EFFECTS OF PROCEDURALIZATION OF AN L2 ON COGNITIVE ABILITIES: LOOKING FOR THE THRESHOLD OF BILINGUAL BENEFITS

Laura Babcock

Scuola Internazionale Superiore di Studi Avanzati

Elizabeth Krawczyk

Georgetown University

Jeffrey Scialabba

Ayn Rand Institute

Abstract:

Much recent research in the fields of SLA and Bilingualism has focused on the social, linguistic, and cognitive benefits of bilingualism (Cook, 1997; Bialystok, 2001; Bialystok, 2004; Sanz, 2000; Sanz, 2007). This research has sought to establish the nature of these benefits and the point at which they emerge. Cummins' Threshold Theory posits that these benefits are determined by level of proficiency in both languages (Cummins,

1976). This threshold, however, has not been adequately operationalized. Based on the Declarative/Procedural model in L2 speakers (Ullman, 2001a; Ullman, 2005), we hypothesized that the onset of use of procedural memory in both languages marks the emergence of cognitive benefits for bilinguals. In a preliminary attempt to investigate this effect, we measured the verbal and non-verbal memory of participants before and after learning an artificial language to high proficiency. ERP measures were used to determine reliance on procedural memory during L2 language processing. The results indicate that the use of procedural memory during L2 language processing may affect non-verbal memory measures; no effects were found for verbal memory measures. Contrary to the hypothesis, however, the participants who used procedural memory showed lower scores on these measures than those who did not use procedural memory. These results suggest that procedural memory and non-verbal memory may be related.

Keywords: bilingualism, cognitive processes, brain.

Introduction

The history of research on bilinguals and their cognitive abilities runs the gamut from claims of detrimental effects to superior abilities in multiple areas. Recently this area of research has received much attention in the fields of Second Language Acquisition (SLA) and Bilingualism. The results continue to be varied, however, as many studies have found positive cognitive effects for bilinguals, especially in the areas of control processes (Bialystok, 2005) and the ability to distinguish meaning and form (Bialystok, 1988; Cummins, 1978; Edwards and Christopherson, 1988; Eviatar and Ibrahim, 2000; Feldmen and Shen, 1971). One proposal for understanding this varied evidence is Cummins' Threshold Theory (Cummins, 1976), which posits that the effects of bilingualism on cognitive abilities is dependent on proficiency in the two languages. Specifically,

he suggests that bilinguals need to have "age appropriate" levels of proficiency in both languages to reap the cognitive benefits of bilingualism. Cummins, however, does not adequately define "age appropriate" levels and several researchers have called for research which operationalizes this threshold (Sanz, 2007; Lado, 2006).

In addition to the question of enhanced cognitive abilities, the fields of SLA and Bilingualism have become increasingly interested in the similarities between the benefits seen in bilinguals and experts in other fields, such as chess and physics (de Groot, 1965; Heller & Reif, 1984). Some researchers have gone as far as labeling bilinguals as experts in the field of language learning (Nayak et al., 1990, Ramsay, 1980). This line of research opens the door to better understanding the benefits bilinguals receive from their expertise in languages. Combining this approach with Cummins' Threshold Theory may lead to advances in understanding the benefits seen in bilinguals.

Background

Cognitive Effects of Bilingualism

A number of studies have shown advantages for bilinguals in areas ranging from control processes to divergent thinking. Here, we will review some of these findings. Landry (1974) found an advantage for bilinguals on standard tests of divergent thinking, which consider measures of flexibility, originality, and fluency; however, these effects were only seen after 6th grade. Higher scores were also found on the "unusual uses" test of creativity in bilingual children compared to monolingual children by Lambert, Tucker, and d'Anglejan (1973). Based on tests of verbal and non-verbal IQ, Peal and Lambert (1962) posited that bilinguals have a more diversified set of mental abilities. Diaz (1985) also found that bilinguals are more creative, as

well as have advantages in conceptual development and analogical reasoning.

