5-2013

# The Effects of Pedagogical Conditions on Second Language Acquisition

McCall Evonne Sarrett
msarrett@utk.edu

# The Effects of Pedagogical Conditions on Second Language Processing

McCall Sarrett

University of Tennessee-Knoxville

Abstract

In this project, we designed a study to compare variations in language processing via behavioral and neurocognitive measures when participants are instructed in a language (in this case, Latin) using different computer-administered pedagogical methods: implicit and explicit. The resultant knowledge is measured behaviorally (via a battery of linguistic tasks) and using electroencephalography (EEG) to measure event-related potentials (ERPs) in the brain. In addition, the effects of individual differences in memory are explored, as participants perform memory tests measuring declarative memory (memory that is generally verbalizable), and procedural memory (memory for skills and procedures, which is not verbalizable).

Participants included 9 college students who were native English Speakers, right-handed, ages 18-35, with no exposure to Latin or any other case-marking language and no more than 2 years of college exposure to any other language. Predictions were as follows: (1) participants may show different ERP signatures and different profiles of language learning, varying with the method of instruction; (2) strength of declarative and procedural memory may influence proficiency and/or neural processing of the second language.

In the piloting of this study, it was found that the proficiency of the language learners on the behavioral tests did not reach a satisfactory level, thus indicating that more language training would be necessary in order to collect meaningful ERP data. Due to time constraints, we were unable to invite participants back for the requisite extra training and collection of data; thus, the crux of this paper is the exploration of the behavioral data collected and the analyses between the correlations of selected variables.

The Effects of Pedagogical Conditions on Second Language Processing

Intuitively and subjectively, it is largely assumed that there is a difference in language processing between native speakers of a language and those who learn it as a late-acquired second language (L2). Through recent research in second language acquisition (SLA) and cognitive neuroscience, the field has been able to observe these differences in neural processing between native speakers and L2 learners and have indeed found some differences (see below). One area, however, that is not well understood is the difference in L2 processing between speakers who have learned that language in different ways. Does an L2 learner process the language differently if they learn it in a classroom explicitly versus if they implicitly pick it up from living in a country where that language is spoken? This question is beginning to be explored in this pilot study, where we use a small chunk of Latin language and grammar as a sort of miniature language to be taught to participants using two different modes of instruction in an attempt to compare differences in profiles of language learning and ERP signatures in the brain. We further question whether strength of different types of memory systems, i.e. declarative and procedural, may be correlated with how well a language is learned given a specific learning environment, i.e. more explicit instruction and practice versus more explicit instruction and practice.

Predictions are that (1) participants may show different ERP signatures and different profiles of language learning, varying with the method of instruction; (2) strength of declarative and procedural memory may influence proficiency and/or neural processing of the second language.

These predictions are based on previous findings in second language acquisition literature regarding implicit and explicit instruction (Norris and Ortega, 2000), on the

Declarative/Procedural model (Ullman, 2005) and on findings that investigate this model via artificial languages (Morgan-Short et al., 2012).

## Background

### The Declarative-Procedural Model

Although some theories sustain that native languages (L1) are basically equal to second languages in the way they are processed, as held by connectionism (Elman et al., 1996; Seidenberg, 1997). However, some evidence suggests that while semantics may be processed in L2, even at initial levels, in similar ways to L1, syntactic processing is not native-like until a high proficiency is reached (Friederici, 2002; Friederici, Steinhauer & Pfeifer, 2002; Hahne, 2001; Hahne & Friederici, 2001). The predictions in this study are based on these previous findings and on the L2-related predictions of the Declarative-Procedural Model of language set forth by Ullman (2005).

The Declarative-Procedural model states that there is a fundamental distinction between our mental lexicon, which consists of the building blocks of language, and our mental grammar, which is the compositional system of that language. This is in stark contrast to connectionist accounts of SLA mentioned earlier and instead hypothesizes that the mental lexicon and the mental grammar are processed largely by two distinct memory systems: the declarative memory system and the procedural memory system, respectively. Both are known to have a role in non-linguistic functions as well, but in this paper we focus on the role that they play in both L1 and L2 processing.

Declarative memory, generally speaking, is memory for facts, events, and arbitrary relations, which is available to conscious awareness (usually, but not necessarily) and is able to be verbalized. Examples might include knowing your state capitals or recalling a specific life event. It is thought to subserve the mental lexicon, i.e. semantics or word-meaning pairs and irregular, non-predictable forms of words, with its underlying neural substrates being based in the temporal lobe, especially in the hippocampus.

The procedural system underlies motor and cognitive skills and habits, such as riding a bicycle. It is just one of the types of implicit memory, that is, memory that is non-conscious or automatic. In L1, it generally subserves the mental grammar, which includes the predictable compositional forms of language, and it is thought to be tied to the basal ganglia and frontal lobe.

When learning an L2, adults are hypothesized, according to the Declarative/Procedural Mode, to rely more on the declarative system not only for lexical but also for grammatical aspects of language (i.e., memorizing declarative grammar rules). With sufficient experience and higher proficiency, however, it is thought that adults can begin to process L2 grammar using the procedural system, much like an L1 learner of the language. This hypothesis is supported by various lines of evidence including behavioral data, electrophysiological (EEG) data, neuroimaging data, and patient data (Ullman, 2005). Since this project seeks to use EEG as its measure of cognitive processing, we will begin by reviewing ERP data from previous literature. Although in the end we did not collect ERP data in the pilot study, because it was a part of the design of the study and we still conceive of it as a part of what will be the full-scale study, we include it here.

**ERPs Related to Lexicon and Semantics**

Due to its high temporal resolution (on the order of milliseconds), ERPs can be used to examine the electrical potentials on the scalp immediately (0-1000 ms) following the presentation of a written or auditory stimulus. It is found that certain patterns are characteristic to different types of language stimuli, including to linguistic violations, and these patterns, or components, are distinguished by their polarity (either positive or negative), amplitude, latency, and scalp distribution. What has been elicited in response to lexical and semantic violations, and is thought to be reflective of the declarative memory system (Ullman, 2004), is the N400 (See Figure 1). The N400 gets its name as a negativity that occurs approximately 400 ms post-stimulus onset and is usually located in the centro-parietal area. It may be linked to bilateral temporal lobe structures that once again are thought to be neural correlates of lexical processing.

