correlated with incidental learning of Samoan, but not with learning in implicit or explicit conditions. He concluded that incidental learning involves "the ability to process for meaning while simultaneously switching attention to form during problems in semantic processing," and that this is "an ability strongly related to [WMC]" (p. 55). Perhaps when the target system has meaning, as in natural language, rather than only form, as in meaningless letter strings (used in the implicit learning conditions in Reber et al. (1991) and Robinson (2005b)), learners rely more on general cognitive processes such as WM when explicit instructions are not provided. A study on Japanese 5^{th} grade students learning English (Ando, Fukunaga, Kurahashi, Suto, et al., 1992) found a complex interaction between WM and L2 success in naturalistic learning conditions. Children with high WMC benefited from an explicit, form-focused teaching approach, whereas children with low WMC benefited from an implicit, communicative teaching approach.

The research outlined above demonstrates that the relationship between WMC and learning in explicit and implicit conditions is a complicated one, and there is thus far little consensus as to how this component of aptitude interacts with learning conditions.

The Current Study

In this study, we seek to contribute to the research on the role of WM in L2 acquisition, and how this ability interacts with learning conditions. We are interested in whether WMC influences an individual's ability to learn L2 syntax, and whether it has a differential effect on learning under implicit or explicit learning conditions.

Method

Participants

Participants were 62 native speakers of English with no background in German or any other V2 language. They were assigned to one of two experimental groups: *incidental* and *rule-search*. The groups did not differ across the variables age, gender, occupation, and number of languages acquired, all $p > .05$. Participants were compensated for their participation.

Stimulus material: Semi-artificial language

A semi-artificial language consisting of English words and German syntax was used to generate the stimuli for this experiment (see also Rebuschat, 2008). The advantages of using a semi-artificial language of this nature are that the grammatical complexity of natural languages is maintained and semantic information is present. The sentences generated by the grammar follow three specific verb-placement rules, each associated with a specific syntactic pattern (see Table 1). A total of 180 sentences were drafted for this experiment. See Rebuschat (2008) for a more extensive description of the system.

Table 1: Descriptions of the verb-placement rules.

Rule	Description (*Example*)
V2	Place finite verb in second phrasal position of main clauses that are not preceded by a subordinate clause. (*Today* **bought** *John the newspaper in the supermarket.*)
V1	Place finite verb in first position in main clauses that are preceded by a subordinate clause. (*Since his teacher criticism voiced,* **put** *Chris more effort into his homework*.)
VF	Place finite verb in final position in all subordinate clauses. (*George repeated today that the movers his furniture* **scratched**.)

Training set The training set consisted of 120 sentences (40 for each rule) and was subdivided into 60 plausible and 60 implausible constructions. For example, a sentence like "Chris entertained today his colleagues with an interesting performance." was semantically plausible, while "After his wife a thief surprised, communicated George with the police banana." was semantically implausible.

Testing set The testing set consisted of 60 new sentences subdivided into 30 grammatical and 30 ungrammatical items, all of which were plausible. Ungrammatical sentences were similar to the grammatical ones, but the VP position was incorrect. With the exception of a limited number of function words, no words were repeated from the training set, making the test analogous to the transfer paradigm in Artificial Grammar Learning research (Reber, 1969).

Procedure

Subjects attended two sessions: an artificial language learning session and a WM session. For each subject, the WM session occurred at least one day and no more than two weeks after the artificial learning language session. Stimuli and instructions for both sessions were presented on a Macintosh computer using SuperLab, version 4.

Artificial Language Learning This session consisted of two parts: an exposure phase, during which subjects were presented with 120 instances of the artificial language in random order, and a testing phase.

Exposure phase. Participants in the *incidental condition* (n = 31) were asked to listen the 120 sentences of the training set, repeat each sentence after a delayed prompt (1,500 ms) and judge the semantic plausibility of each sentence. Importantly, subjects were not told that the syntax underlying the training sentences followed a rule system, nor were they told that there would be a test on word order after the exposure phase. Instead, these subjects were simply informed that they were taking part in a sentence comprehension experiment that sought to investigate how scrambling affects our ability to understand the meaning of sentences. Subjects were thus exposed to the artificial language under incidental learning conditions.