In addition to these studies, Bialystok (2005) found that bilinguals performed better on a Simon task--which assesses control processes--than did monolinguals, at most age groups. Specifically, bilinguals were better able to ignore non-relevant information and focus on the task at hand. This advantage was found for children as well as older adults, suggesting that the effect is not limited to children, as many of the studies mentioned above might suggest. In another experiment, which focused on children, Bialystok (1988) found that bilingual children were more willing to accept the possibility of switching the names of the sun and the moon and were more able to play along with the switching game than monolingual children. Similar evidence of an increased ability to recognize the arbitrariness of a name has been found in bilingual children by others as well (Cummins, 1978; Edwards and Christopherson, 1988; Eviatar and Ibrahim, 2000; Feldmen and Shen, 1971).

Negative cognitive effects have also been found in bilinguals. Ransdell and Fischler (1987) found that bilinguals were slower at data-driven tasks such as list recognition and lexical decision. Edwards and Christopherson (1988) found that bilinguals performed worse on Grammaticality Judgment Tests (GJTs) than monolinguals. Critically, it is not clear if these studies controlled for proficiency level in both languages.

In fact, more recent research has found that benefits were related to the level of proficiency in the second language. Ricciardelli (1992) found that children with a high level of proficiency in two languages performed better on a GJT than those with a high level of proficiency in one and a low level in the other, who in turn performed better than those with low levels of proficiency in both languages. Gathercole

(1997) also found that performance on GJTs was dependent on proficiency level. Segalowitz and Frenkiel-Fishman (2005) found that the variance in speed of attention control in bilinguals was accounted for by their level of proficiency in the language of testing.

All of this seemingly contradictory evidence may perhaps be unified under a theory such as Cummins' Threshold Theory (Cummins, 1976). The Threshold Theory posits that there are three levels of bilingualism, with a threshold between each level. At the lowest level of bilingualism, children have low levels of proficiency in both languages and can be expected to show negative cognitive effects. At the second level, children have age appropriate skills in only one language and show neither positive nor negative cognitive effects. At the third, and highest, level, children exhibit age appropriate skills in both languages and can be expected to reap positive cognitive effects. While this theory is accepted by many, problems do exist with it. First, the theory was based on child bilingualism, not adult bilingualism; however, some suggest that the model can be extended to adulthood bilingualism (Lado, 2006). The second and larger issue is the lack of operationalization of the thresholds, in particular the second threshold. Both Sanz (2007) and Lado (2006) have called for research examining this threshold. One hypothesis is that biliteracy represents the upper threshold (Sanz, 2007). The hypothesis presented in this paper suggests that the literature on experts and novices may help define the threshold for cognitive benefits in bilinguals. Specifically, we posit that expert language learners have attained the highest level of bilingualism and have done so utilizing the same processes as experts in other fields.

Bilinguals as Experts

A large literature exists on the differences between novices and experts in a number of fields, most notably the game of chess

(de Groot, 1965; de Groot & Gobet, 1996). In these fields experts are defined by their better memory, better and more elaborated problem representations, and different problem-solving strategies. Additionally, their expertise is based on knowledge, not basic capacity, and they became experts through extensive practice (Green & Gilhooly, 1992). Many researchers in this field have suggested ways in which novices become experts, one suggestion, (the ACT theory) made by Anderson (1993), is that some sort of proceduralization of the knowledge has occurred.

The Atomic Components of Thought (ACT) Theory is a domain-general rule-based cognitive theory which describes the processes underlying skill acquisition (Anderson, 1993). The ACT theory posits two types of memory/knowledge: declarative memory/knowledge, used to store facts, and procedural memory/knowledge, which consists of production rules. The facts stored in declarative memory/knowledge are available for reporting while the production rules of procedural memory/knowledge are manifested only in performance (Anderson, 1993). Additionally, these two types of knowledge have been tied to particular brain structures: declarative knowledge is tied to the temporal lobes, hippocampus, and ventrolateral prefrontal cortex and procedural knowledge is tied to the basal ganglia and associated regions (Anderson et al., 2004). The ACT theory further claims that knowledge is initially declarative based, but becomes procedural based through a process of compilation. Once this knowledge is proceduralized, practice leads to automation. Crucially, two fields have utilized this model of psychological processes: the expert/novice literature discussed above, and SLA.