In L1, an N400 is consistently and strongly produced when participants are presented with a semantic error (Friederici, Pfeifer & Hahne, 1993; Hahne, 2001; Ullman et al., 2000; Weyerts et al., 1997). For example, "Toast tastes good with #socks;" as compared to "Toast tastes good with butter." In L2, the N400 is a robust component that has been elicited at both low (Hahne, 2001; Hahne & Friederici, 2001; Ojima et al., 2005; Weber-Fox & Neville, 1996) and high proficiency (Hahne, 2001; Hahne & Friederici, 2001; Ojima et al., 2005).

**ERPs Related to Syntax**

In grammatical processing, there are two ERP components that have been widely observed, the left anterior negativity (LAN) and the P600 (See Figure 2). The LAN is usually observed about 300-500 ms after a target word is presented and is linked to left frontal structures in the brain. It is posited to reflect grammatical or procedural memory (Ullman 2001, 2004). The P600 is a positivity occurring at 600 ms and has been linked to posterior brain regions, and while it is usually seen in conjunction with the LAN, it is not thought to reflect procedural memory, rather it is thought to be involved in conscious processing, repair, and reanalysis of a violation of grammar or syntax.

The LAN or the LAN and the P600 have been elicited in L1 speakers fairly consistently (Barber & Carreiras, 2005; Coulson, King & Kutas, 1998; Friederici, 2002; Friederici, Pfeifer &

Hahne, 1993; Neville et al., 1991; Penke et al., 1997; Rodriguez-Fornells et al., 2001; Steinhauer et al., 2006), but these components are seen less often in L2 speakers, usually only observed at high levels of proficiency. Ojima et al. (2005) observed a LAN when presenting English violations to high proficiency English L2 speakers who were native speakers of Japanese, while a LAN and a P600 has been observed in high proficiency L2 learners by Friederici, Steinhauer, and Pfeifer (2002). The LAN, however, is not observed at low proficiency, according to Hahne (2001), indicating at the very least that there are differences in ERP components in early stages of L2 acquisition as compared to L1, and possibly that the procedural memory system may not yet play a role in L2 processing at this stage.

**ERPs and Instructional Method**

While the role of instructional method on second language instruction has been widely investigated, there is still not a clear consensus as to what type of instructional domain yields the greatest effectiveness (Norris and Ortega, 2000). Measures of learning in these contexts has been explored in a variety of ways, and recently using EEGs to explore the acquisition of an artificial language suggests that L2 learners taught with an implicit mode of instruction show more native-like ERP signatures in response to syntactic and at high proficiency (LAN and P600 observed; at low proficiency, an N400 was observed). In contrast, learners of L2 taught with an explicit mode of instruction showed no significant effects at low proficiency, but did show a P600 at high proficiency, accompanied by an anterior positivity (not a LAN) (Morgan-Short et al., 2012). Further, effects of individual differences of strengths of declarative and procedural memory are being explored. 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 have been indicated using an artificial language to look at L2 acquisition (Morgan-Short et al., in press). These findings in artificial languages will be useful for approaching the effects of L2 pedagogy and the roles of declarative and procedural memory in natural L2s.

**The Present Study**

Drawing from the previous literature, this study seeks to explore behaviorally the differences in L2 processing varying with method of instruction, and it would be expected to see the ERP components mentioned when participants reach a sufficient level of proficiency. Specifically, we would expect to see a difference especially in grammar or syntactic processing, such that the implicit mode of instruction would be more likely to lead to a native-like LAN, while the explicit mode of instruction may more often elicit the N400. These predictions follow those in line with the declarative-procedural model of SLA, in that second language learners may initially depend more on the declarative memory system for grammar and lexicon, especially when taught explicitly as opposed to implicitly (Ullman, 2005). However, while we might expect to see more native-like ERP signatures with the implicit mode of instruction, we still expect that the explicit mode will yield better behavioral results (Norris & Ortega, 2000), though the literature on this topic is still somewhat in dispute.

Likewise, we would expect to see a similar variation depending on strength of declarative and procedural memories. Since the LAN has been implicated in the processes of the frontal structures of the procedural memory system, we would expect to see participants with stronger measures of procedural memory to show a stronger LAN; similarly, we would expect participants with stronger measures of declarative memory to show a stronger N400 effect.

While L2 and L1 show differences in processing that seem to converge at high L2 experience and proficiency, the present study seeks to explore any differences between different groups of L2 learners themselves based on the method of training as well as on individual differences in declarative and procedural memory.

## Materials and Methods

In order to explore the principle questions of this study, there are three sessions in this experiment: one session for cognitive tasks, in order to measure strengths of the declarative and procedural memory systems; one for Latin language training, designed to examine the differences in language learning on various tasks in relationship to implicit and explicit modes of pedagogy; and one for electroencephalography (EEG) testing as participants complete a grammaticality judgment task, in order to explore possible differences in language processing in the brain.

### Participants

Participants are monolingually raised native English speakers, ages 18-35, with some college education. Due to the lateralization of language in the brain, handedness was also controlled for (Knecht, 2000); all participants were confirmed to be right-handed. Participants were excluded if they had more than one native language, if they had more than two years of any college language classes, or if they had any exposure to any case-marking languages, such as German, Arabic, or Latin itself. Using a demographic survey and background history questionnaire, it was confirmed that participants had no known learning or cognitive disorders, and no medical history that may be relevant to the collection of EEG data (e.g., invasive brain procedures or long term psychiatric medication).

### Cognitive Tasks

Declarative memory is measured using the Modern Language Aptitude Test (MLAT), Part V, and the continuous visual memory test (CVMT). Procedural memory is measured using the Alternative Serial Response Task (ASRT) and the dual-task version of the weather prediction task.

**Modern Language Aptitude Test (MLAT), Part V**

The MLAT, Part V is a (linguistic) measure of declarative memory, which probes how easily a participant is able to learn twenty-four vocabulary words in Kurdish and their English translations (Carroll & Sapon, 1959). Participants are first given two minutes to study a paired list, followed by a two-minute fill-in-the-blank translation practice session with the list. Finally, both the list and the practice fill-in-the-blank sheet are taken up, and participants complete a twenty-question multiple-choice translation test for which they are allowed four minutes. Scoring is either correct or incorrect for each item; the maximum score is 24. Scores used in analysis from this task were simply calculated as number correct out of 24 possible.