Participants in the *rule-search condition* (n = 31) were asked to listen to the same 120 sentences. They were told, at the beginning of the experiment, that the word order of the sentences was determined by a "complex rule-system" and that their task was to listen carefully to each sentence and to discover the word-order rules. Subjects were also informed that they would later be tested on the rules. Subjects were thus exposed to the artificial language under intentional learning conditions.

Testing phase. After exposure, subjects in the incidental condition were told that the word order of the previous 120 sentences was determined by a "complex rule-system;" subjects in the rule-search condition were reminded of this fact. All subjects then listened to 60 new sentences, as described in the testing set. For each sentence, subjects were asked to judge whether the sentence followed the rule system of the sentences in the exposure phase (GJT), report how confident they were in their judgment (*not confident, somewhat confident, very confident*), and indicate the basis of their judgment (*guess, intuition, memory, rule*). The confidence ratings and the source attributions were used as subjective measures of awareness (Dienes, 2008).

After the testing phase, subjects completed a debriefing questionnaire which prompted them to verbalize any rule or regularity they might have noticed during the course of the experiment. The questionnaire also asked participants to provide their gender, age, nationality, occupation and language background.

Working memory assessment On a separate day, subjects performed two WM tasks, the *operation-word span task* (OWST, Turner & Engle, 1989; Unsworth, Heitz, Schrock & Engle, 2005) and the *letter-number ordering task* (LNOT, Wechsler, 1997). The order in which subjects completed these tasks was counterbalanced. In the OWST, subjects saw an equation and word appear on the computer screen. They read the equation out loud, stated whether the answer provided was correct or not, and then read the following word out loud. For example, if the participant saw "IS (6 x 2) + 1 = 10 ? CAT," they would say "Is six times two plus one equal to ten…no…cat." Once the subject said the word, the experimenter advanced to the next operation and word in the set. There were 12 sets overall, with two to five words in a set. At the end of each set, a cue appeared to prompt participants to write down all of the words that they could remember from that set. Subjects were awarded one point for every word remembered in the correct order, for a total possible score of 42 points.

The LNOT is part of the WAIS-III Intelligence Scale (Wechsler, 1997). We used an English version of the task, adapted from van den Noort et al. (2006). In this task, the experimenter read aloud series of letters and numbers, from two to eight digits long. The subject was asked to repeat the numbers in numerical order and then the letters in alphabetical order. For example, if the experimenter read "W-1-K-5," the subject would repeat "1-5-K-W." Subjects received one point for every series repeated back correctly, for a maximum of 21 points. If a subject missed three series in a row, the experimenter discontinued the task and counted the subject's score from all of the previous series.

Results

Seven participants were excluded from the analysis because they did not follow directions correctly in the testing phase, for a total of 29 participants (23 women, 6 men, M_{age} = 21.4 years) in the incidental group and 26 participants (20 women, 9 men, M_{age} = 23.0 years) in the rule-search group.

Grammaticality Judgments

The analysis of the GJT showed that the incidental group classified 58.9% (*SD* = 8.6%) of the items correctly, while the rule-search group classified 71.2% (*SD* = 15.5%) of the items correctly. This difference was significant, $t(53) = 3.692$, $p < .001$. Both the rule-search group, $t(25) = 6.977$, $p < .001$, and the incidental group, $t(28) = 5.563$, $p < .001$, performed significantly above chance. The training phase thus produced a clear learning effect in both experimental groups, with an advantage for the rule-search group.