DeKeyser (2001) posits that ACT theory accurately describes the process by which adults learn a second language. Thus, language learners and experts in other fields appear to be using the same

process to develop and store their knowledge. Due to perhaps this similarity and others, bilinguals have been increasingly compared to experts in other fields and therefore have been considered experts in the field of language learning. These similarities include better strategy usage (Ramsay, 1980; Nayak et al., 1990) and the ability to recall *pertinent* knowledge (Bransford, 1999; Bialystok, 2005). Thus, if we take the connection between experts and bilinguals to be true, and adhere to the ACT theory, which suggests that proceduralization is what allows experts to perform better than novices, then it follows that bilinguals perform better than monolinguals in some areas, particularly the ability to identify pertinent information, due to proceduralization of the relevant knowledge (i.e. both languages). The question of how to assess proceduralization of the language in these individuals then arises.

Neurocognitive Aspects

The Declarative/Procedural (DP) model is a neurocognitive complement to the ACT theory which specifically addresses language. The DP model posits that the mental lexicon and grammar are stored in two separate, domain-general memory systems (Ullman, 2001b, 2004; Ullman et al., 1997): the declarative memory system and the procedural memory system. These memory systems have been studied extensively in humans and animals in various domains. The declarative memory system houses the mental lexicon as well as knowledge about facts and events and relies on the temporal lobes. The procedural memory system underlies mental grammar, implicit learning, and motor and cognitive skills. This system depends on frontal lobe and basal ganglia structures. The two memory systems interact both *cooperatively*, working together in the acquisition and use of knowledge, and *competitively*, in that enhanced learning in one

system may lead to suppression of the other, leading to what Ullman calls a "see-saw effect." For example, damage to the hippocampus, a declarative memory structure, can lead to enhanced basal ganglia operation; conversely damage to the basal ganglia has been found to enhance declarative memory (Ullman, 2004).

The DP model proposes that language processing is primarily a function of the interaction between these two memory systems. In native language processing, the declarative memory system is posited to underlie the mental "lexicon", storing arbitrary word-form information and underlying the learning of new words, while the procedural memory system is posited to subserve the rule-based, compositional domains of language, such as syntax and morphology. The DP model also makes specific predictions regarding second language acquisition, when the acquisition occurs after childhood. Initially, all learning and use of a second language is expected to occur in declarative memory, including not only the mental lexicon but also the formation and execution of explicit grammatical rules. With increased practice and proficiency, however, the computation of the compositional aspects of language shifts from the declarative memory system to the procedural memory system (proceduralization). *At this point, learners are native-like not only in proficiency but also in their underlying dependence on the two memory systems.* Both behavioral (Babcock et al., in prep) and ERP data (Morgan-Short, 2007) suggest that this characterization of high proficiency learners is accurate.

The benefit of this model is that it bases proceduralization on memory system usage which can be tested through various methods, including neuroimaging and highly sensitive behavioral measures. Importantly, this attention to the underlying brain mechanisms allows for identification of L2 learners who not only exhibit L1-like behavior, but also native-like processes.

Research Questions

By utilizing theories from several fields, a possible understanding of Cummins' threshold for cognitive benefits in bilinguals can be reached. In particular, the literature on experts and novices lends a possible definition of a threshold to enhanced performance that can be extended to bilinguals given the similarities between them and experts. Cognitive neuroscience adds an understanding of the specific brain structures involved in the process of proceduralization. Thus, the research questions for this study are:

1. Is proceduralization of both languages the threshold needed for cognitive benefits?
2. Will learners who have proceduralized an additional language show enhanced cognitive abilities as compared to those who did not proceduralize the language?

It is hypothesized that proceduralization is related to Cummins' threshold and that learners who have proceduralized an additional language will show benefits on measures of cognitive ability.

Methods

Research Design

This study investigated whether proceduralization of a second language is the threshold beyond which benefits are seen in bilinguals. Making use of an ongoing research project (Morgan-Short, 2007) behavioral and neurocognitive evidence was used to address the research question. Testing occurred in three phases. The first phase consisted of four cognitive tests and a background questionnaire. In the second phase participants learned an artificial

language, BROCANTO2, in three sessions each no more than 5 days apart. During the final BROCANTO2 learning session, once participants had reached high proficiency in BROCANTO2, ERP data were collected while participants listened to and judged correct and incorrect BROCANTO2 sentences. The final phase of this study assessed retention of BROCANTO2 and occurred 3 to 6 months later. Participants received retraining on BROCANTO2 after which ERP data were collected as previously described. Additionally, participants were retested on the cognitive tests they completed in the first phase of this study.