**Continuous Visual Memory Task (CVMT)**

The CVMT is another measure of explicit memory (in this case, measured via a non-linguistic task) in which subjects are presented with a graphic design on a computer monitor and asked to key in whether or not they have seen the design previously (Trahan & Larrabee, 1988; Carpenter, 2008; Carpenter, Morgan-Short, & Ullman, 2009; See Figure 3).

To analyze this data, d’ scores were calculated. This score is a sensitivity index that gives a little more information than a simple report of response accuracy because it takes into account hits, misses, false alarms, and correct rejections in discrimination of stimuli. It is calculated from the hit rate and the false alarm rate (See Table 1). A d’ score of zero corresponds to chance level, while 4.65 indicates perfect ability to discriminate.

**Alternative Serial Response Task**

The alternative serial response task is measure of implicit learning in which participants must follow a pattern on the screen, which consists of four open circles filling in one at a time (See Figure 4). While there is a pattern, it is unable to be picked up on explicitly, thus learning is measured via increased accuracy and decreased reaction time.

For the Alternative Serial Response Task (ASRT, one of the measures of procedural learning), learning is measured by calculating the differences between reaction times on the last block and the first block to the patterned triplets (a series of three of the filled-in dots that is intended to reoccur throughout the blocks, which participants should pick up on implicitly) and to the random triplets (a series of three filled-in dots that is randomized, whose reaction time should in theory only vary from block 1 to block 10 because of better muscle memory and / or skill at completing the task, but not due to implicit learning). The difference calculated from these two measures was used as the participants' score for implicit learning (Howard & Howard, 1997).

**Weather Prediction Task**

In this probability-learning task, participants are asked to predict the weather (rainy or sunny) based on a set of cards that have varying probabilities to what the weather outcome will be (Knowlton, Squire, & Gluck, 1994; See Figure 5). Since here the dual-task version was used, participants are simultaneously asked to count a series of tones (low and high) that are presented randomly over blocks. The dual-task version ensures that participants can only pick up on the probability pattern implicitly, thus using the procedural memory system, by having them explicitly direct part of their attention to something else, in this case, tone counting. The last block is compared with the first block as a measure of implicit learning.

Scores from the Weather Prediction Task were not included in the analysis due to repeated difficulties in running the program on the computer, resulting in a loss of data for all but 2 participants.

**Latin Language Training**

In this study, a small, controlled portion of Latin is used as the second language taught to participants. Unlike English, Latin has subject (nominative) and object (accusative) cases for nouns, and the learning of this noun case morphology is the focus of this study. Participants are first taught Latin vocabulary needed for purposes of the study. Vocabulary presentation is timed and presented both orally and in writing, together with pictures depicting the words being learned and their English translations. For example, a picture of a monster is presented on the left side of the screen, with "*belua*" and "*beluam*", first spoken and then in writing, and on the right side, a picture of multiple monsters is presented with "*beluae*" and "*beluas*" both spoken and in writing, Finally, the word "MONSTER" in English appears at the bottom of the screen (See Figure 6). After vocabulary training, participants complete a vocabulary assessment. The training/assessment process is repeated as needed until subjects show mastery of the vocabulary, at which point the language training program allows the participants to advance. Participants are then asked to complete a computerized grammar pretest, consisting of sentence-level written and aural interpretation tasks, a grammaticality judgment task, and a sentence production task. The score on the pretest is used as a baseline for comparisons made using the scores on the posttest.

After the vocabulary training and grammar pre-test, subjects are randomly assigned to one of two groups: an explicit training group or an implicit training group (See Figure 7). Both implicit and explicit training will expose learners to aspects of Latin case morphology and verb agreement. Specifically, participants are exposed to nominative (subject) and accusative (object) case morphology (for masculine and feminine, singular and plural nouns) and to singular and plural verb conjugations.

Those in the explicit training group are given a computerized explicit explanation of these aspects of Latin, much like one would experience in a classroom setting (See Figure 8a). They are given practice questions throughout the grammar lesson. Afterwards, they practice interpreting these aspects of the language. Various tasks (aural and written) require participants to interpret Latin sentences in a multiple-choice format, focusing on the subjects and objects of sentences. In this condition, participants receive explicit, metalinguistic feedback to their responses. For example, they are given a sentence to translate such as: *Basiant dei feminas*. (Kiss$_{PL}$ gods$_{NOM\text{-}PL}$ women$_{ACC\text{-}PL}$ – ‘The gods kiss the women’) along with two possible answers: “The gods kiss the women.” and “The women kiss the gods.” Upon choosing a response, feedback is given, such as, “Good job! -i is the subject case ending; -as is the object case ending.” (See Figure 8b).

Those in the implicit training group do not receive an explicit grammar lesson with explanations of the language. Instead, they receive meaningful language input, via pictures and their corresponding Latin sentences, presented both written and aurally, similar to meaningful but un-explained input that one would experience in an immersion setting (See Figure 9a). They then undergo practice with the language identical to that of the explicit group, except that the practice only provides implicit feedback, for example, “Good job!” or “Sorry, that’s incorrect!” (See Figure 9b), with no metalinguistic information.

After the participants finish the explicit or implicit language training, they complete a language posttest similar to the pretest, consisting of the same series of computerized sentence-level tasks: written interpretation, aural interpretation, grammaticality judgment, and sentence production. Scoring on the first three tasks (which are multiple choice) is a simple tally of correct items. The fourth task was not analyzed in the current study (see below).

**Electroencephagraphy (EEG)**

ERPs will be collected as subjects read (on a computer monitor) a variety of Latin sentences with and without violations of case morphology, verbal agreement, and semantics. Sentences are presented one word at a time, so that ERPs to target stimuli may be measured, and participants respond to each sentence indicating whether the sentence is acceptable in Latin. ERP signatures exhibited in response to violations (as compared to matched correct sentences) will be examined in order to assess which ERP components, previously discussed and identified in both first and second language use, are present. No feedback will be given during the ERP testing phase.