The rule-search group endorsed 75.9% (*SD* = 16.9%) of grammatical items and 33.4% (*SD* = 19.5%) of ungrammatical items. The incidental group endorsed 71.7% (*SD* = 14.2%) of grammatical items and 54.0% (*SD* = 19.1%) of ungrammatical items. A mixed ANOVA with *learning condition* as a between groups factor and *grammaticality* as a within groups factor revealed a main effect of *grammaticality*, $F(1,53) = 27.415$, $p < .001$, and an interaction between *grammaticality* and *group*, $F(1,53) = 5.937$, $p < .05$. Bonferroni corrected post-hoc tests showed that both groups endorsed more grammatical items than ungrammatical items, $p < .001$. That is, performance on the GJT was driven by memory for previously encountered syntactic patterns.

Bonferroni corrected post-hoc tests also showed that the difference between rule-search and incidental groups on the endorsement of grammatical items was not significant, $p > .05$, i.e. neither group was more likely to correctly endorse grammatical strings, but the difference between groups on ungrammatical strings was significant, $p = < .001$. The lower accuracy of the incidental group was therefore largely due to poorer performance on ungrammatical items.

Confidence Ratings and Source Attributions

Subjective measures reveal the extent to which knowledge was conscious or unconscious. Dienes (2008) distinguishes two types of knowledge, namely *structural knowledge* and *judgment knowledge*. Structural knowledge is knowledge about the structure of sequences in the language, and judgment knowledge is knowledge about whether test items share this structure with training items. Structural and judgment knowledge are conscious only if the individual is aware of having that knowledge (e.g., *I know the verb comes in final position in a subordinate clause* or *I know that this sentence is not like the training sentences*).

Confidence ratings show that participants in the rule-search group had unconscious judgment knowledge by the *guessing criterion*. That is, their accuracy was above chance when they said they were not confident (67%), indicating that subjects in this group acquired knowledge about the grammaticality of test sentences, but they were not aware of having acquired that knowledge. Learners in the incidental group were at chance when they said they were not confident (56%); the *guessing criterion* for unconscious judgment knowledge was thus not satisfied for this group.

The Chan difference score was computed in order to establish whether learning was implicit by the *zero-correlation criterion*. The difference between the average confidence for correct and incorrect judgments was not significant for either group, $p > .05$; they were not more confident in correct decisions than in incorrect ones. This indicates unconscious judgment knowledge by the *zero-correlation criterion*. At least for some of the knowledge, subjects were not aware of the fact that they had acquired it during the exposure phase.

An analysis of source attributions shows that there were no differences in accuracy between groups based on which type of knowledge they reported using for classification judgments. Both groups performed significantly above chance when basing their judgments on intuition (rule-search: 64%; incidental: 61%), memory (70%; 58%), and rule (69%; 59%), but only the rule-search group performed above chance when basing their judgments on a guess (69%; 55%). Above-chance accuracy when using intuition suggests that participants in both groups developed at least some unconscious structural knowledge of the grammar.

Above chance accuracy when reporting high confidence (69%; incidental: 61%) and basing judgments on memory and rules (see above) suggests that both groups also developed conscious structural and judgment knowledge.

Verbal Reports

In both the incidental and rule-search groups, most participants verbalized incorrect rules for the artificial grammar. Participants frequently mentioned that the verb could appear at the end of the sentence, but they did not indicate that this was only possible in a subordinate clause. A few participants provided examples of one sentence type, but not the other two. However, three participants in the rule-search group provided examples of all sentence types and verbalized all verb-placement rules.

Working Memory Tests

The average score on the OWST was 29.6 (*SD* = 5.5) for the incidental group and 29.6 (*SD* = 5.0) for the rule-search group. For the LNOT, the average scores were 13.0 (*SD* = 2.8) for the incidental group and 12.1 (*SD* = 2.5) for the rule-search group. The difference between groups was not significant on the OWST or the LNOT, $p > .05$.

When all subjects were analyzed together, there were no correlations between accuracy on the GJT and performance on either the OWST, $r = .117$, $p > .05$, or the LNOT, $r = .223$, $p > .05$. For the incidental group, there were no significant correlations between accuracy on the GJT and either WM task (OWST: $r = .168$, $p > .05$; LNOT, $r = .182$, $p > .05$). For the rule-search group, there was no correlation between accuracy on the GJT and performance on the OWST, $r = .117$, $p > .05$, but there was a significant correlation between accuracy on the GJT and performance on the LNOT, $r = .477$, $p < .05$, suggesting that according to one of our WM measures, WM predicts learning only in one learning condition.