Participants

We tested 42 native English speaking healthy adults between 18 and 40 years of age. All participants were given a pre-screening questionnaire, and, if they met the criteria for the study, a full questionnaire asking for a range of detailed information, including age, education, handedness, medical history, and language background. All participants were right-handed (Oldfield 1971), had no history of drug or alcohol dependence, had no known personal or family history of neurological, psychiatric, or learning disorders, and had normal hearing and normal or corrected to normal vision. At the time of the study, they were all enrolled in college or had completed at least 4 years of college study. Further, due to the nature of the study and that the artificial language learned was similar to Romance languages, no participant had ever been fluent in a language other than English and exposure to Romance languages was limited in the following ways: (a) Participants could not have studied any Romance language for more than one year in college and for more than three years total, (b) all formal exposure to Romance languages must have occurred at least two years prior to study participation,

and (c) participants must not have lived for more than two weeks in a Romance language immersion environment. All subjects gave informed consent, and were paid for their participation. At the end of the study only 21 participants completed all three phases of the study and 2 of these participants had to be excluded from statistical analyses because of the large number of artifacts in their ERP data. Thus, data from 19 participants were used in the statistical analyses.

Materials

Four measures of cognitive ability, Weather Prediction task, CVMT, MLAT Paired Associates, and CVLT-II, were used to gauge varying types of aptitude and memory before learning BROCANTO2 and at the retention phase. The Weather Prediction (WP) task measures non-linguistic procedural learning using a dual task paradigm which increases the likelihood of implicit learning (Foerde, Poldrack et al. in press). Participants are asked to predict the weather (rain or sun) based on a grouping of one, two, or three cards. Feedback is given after a response is made; however, the design is probabilistic, thus a given grouping of cards does not indicate one weather condition 100% of the time, but rather between 10% and 90% of the time. The probabilistic nature of the task lends to increased reliance on implicit learning. In addition to this primary task, participants hear a series of high and low tones throughout the task and are asked to count the number of high tones they hear. Periodically they are asked to report the number of high tones they have heard. The purpose of this distracter task is to cause participants to process the weather prediction information implicitly. Participants received eight blocks of 40 items each in the training phase of this task, after the training phase a ninth block was presented and used for testing. This block had no tones to count nor was feedback given to the participants'

responses. After the final block a questionnaire assessing explicit knowledge of the weather predictions was given. Prior to learning BROCANTO2 the full task was given to participants, however at retention only the testing block and questionnaire were given.

The CVMT (Continuous Visual Memory Test) tests non-verbal declarative learning using a series of grouped images. In each group of similar images one image occurs seven times, whereas the others occur only once. Additionally, there are some images which occur only once and belong to no group. Participants are shown the images and asked to reply "new" or "old" to each image. Thirty minutes after this training phase participants were shown all images from one group and asked to identify the image presented most frequently. They also completed a visual discrimination test to verify that they could differentiate between the images. During the retention phase participants were only shown the groups of images and asked to identify the one image that had been presented most during the training they received before learning BROCANTO2.

The MLAT (Modern Language Aptitude Test) Paired Associates is a verbal declarative memory task. In the task participants are given a list of 24 foreign words with their English counterparts and are given two minutes to memorize the list. After the two minutes the list is removed and participants are given a five-way multiple choice test to assess their knowledge. When completed during the retention phase only the multiple choice test was given, not the original list of paired words.

The CVLT-II (California Verbal Learning Test-II) is also a test of verbal declarative memory, but additionally assesses semantic and serial clustering. Participants are read a list of 16 words, which belong to 4 semantic categories, and are then asked to recall as many words as they can. This exact procedure repeats four times with the

same list of words. An interference list of 16 words in 4 semantic categories, two of which overlap with the first list, is then read and participants are asked to recall this list. Immediately following, participants are asked to recall the first list using three methods: free recall, cued recall using semantic category, and yes/no recognition where items listed are from both lists and distracter items. Finally, participants are asked to recall using the same three methods above after a delay of 20 minutes. During the retention phase the list was not repeated and testing followed the three methods model.