In particular, we will use a violation paradigm to examine participants' brain responses to violations of syntactic (sentence structure / grammar-related) and semantic (related to meaning) rules of Latin. In this study, three different types of violations are used: double case violation, subject verb agreement violation, and semantic violation (See Table 2). The double case violation and subject verb agreement violation are violations that are both syntactic in nature; that is, they violate structure or grammar-related rules of Latin. For example, double case violation might be one where both nouns are in object case and there is no subject, *"Belaue visitant regina*" (roughly, "The monsters$_{\text{SUBJECT}}$ visit the queen$_{\text{SUBJECT}}$"). A subject-verb agreement violation is one that is easier to conceptualize in English; for example, "*Beluae*$_{\text{PL}}$ *visitat*$_{\text{SING}}$ *reginam*" ("The monsters$_{\text{PL}}$ visits$_{\text{SING}}$ the queen.") A semantic violation in this study is one where a picture is presented, and then the corresponding sentence does not accurately describe what is happening in said photograph. For example, a photograph of "*Beluae visitant* *reginam*" ("The monsters visit the queen") is shown, and the sentence "*Beluae visitant coquum*" ("The monsters visit the cook") is subsequently presented. Syntactically and semantically correct sentences will also be presented, as a baseline.

The stimuli examined include 48 correct syntactic sentences and 48 incorrect sentence counterparts for comparison of each type of syntactic violation, and 48 correct semantic sentences with 48 incorrect sentence counterparts, for a total of 240 sentences. Due to the relatively free word order structure of Latin, sentences are balanced so that there are 8 of each word order type: subject-verb-object, subject-object-verb, verb-subject-object, verb-object-subject, object-subject-verb, and object-verb-subject. Due to the possibility of wrap-up effects, the sentences are also balanced for which position the target word is in, second or third position; the target word never occurs in the first position as this could never create any of the types of violation examined here. Since ERP data collection is very sensitive to noise from movements like blinking or jaw-clenching, the beginning of each block contains 5 "practice" sentences, whose ERP data is not analyzed, in order to accustom the subjects to the rhythm of the ERP collection.

**Analysis**

We use repeated-measures ANOVAs to examine language learning resulting from the different training conditions (pre-test and post-test for each task in order to explore effects of training and effects by group. Note that we decided not to analyze the fourth task -- sentence production -- because the task itself and its scoring are more complex.

Pearson correlations between scores on the cognitive tasks and the effects of Latin language training are investigated to see if any correlations exist between strengths of declarative and procedural memory, as measured by the tasks, and learning during the Latin Language

Training. These are run across all participants as well as separately by each group in order to see if any effects are present across all participants or within any training group.

## Results

Descriptive statistics for Latin learning are shown in Table 3 and Graphs 1, 2, and 3. There was found to be no significant main effect of Time or Group, and no significant Time by Group interaction (all *p*s > 0.188). Thus, there was no evidence of learning over time across the two groups, and no differences in learning across time between the two groups.

Correlations between gains on the Latin language training (Posttest score – pretest score for each task) and scores on each of the cognitive tasks were run using SPSS. No correlation between scores was shown to be significant (all *p*s > 0.112).

## Discussion and Conclusions

Since no significant learning was observed in the Latin language learning according to the ANOVAs by time, running subjects in the ERP grammaticality judgments would not yield any meaningful data. Thus, from this pilot data collection, we learned that participants may need more than one session with the Latin language training in order to make any significant strides towards language proficiency before ERPs could be run to see if the signatures varied by pedagogical condition. However, this result was still surprising since previous studies have shown significant results using the same computer administered training. The participant populations in these previous studies, however, was greater than that of the present study (all *n*s per group > 11; see Sanz et al. 2009, Stafford, Sanz, & Bowden 2010, and Stafford, Bowden and Sanz 2012). Thus, for future studies, an a priori power analysis may be performed in order to measure how many participants per group would be needed in order to yield significant data.

Because these trends were not significant given the sample size, there are few conclusions to be drawn from the data. However, if these trends shown in graphs 1, 2, and 3 were significant, more interpretations would need to be extrapolated. For example, the implicit group showed greater improvements on the written interpretation task than the grammaticality judgment task as compared to the explicit group, while both improved on the aural interpretation task. If these effects were borne out in a larger sample size and found significant, it would make sense that the explicit group showed greater improvement on grammaticality judgment, as they are given explicit metalinguistic feedback informing them of the grammar of the language, which would be expected to lead to greater behavioral gains, at least short term (Norris & Ortega, 2000). Explanations for the differences found on the written and aural interpretation tasks might be more difficult to derive from the given data.

The future project for which the present study was a pilot would likely include more sessions for the participants to interact with the language, giving them both more exposure and more practice to better their proficiency. However, it may be interesting to record ERPs before necessarily seeing any behavioral difference, as a neural effect may be able to be observed without a significant behavioral effect. Analyzing the results in this way could also help us really define what it is that we mean when we say "proficient," and how it is that we measure it. Because we are able to measure learning both by behavioral scores and by variations in ERP signatures, it may be useful to look at the interaction between the two when discussing second language proficiency. What makes native-like ERP signatures "better"? Conversely, what makes high behavioral scores "better"? For any language acquisition study, it will be necessary to consider exactly what our definition of proficiency really is. Any results arising from these further explorations could lead to further discussion about the neural mechanisms of second language acquisition.

The lack of significant interactions between the scores on the cognitive tasks and the gains in Latin Language Training could be due to a number of factors; it is likely due in part to the effect of small sample size (N=9), and a pattern may emerge with increasing numbers of participants. In addition, with more Latin training, significant differences may emerge. Specifically, with the Dual-Task Version of the Weather Prediction Task, it may be beneficial to ensure that before beginning the task, participants are able to discriminate between low and high tones, and that the presentation volume throughout any task with an aural component, including the Latin tasks, be calibrated and held constant in attempts to eliminate this as a variable in the future. And although we did not have reason to believe this to be an issue in the collection of data, we consider it to be a limitation of the present study due to the fact that it was not carefully measured.

Further, though handedness was controlled for, as mentioned previously, by a simple questionnaire, it may be necessary to utilize a more sophisticated scale of handedness, such as the Edinburgh Handedness Inventory, which gives a more nuanced handedness score and may be useful in further controlling for the role of handedness in language acquisition ERPs.

Language experience and background may also play a role in Latin language learning. While we ensured that participants had no more than two years of a college level language, experience in high school or any experience up to the two-year marker in college level language may affect the outcomes of the Latin tasks, especially if that language experience is with a romance language. Future studies may be able to explore this question by controlling language experience by group in order to examine any behavioral or neural effects.