As mentioned above, three participants demonstrated awareness of the rules of the artificial language. These subjects were 98%, 100%, and 100% accurate on the GJT. Their scores on the LNOT were 16, 15, and 17, and their scores on the OWST were 25, 35, and 23, respectively. Because the three verbalizers seem to behave differently than the rest of the participants in the rule-search group, further analyses were conducted without these participants. When the relationship between WM and accuracy for the rule-search group was analyzed without the verbalizers, the correlation between LNOT and accuracy disappeared, $r = .156$, $p > .05$, and the correlation between OWST and accuracy remained non-significant, $r = .309$, $p > .05$.

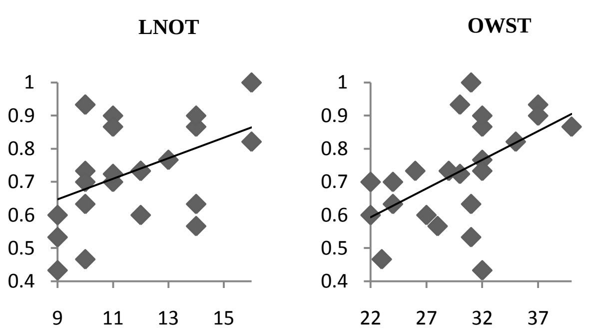

Figure 1. Correlations between WM and accuracy on grammatical items for the rule-search group, excluding verbalizers. The y-axis represents accuracy; the x-axis represents WM score for LNOT (left) and OWST (right).

Because of the differential performance observed for grammatical and ungrammatical items, we analyzed these items separately. These analyses revealed a positive correlation between WM and accuracy on grammatical items for participants in the rule-search group only (excluding verbalizers; see Figure 1). This was true for both the LNOT ($r = .424$, $p < .05$) and the OWST ($r = .542$, $p < .05$). The correlations were not significantly different, $t < 1$. There were no other significant correlations between WM score and accuracy.

Discussion

The analysis of the GJT produced similar results to those found previously for incidental and rule-search conditions in this semi-artificial language (Rebuschat, 2008). Both incidental and rule-search groups performed significantly better than chance, and this observed learning effect appears to be driven by the correct endorsement of grammatical items. That is, both groups memorized learned patterns from the training phase. The rule-search condition gave learners an advantage in distinguishing learned patterns from novel patterns, as demonstrated by their superior accuracy on ungrammatical items. Confidence ratings, source attributions, and verbal reports indicate that subjects in both groups had at least some explicit knowledge of the grammar they were exposed to, and also that subjects in both groups had at least some implicit knowledge, which is expected based on previous findings (Rebuschat, 2008, Exp. 6).

Our analyses of individual differences suggest that WM does affect an individual's ability to learn L2 syntax, but this effect is apparent only in certain conditions, and for certain items. Overall, we found no correlations between either WM test and accuracy in the incidental group, nor for the OWST and accuracy for the rule-search group. The LNOT appeared to predict learning for the rule-search participants, but the importance of assessing awareness becomes clear, as we determined that this correlation was entirely driven by three exceptional learners who were able to verbalize rules on the debriefing questionnaire. These three learners had very high WMC as measured by the LNOT, which may have helped them discover rules. However, many other variables could have influenced their performance. Therefore, we cannot make any claims about the relationship between the WM skills involved in the LNOT and L2 acquisition based on these participants alone.

A deeper analysis of the relationship between WM and item classes revealed that in the rule-search condition, WM predicted performance on grammatical items only. Thus, while WM, as measured by the OWST and the LNOT, does not appear to affect an individual's ability to learn L2 syntax in incidental learning conditions, it may play a role in helping learners in more explicit conditions incorporate positive evidence from a new language.