This study used BROCANTO2, an artificial language based on BROCANTO (Friederici et al., 2002; Opitz & Friederici, 2002; Opitz & Friederici, 2003), which is a fully productive artificial language that adheres to the universal requirements of natural languages. BROCANTO has a limited lexicon of phonologically feasible words in German and syntactic rules similar to English. BROCANTO2 differs in that the lexicon is phonologically feasible in English and the grammar resembles Romance languages, particularly Spanish. These changes reduce and control for effects of L1 transfer in both phonology and syntax.

An artificial language, rather than an existing natural language, was chosen for use in this study for a number of important reasons. First, the objective of the study was to look at effects of proceduralization of a second language, but as mentioned above proceduralization of a full language is time consuming. Thus for feasibility reasons the limited system of an artificial language was desirable. Additionally, previous studies using BROCANTO, the language upon which BROCANTO2 is based, found brain activity typical of natural language processing (Friederici et al., 2002; Opitz & Friederici, 2002; Opitz & Friederici, 2003). Second, as mentioned above, phonological differences between the L1 and testing language

could be minimized. This is important because difficulties with a new phonological system could be confounded with the learning of the lexicon and grammatical system. Third, syntactic differences between the L1 and testing language could also be controlled so that effects of transfer could be systematically examined. Finally, constraints regarding stimulus material exist when using ERP measures (e.g. acoustically identical baseline periods and no co-articulation across word boundaries). The design of the artificial language could accommodate these constraints, whereas a natural language could not as easily.

The lexicon of BROCANTO2 consisted of 13 nonce words (see Appendix I), which were confirmed to be nonce words in English and did not exhibit phonological difficulties for native English speakers (Morgan-Short, 2007). Each item in the lexicon belonged to one of five classes of words: nouns, adjectives, determiners, verbs, and adverbs. Further, each noun was classified as either masculine or feminine, adjectives and determiners had both masculine and feminine forms, and verbs were classified as transitive, intransitive, or both.

The grammar of BROCANTO2 is based on universal requirements of natural language and is fully productive. It exhibits a fixed subject-object-verb order of verbal phrases, which show no morphological features. The noun phrase, however, displays agreement between determiner, adjective, and noun based on morphological markers. Additionally, the noun phrase follows a noun-adjective-determiner order. When present, adverbs appear after the verb. This grammar is unlike English in many respects; however, it is similar to structures found in Romance languages. The post-nominal determiner while not widely found in Romance languages, does occur in Romanian (Mallison, 1986) as well as some non-Romance languages such as

Basque (Hualde & Ortiz de Urbina, 2003), Malay (Lewis, 1969, as cited in Morgan-Short, 2007) and Zapotec (Peñafiel, 1981, as cited in Morgan-Short, 2007).

Participants learned BROCANTO2 through a computer board game. The game is similar to chess in some ways. Abstract tokens, the nouns in BROCANTO2, are used as playing pieces and can be further distinguished by their shape, round or square (BROCANTO2 adjectives). These pieces can be moved, swapped, captured, and released, actions corresponding to the BROCANTO2 verbs, either vertically or horizontally (BROCANTO2 adverbs). Before learning BROCANTO2, participants were shown possible moves in the computer board game without any linguistic input.

Participants learned BROCANTO2 either explicitly or implicitly, however for this study, no distinction was made between the two groups. The explicit group received metalinguistic lessons about the grammar of BROCANTO2 followed by exposure to meaningful aural examples of BROCANTO2 phrases and sentences which corresponded to game constellations and moves. The implicit group received only exposure to meaningful aural examples; however, time was controlled for such that both groups received approximately 13 minutes of training total. Following training, participants completed practice modules, which alternated between comprehension and production practice. During comprehension practice, participants heard a sentence in BROCANTO2, which corresponded to a game move, and were asked to make the corresponding move. During production practice participants viewed a game move and were asked to orally produce the corresponding sentence. In both cases correct responses earned the participants 10 points and incorrect responses cost them 10 points from a running total visible on the computer screen.