While the results of this pilot study were in some ways inconclusive, they do set the stage for the continuation of this line of research. In further investigation with this study, it will need to be ensured that participants are learning the material during the Latin language training, by either running participants through the lessons or practice multiple times to increase amount of exposure or setting up the explicit and implicit conditions differently to make the sessions more effective. Implications of this study may affect the way we approach second language acquisition and pedagogy, and for that reason are important to pursue further, in order to be able to learn and teach more effectively and foster better communication among us all.

**Graphs and Figures**

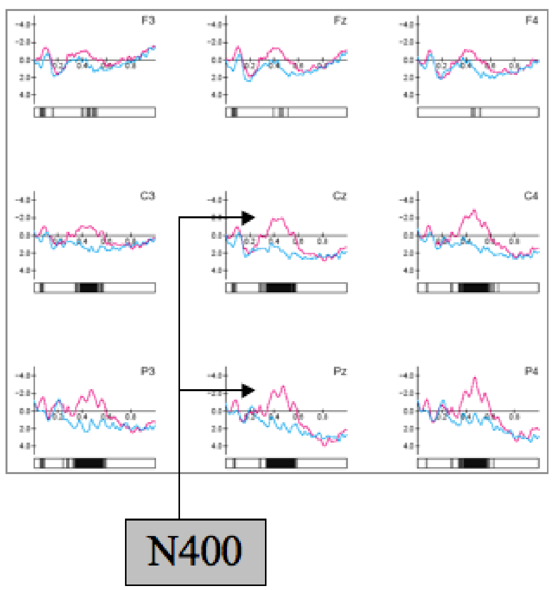



*Figure 1.* ERP components, N400, elicited during semantic and lexical violations. (Figure from Bowden et al. 2007)

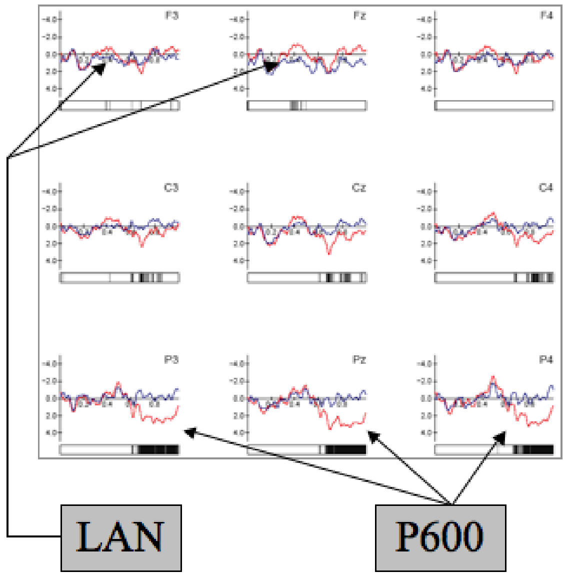



*Figure 2*. ERP components, LAN and P600, elicited during syntactic and grammatical violations. (Figure from Bowden et al. 2007).

*Figure 3*. Examples designs encountered on the Continuous Visual Memory Task

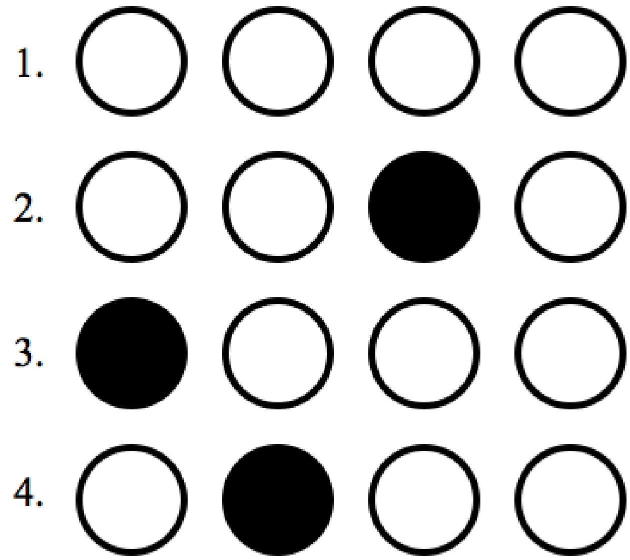



*Figure 4*. An example of the starting screen (1) and a triplet (2, 3, 4) on the Alternative Serial Response Task.

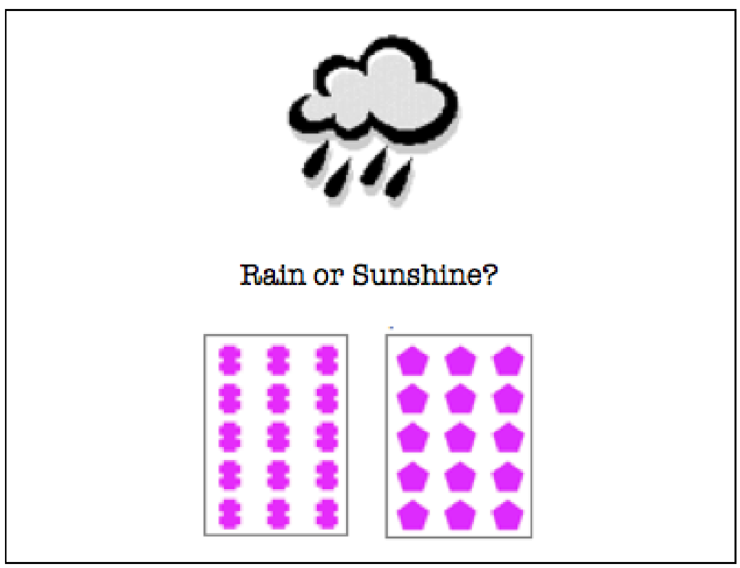



*Figure 5*. An example of a question asked on the Weather Prediction Task.


*Figure 6*. Screenshot from the Latin Language Training vocabulary lesson.