Our findings support Reber et al's (1991) claim that individual differences influence learning in explicit, but not implicit conditions. However, the differential effect of WMC on grammatical and ungrammatical items suggests that this relationship is complex. Furthermore, while we can make some claims about the predictive value of WM in L2 learning under implicit and explicit conditions, there are several limitations of this study that must be acknowledged.

Firstly, untimed GJTs such as the ones used in this study might favor explicit processes. In the future, it would be worth employing a learning measure that favors implicit processes, e.g. the elicited imitation task proposed by Ellis (2005). Additionally, some studies have found that correlations become significant when comparing aptitude measures and *delayed* posttests, (Erlam, 2005; Mackey et al., 2002), so it would be of interest to include a retention phase. Also, we explored the role of WM, which involves the control of attention and may therefore be more related to explicit than implicit learning. Contrary to the claims of Reber et al (1991), recent evidence suggests that implicit learning is indeed an ability that varies across individuals (Kaufman, DeYoung, Gray, Jiménez, et al., 2010). Yet as Kaufman et al. found, typical cognitive abilities, such as IQ and WM, do not correlate with implicit learning outcomes, whereas other cognitive abilities do. The incorporation of cognitive factors that might draw on some of the same processes as implicit learning might yield different results. Finally, it is crucial to note that this study investigates the relationship between WM and *learning in two different conditions*, not *types of learning*. As shown by measures of awareness, implicit and explicit conditions do not necessarily nor exclusively result in implicit and explicit knowledge, respectively; the relationship is more complex.

Acknowledgements

The authors would like to gratefully acknowledge the financial support of this study by a research grant from the Federal University of Santa Catarina, Brazil, as well as Luke Amoroso, Phillip Hamrick, Katie Jeong-Eun Kim, Julie Lake, and John Williams for their feedback.

References

Ando, J., Fukunaga, N., Kurahashi, J., Suto, T., Nakano, T., and Kage, M. (1992). A comparative study of the two EFL teaching methods: The communicative and grammatical approach. *Japanese Journal of Educational Psychology, 40*, 247–256.

de Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition, 19*, 249-276.

DeKeyser, R. (2003). Implicit and explicit learning. In C. J. Doughty, & M. H. Long, *Handbook of Second Language Acquisition*. Oxford: Blackwell.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Progress in Brain Research , 168*, 49-64.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*, 141-172.

Engle, R. W. (2001). What is working-memory capacity? In H. L. Roediger III & J. S. Nairne (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder*. Washington, DC: American Psychological Association.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research, 9*(2), 147-171.

Kaufman, S.B., DeYoung, C.G., Gray, J.R., Jiménez, L., Brown, J., Mackintosh, N. (2010). Implicit learning as an ability. *Cognition, 116*, 321-340.

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

Miyake, A. & Friedman, N.P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In Healy, A. & Bourne, L. (eds.), *Foreign Language Learning: Psycholinguistics studies on training and retention*. Mahwah, NJ: Lawrence Erlbaum.

Reber, A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology, 81*, 115-119.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General , 118* (3), 219 - 235.

Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(5), 888-896.

Rebuschat, P. (2008). *Implicit learning of natural language syntax*. Unpublished PhD thesis, University of Cambridge.

Robinson, P. (2002). *Individual differences and instructed language learning*. Amsterdam: Benjamins.

Robinson, P. (2005a). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46-73.

Robinson, P. (2005b). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning. *Studies in Second Language Acquisition, 27*, 235-268.

Sagarra, N. & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua, 120*(8), 2022-2039.

Skehan, P. (2002). Theorizing and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning*. Amsterdam: Benjamins.

Turner, M. L. & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127-154.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498-505.

van den Noort, M. W. M. L., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and Working Memory Capacity. *European Psychologist, 11*, 289-296.

Wechsler, D. (1997). *The Wechsler Adult Intelligence Scale: Revised.* New York: The Psychological Corporation.