Procedure

Original testing consisted of four sessions and retention testing of one session. During the first testing session participants completed the background questionnaire and the four tests of cognitive ability. The second through fourth sessions were used to learn BROCANTO2. During the second session participants were introduced to the rules of the computer board game and received game token name training. Participants then received training on the language based on their group membership, explicit or implicit, which was randomly assigned, followed by practice with alternating blocks of comprehension and production. Once participants scored above chance, which was 45% correct, on two consecutive practice blocks their learning was assessed behaviorally and through ERP measures. During the ERP data collection, which was completed first, participants heard correct and incorrect sentences in BROCANTO2 and made grammaticality judgments. They were also given a non-ERP grammaticality judgment test (GJT) following ERP measurement. On the third day of testing participants again received training based on their group membership followed by alternating comprehension and production practice. On the fourth day of testing, participants completed practice followed by behavioral and ERP assessment, which was identical in method to the testing on day 2. Ideally participants reached 95% proficiency before assessment. Additionally, participants completed a speeded GJT, a written GJT, a free response task, and a debriefing questionnaire. When participants returned for retention testing 3 to 6 months later they were given eight block of practice, which alternated comprehension and production. They then completed all assessments that were given on the fourth day of testing. Additionally, they completed the testing phases of the four cognitive ability tests.

Scoring and Data Analysis

Scoring of the MLAT Paired Associates and CVMT was straightforward; the score was the number of correct responses, for the CVMT the visual discrimination test was not included. For the Weather Prediction task the score was the percentage of correct predictions, where a correct prediction was considered choosing the weather with a greater probability of occurring, note that cases with 50% probability were not considered. The CVLT-II yielded three scores used in our analyses; these were the responses to the delayed recalls and recognition. The free recall and cued recall scores were calculated by dividing the total number of items (16) by the number of items recalled. The yes/no recognition test made use of d-prime scoring, where the score is the proportion of correct responses minus the proportion of false alarms. These three CVLT-II scores were then combined using factor analysis which yielded one factor. Subsequent discussion of the CVLT-II scores will make use of only this combined factor.

In this study, participants were grouped based on their proceduralization of BROCANTO2. To operationalize the concept of proceduralization the ERP data were used. ERPs (Event-Related Potentials) measure changes in the electrophysiological activity in the brain related to a specific stimulus. ERP activity is characterized by four aspects: (a) the polarity of the electrical change, either positive (P) or negative (N), (b) the latency, referring to when the peak occurs, (c) the duration, how long the peak lasts, and (d) the scalp distribution, what areas of the scalp show the response (i.e. frontal, parietal, temporal, occipital, anterior, posterior, central) (Morgan-Short, 2007). Typical responses to stimuli based on these four characterizations are called ERP components. These ERP components can be used to interpret data given that they change

depending on the behavioral task requirements. In language, a few ERP components have been identified to indicate various types of linguistic processing (Friederici, 1995; Gleason & Ratner, 1998; Osterhout & Holcomb, 1995). Additionally, ERP components have been tied to different memory systems. Centro-parietal negativities (N400s), occurring between 250 and 600 ms, are thought to indicate declarative memory usage (Ullman, 2001a, 2001c). While left anterior negativities (LANs) occurring in the 150-500 ms time window have been posited to reflect use of the procedural memory system (Ullman 2001a, 2001c). Additionally, positive posterior components (P600s), occurring between 600 and 900 ms, may involve the basal ganglia (Friederici & Kotz, 2003; Friederici, Kotz, Werheid, Hein, & von Cramon, 2003; Friederici, von Cramon, & Kotz, 1999). Therefore ERP data can be used to determine which memory system, declarative or procedural, is used during processing.

A number of studies have investigated the brain responses to native language using ERPs. Violations of syntax have been shown to lead to LANs (Friederici, Gunter, Hahne, & Mauth, 2004; Friederici, Pfeifer, & Hahne, 1993; Neville, 1991; Rosler, Putz, Friederici, & Hahne, 1993) and P600s (Friederici, 1995; Friederici et al., 2004; Osterhout & Holcomb, 1992). This biphasic (LAN and P600) response is characteristic of native language processing of syntactic violations and has been thought to indicate procedural memory usage. Thus, for this study, proceduralization was originally operationalized as the presence of this biphasic response.