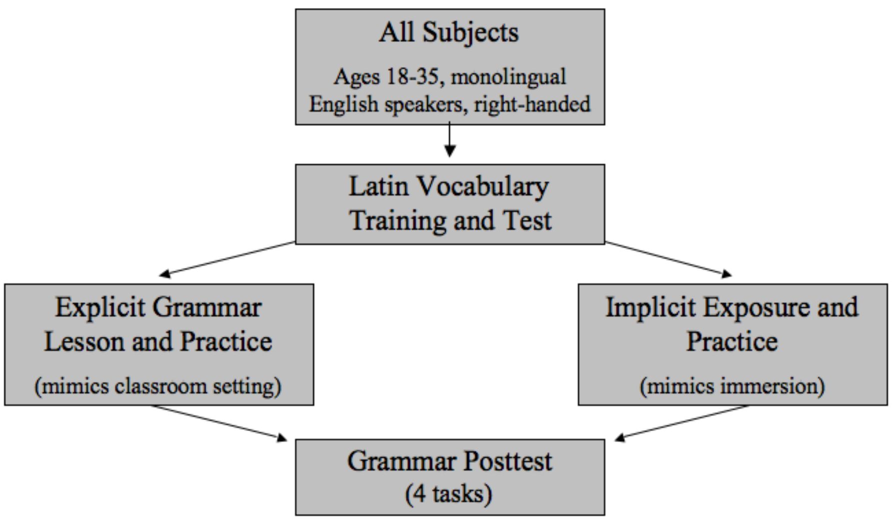



*Figure 7*. Latin Language Training Set Up.

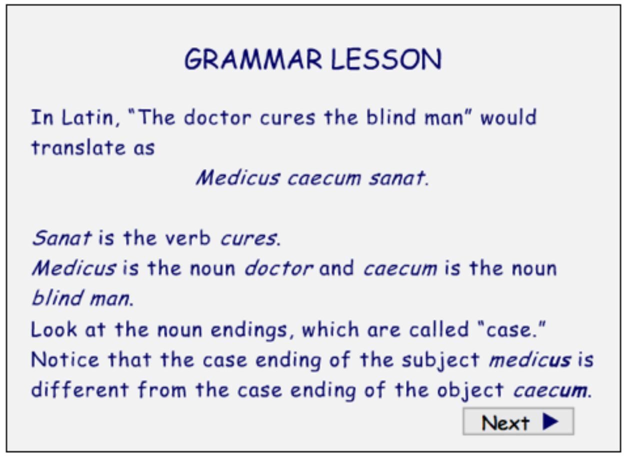



*Figure 8a*. Screenshot from the grammar lesson used in the explicit group.

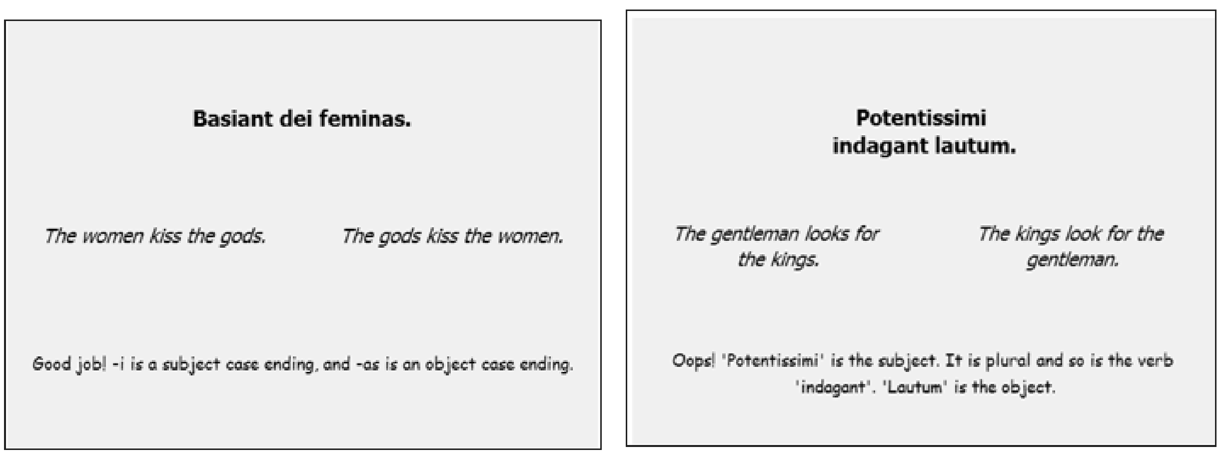



*Figure 8b*. Screenshots from the practice of the explicit group with explicit feedback.


*Figure 9a*. Screenshot from the exposure PowerPoint used with the implicit group.

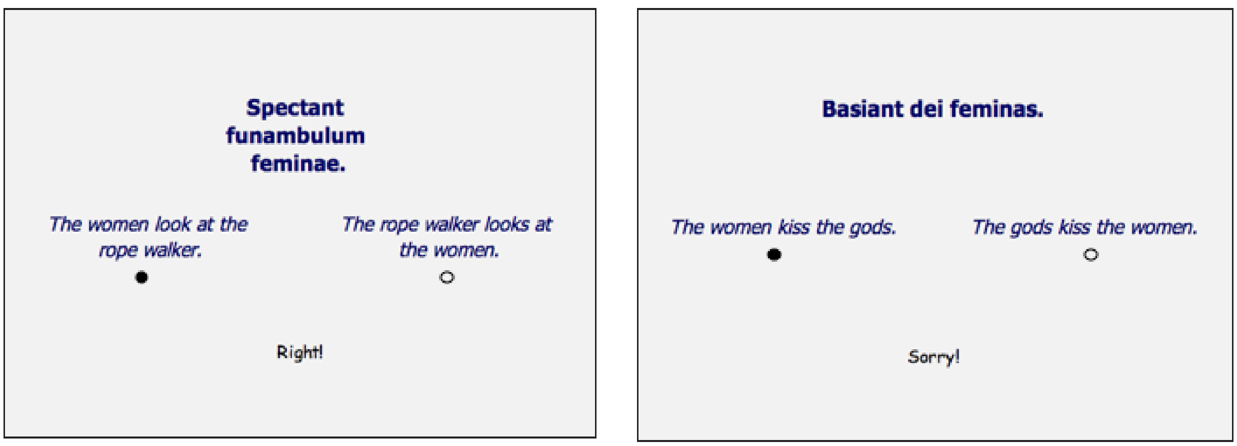



*Figure 9b*. Screenshots from the practice of the implicit group with implicit feedback.