Typically, ERP studies average ERP data over items and subjects, however, in this study the ERP data were only averaged over items, not subjects, as individual ERPs were used to classify the participants. For each participant four sets ERPs were examined. These corresponded to two sets taken at the end of original testing,

or high proficiency, and two taken during the retention phase. At each time one set showed the response to violations of agreement in BROCANTO2 compared to matched control items and the other to violations of phrase structure, or word order, also compared to matched control items. This led to four ways of classifying the proceduralized and non-proceduralized groups, which was important given the highly innovative nature of the hypothesis and fluidity in proceduralization of an L2.

Upon examining the ERPs, no LANs were found in any individual; therefore the operationalization of proceduralization was limited to the presence of a P600, which is typically seen in response to syntactic violations in native language processing. A P600 was defined as significant positivities in the 600 to 1200 ms time window in multiple electrodes in the posterior area. This definition left the following number of participants in the proceduralized group: 4 of 17 for agreement at high proficiency, 6 of 17 for phrase structure at high proficiency, 5 of 17 for agreement at retention, and 11 of 17 for phrase structure at retention. At both high proficiency and retention testing 2 subjects, different for each time, were excluded due to large amounts of alpha waves in their ERPs, suggesting drowsiness and perhaps inattention to the stimulus. After determination of the groups was completed, the ERPs of individuals in the same group were averaged, to verify group cohesion. In all cases the proceduralized groups showed significant P600s, and in one case a significant LAN, and all non-proceduralized groups showed no evidence of a P600 or LAN.

To analyze the data, factor analysis was first used to determine any underlying factors among the cognitive tests. These factors were then used in ANCOVAs to test for a difference in cognitive abilities at retention between the proceduralized and non-proceduralized groups.

Results

Factor Analysis

Two factor analyses using principal components extraction with Varimax rotation were completed, one on the scores of cognitive ability before learning BROCANTO2 and the other on the scores at retention. Two factors emerged for both testing times with identical loadings. The scores from the MLAT Paired Associates and the CVLT-II formed one factor, named verbal memory, and the scores from the Weather Prediction task and CVMT another factor, named non-verbal memory.

Verbal and non-verbal memory scores were constructed for each participant. Scores on the individual tests were z-transformed then the z-scores on the MLAT Paired Associates and CVLT were combined to create a verbal memory score and the z-scores on the Weather Prediction task and CVMT were combined to create a non-verbal memory score. This was done for scores both before learning BROCANTO2 and from retention testing.

ANCOVAs

A total of eight ANCOVAs were run on the data where the dependent variable was either verbal memory or non-verbal memory at retention, the covariate was the corresponding memory measure at original testing, and the independent variable was group based on proceduralization. ANCOVAs, as opposed to ANOVAs, were used because the original scores influenced the scores at retention. The correlation between non-verbal memory at original testing and at retention was significant ($r = .694$, $p = .001$), while the correlation between verbal memory at original testing and at retention was large, but non-significant ($r = .194$, $p = .441$). However, given the small

sample size (N = 19), this large correlation warranted inclusion as a covariate.

As previously mentioned there were four ways of dividing the participants into proceduralized and non-proceduralized groups: agreement at high proficiency, phrase structure at high proficiency, agreement at retention, and phrase structure at retention. Results from all eight ANCOVAs are presented below (Table 1), however, only two yielded significant results. There was a significant difference in scores on the non-verbal measure when the groups were divided based on proceduralization of agreement structures at both high proficiency ($p = .033$) and retention ($p = .016$). The advantage in scores, however, was for participants in the non-proceduralized group (Table 2).

Table 1: ANCOVAs on Verbal and Non-verbal Measures

	Agreement at High Proficiency	Phrase Structure at High Proficiency	Agreement at Retention	Phrase Structure at Retention
Verbal	$F(1) = 0.034$ $p = .857$ $\eta^2 = .003$	$F(1) = 0.744$ $p = .404$ $\eta^2 = .054$	$F(1) = 2.119$ $p = .169$ $\eta^2 = .140$	$F(1) = 0.001$ $p = .979$ $\eta^2 = .000$
Non-verbal	$F(1) = 5.571$ $p = .033$ $\eta^2 = .285$	$F(1) = 1.087$ $p = .315$ $\eta^2 = .072$	$F(1) = 7.449$ $p = .016$ $\eta^2 = .347$	$F(1) = 0.699$ $p = .417$ $\eta^2 = .048$

Table 2: Adjusted Marginal Means for Significant Effects

	Agreement at High Proficiency	Agreement at Retention
Non-proceduralized	.574	.532
Proceduralized	-.897	-1.130

Given that the Weather Prediction task is a procedural memory task and the CVMT is a declarative memory task, it is possible that they behave differently from one another. To test if this was the case four additional ANCOVAs with the scores on the individual tests as the dependent variables and covariates and groups divided by proceduralization of agreement at high proficiency and at retention were run. These comparisons yielded no significant results (Table 3), suggesting that neither test is individually responsible for the significance.