Table 1

*Possible Responses on the Continuous Visual Memory Task*

| | Response: Old | Response: New |
|---|---|---|
| Stimuli: Old | Hit | Miss |
| Stimuli: New | False Alarm | Correct Rejection |

Table 2

*ERP Testing: Types of Violations*

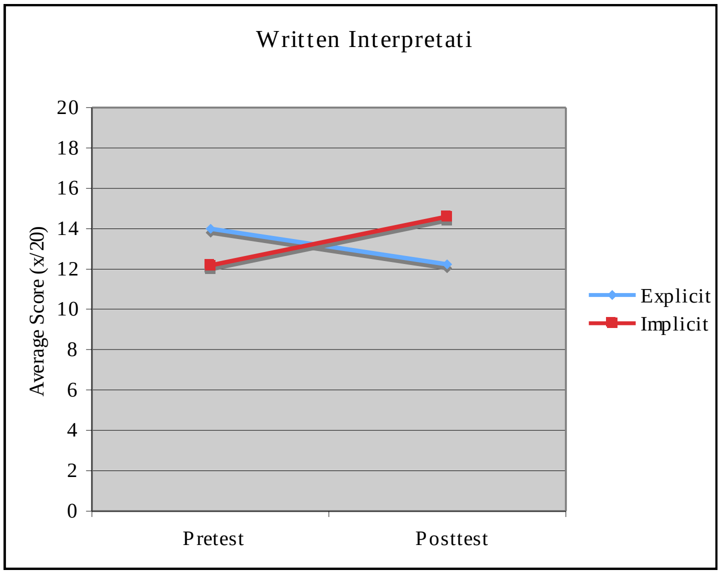
| | Semantic Violation | Syntactic Violation: Double Case | Syntactic Violation: Subject-Verb Agreement |
|---|---|---|---|
| Correct | Beluae visitant reginam<br>Monsters visit queen<br>'The monsters visit the queen' | Beluae$_{NOM}$ visitant reginam$_{ACC}$.<br>Monsters$_{NOM}$ visit queen$_{ACC}$<br>'The monsters visit the queen' | Beluae$_{PL}$ visitant$_{PL}$ reginam.<br>Monsters$_{PL}$ visit$_{SING}$ queen<br>'The monsters visit the queen' |
| Incorrect | Beluae visitant #coquum.<br>Monsters visit cook<br>'The monsters visit the #cook.' | Beluae$_{NOM}$ visitant *regina$_{NOM}$.<br>Monsters$_{NOM}$ visit queen$_{NOM}$<br>'The monsters$_{NOM}$ visit the queen$_{NOM}$.' | Beluae$_{PL}$ *visitat$_{SING}$ reginam.<br>Monsters$_{PL}$ *visits$_{SING}$ queen<br>'The monsters$_{PL}$ *visits$_{SING}$ the queen." |
| Presented with photo |  * | N/A | N/A |

*Note*: All target words, correct and incorrect, are underlined. A pound or number sign (#) indicated a general anomaly; an asterisk (*) indicates a syntactic error. The subscripts NOM and ACC stand for nominative case and accusative case, respectively, while the subscripts PL and SING stand for plural and singular, respectively.
*While this use of a photograph to prime a semantic violation is novel and experimental, pictures have been used in previous ERP studies with successful results (Ganis, 1996).

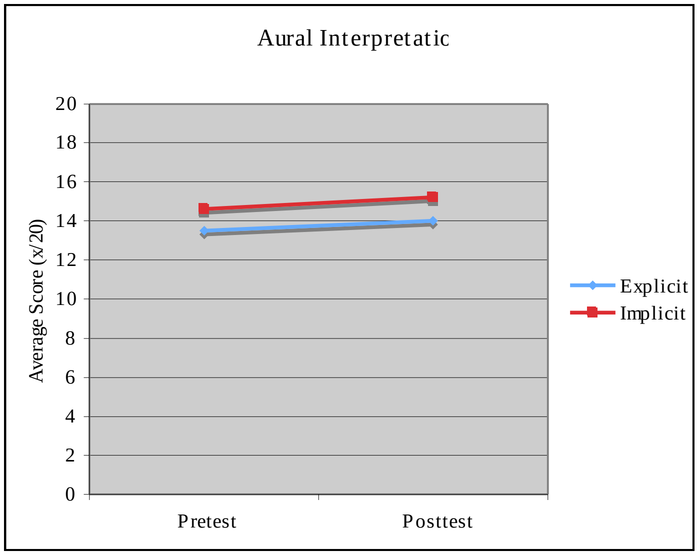
Table 3

*Results, Descriptive Statistics of Latin Language Learning*

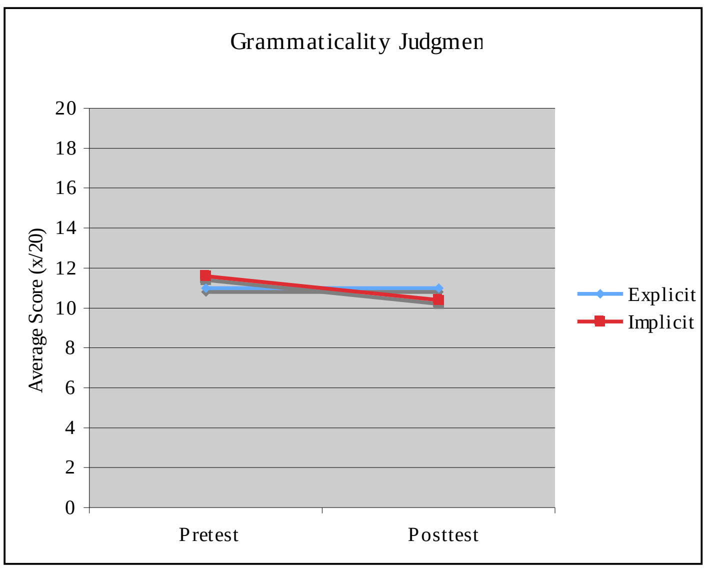
| Task | Group | Pretest Mean | Pretest Std. Deviation | Posttest Mean | Posttest Std. Deviation | Mean Gain | N |
|---|---|---|---|---|---|---|---|
| Written Interpretation asks participants to read a sentence and choose between two pictures | Explicit | 14.00 | 1.826 | 12.25 | 2.217 | -1.75 | 4 |
| | Implicit | 12.20 | 2.168 | 14.60 | 5.273 | 2.40 | 5 |
| | All | 13.00 | 2.121 | 13.56 | 4.157 | 0.56 | 9 |
| Aural Interpretation like Task 1, but participants listen to the sentence rather than reading it | Explicit | 13.50 | 2.646 | 14.00 | 2.000 | 0.50 | 4 |
| | Implicit | 14.60 | 2.191 | 15.20 | 4.207 | 0.60 | 5 |
| | All | 14.11 | 2.315 | 14.67 | 3.279 | 0.56 | 9 |
| Grammaticality Judgment asks subjects to decide whether a sentence is right or wrong | Explicit | 11.00 | 2.944 | 11.00 | 2.160 | 0.00 | 4 |
| | Implicit | 11.60 | 1.517 | 10.40 | 2.191 | -1.20 | 5 |
| | All | 11.33 | 2.121 | 10.67 | 2.062 | -0.66 | 9 |




*Graph 1.* Averages by group on the written interpretation task during Latin Language Training.




*Graph 2.* Averages by group on the aural interpretation task during Latin Language Training.




*Graph 3.* Averages by group on grammaticality judgment task during Latin Language Training.

**Acknowledgements**

Thank you to the College Scholars Program and Dr. Jeffrey Kovac for supporting this project, and also to Dr. Harriet Bowden for her continued guidance in the conceptualization and implementation of this project.

References

Barber, H. a. C., M. (2005). Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience, 17*(1), 137-153. doi: 10.1162/0898929052880101Bates, E., & Elman, J. (1996). Learning rediscovered. *Science, 274*(5294), 1849-1850.

Bowden, H.W. (2007). Proficiency and second-language neurocognition: A study of Spanish as a first and second language. Thesis, PhD; Georgetown University.

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

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

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

Coulson, K., and Kutas. (1998). Expect the Unexpected: Event-related Brain Response to Morphosyntactic Violations. *Language and Cognitive Processes, 13*(1), 21-58. doi: 10.1080/016909698386582Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn Sci, 6*(2), 78-84.

Elman, J. L. (1998). *Rethinking innateness: A connectionist perspective on development* (Vol. 10). The MIT press.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proc Natl Acad Sci U S A, 99*(1), 529-534. doi: 10.1073/pnas.012611199

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. [Research Support, Non-U.S. Gov't]. *Brain Res Cogn Brain Res, 1*(3), 183-192.

Ganis, K., Sereno. (1996). The Search for "Common Sense": An Electrophysiological Study of the Comprehension of Words and Pictures in Reading. *Journal of Cognitive Neuroscience, 8*(2), 89-106.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. [Comparative Study Research Support, Non-U.S. Gov't]. *J Psycholinguist Res, 30*(3), 251-266.

Hahne, A. and Friederici, A. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. Bilingualism: Language and Cognition, 4, pp 123-141. doi:10.1017/S1366728901000232.

Howard, J. H., & Howard, D. V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. Psychology and Aging, 12(4), 634-656.

Knecht, S.; Dräger, B.; Deppe, M.; Bobe, L.; Lohmann, H.; Flöel, A.; Ringelstein, E. B.; Henningsen, H. (2000). "Handedness and hemispheric language dominance in healthy humans". *Brain* **123** (12): 2512–2518.

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

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*(4), 933-947.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative metaanalysis. Language Learning, 50, 417-528.

Ojima et al. (2005). An ERP Study of Second Language Learning after Childhood: Effects of Proficiency. *Journal of Cognitive Neuroscience, 17*(8), 1212-1228. doi: 10.1162/0898929055002436

Penke, M., Weyerts, H., Gross, M., Zander, E., Munte, T. F., & Clahsen, H. (1997). How the brain processes complex words: an event-related potential study of German verb inflections. [Research Support, Non-U.S. Gov't]. *Brain Res Cogn Brain Res, 6*(1), 37-52.

Rodriguez-Fornells, A., Clahsen, H., Lleo, C., Zaake, W., & Munte, T. F. (2001). Event-related brain responses to morphological violations in Catalan. [Clinical Trial Research Support, Non-U.S. Gov't]. *Brain Res Cogn Brain Res, 11*(1), 47-58.

Sanz, C. e. a. (2009). Concurrent Verbalizations: Pedagogical Conditions, and Reactivity: Two CALL Studies. *Language Learning, 59*(1), 33-71.

Seidenberg, M. S. (1997). Language acquisition and use: learning and applying probabilistic constraints. [Research Support, U.S. Gov't, P.H.S. Review]. *Science, 275*(5306), 1599-1603.

Stafford, C. A., Bowden, H. W. and Sanz, C. (2012), Optimizing Language Instruction: Matters of Explicitness, Practice, and Cue Learning. Language Learning, 62: 741–768. doi: 10.1111/j.1467-9922.2011.00648.x

Stafford, Catherine A., Sanz, Cristina and Bowden, Harriet Wood(2010) 'An experimental study of early L3 development: age, bilingualism and classroom exposure', International Journal of Multilingualism, 7: 2, 162 —183, First published on: 21 April 2010 (iFirst)

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

Ullman, M. T. (2001). A neurocognitive perspective on language: the declarative/procedural model. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S. Review]. *Nat Rev Neurosci, 2*(10), 717-726. doi: 10.1038/35094573

Ullman, M. T. (2005). A Cognitive Neuroscience Perspective on Second Language Acquisition: The Declarative/Procedural Model. In C. Sanz (Ed.), *Mind and Context in Adult Second Language Acquisition: Methods, Theory, and Practice* (pp. 141-178). Washington, DC: Georgetown University Press.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S. Review]. *Cognition, 92*(1-2), 231-270. doi: 10.1016/j.cognition.2003.10.008

Weber-Fox, C. M. a. N., Helen J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience, 8*(3), 231-256. doi: 10.1162/jocn.1996.8.3.231

Weber-Fox, C. M. a. N., Helen J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience, 8*(3), 231-256. doi: 10.1162/jocn.1996.8.3.231

Weyerts, H., Penke, M., Dohrn, U., Clahsen, H., & Munte, T. F. (1997). Brain potentials indicate differences between regular and irregular German plurals. [Research Support, Non-U.S. Gov't]. *Neuroreport*, *8*(4), 957-962.