Table 3: ANCOVAs on Weather Prediction and CVMT

	Agreement at High Proficiency	Agreement at Retention
Weather Prediction	$F(1) = 0.305$ $p = .590$ $\eta^2 = .021$	$F(1) = 2.049$ $p = .174$ $\eta^2 = .128$
CVMT	$F(1) = 0.700$ $p = .417$ $\eta^2 = .048$	$F(1) = 1.900$ $p = .190$ $\eta^2 = .119$

Discussion

Many of the results presented above were contrary to the hypotheses and therefore constituted surprising results. The underlying factors aligned according to verbal and non-verbal characteristics, rather than declarative and procedural memory characteristics. Significant results were found only when groups were divided based on proceduralization of agreement structures and only on the non-verbal measure. Additionally, the non-proceduralized group had the advantage, rather than the expected proceduralized group.

Verbal and Non-verbal Memory

The four tests of cognitive ability were hypothesized to primarily measure the declarative and procedural memory systems; however, they were actually more strongly related to verbal and non-verbal memory measures. It is thought that this division of tests occurred due to the more dissociable nature of the verbal and non-verbal memory systems than the declarative and procedural memory systems. Thus, an individual may have superior verbal memory and low non-verbal memory, but such a large difference does not occur with declarative and procedural memory. If this were the case the non-verbal measures would be a more cohesive group than the declarative memory measures and the separation seen would be expected.

Additionally, due to the nature of testing (i.e. no retraining at retention), all four tasks, including Weather Prediction, were likely to tap declarative memory more strongly. Thus, at retention at least, all four tasks might have been better classified as declarative tasks.

Proceduralization of Agreement

No hypotheses were made as to which of the four ways to classify participants as proceduralized would lead to significant results. Looking at the number of participants who proceduralized agreement structures though, it is apparent that proceduralization of agreement was more "difficult" than of phrase structure. This smaller number of participants able to proceduralize the agreement structures suggests that it is a more sensitive measure of proceduralization. Such a sensitive measure is desirable when trying to classify learners as having proceduralized a language because language consists of many components that need to be proceduralized and are done so at different times. The cutoff point for when a learner has proceduralized is difficult to determine, however, this data suggests that agreement

is superior to phrase structure in determining proceduralization of the language.

Advantage for Non-proceduralized Learners on Non-verbal Memory

The advantage for non-proceduralized learners was very surprising given the previous evidence of benefits for bilinguals. The results, though, may be explained based on the interplay of the declarative and procedural memory systems and the declarative nature of all tests at retention. A "seesaw" effect between the declarative and procedural memory systems has been noted (Poldrack & Packard, 2003; Ullman, 2004, 2005). This effect suggests that a highly functioning procedural system may suppress the declarative memory system and vice versa. Thus learners who achieved proceduralization of BROCANTO2 had a highly function procedural memory system which might have suppressed their declarative memory system. In this case these participants would be expected to perform worse on tests of declarative memory than those who did not proceduralize the language. Since all tests had a declarative nature at retention lower scores on all tests would be expected. This, however, was not the case; lower scores were only found on the non-verbal score. At this point it is not clear why verbal memory measures would not evidence the same effect.

Limitations

This study had a number of limitations which prevented reliable data from being collected and analyzed. First, BROCANTO2 is a small artificial language as opposed to a full natural language and thus the participants, even at high proficiency, could not be truly classified as bilinguals. If these participants are not bilinguals, then the benefits

associated with bilingualism would not be seen in these participants. Second, the measures used in this study did not correspond to those which evidenced benefits in previous studies. The extent of bilingual benefits is still being determined and the measures used in this study may not fall within the group of abilities that are enhanced in bilinguals. Additionally, it is unclear what the tests used were measuring at retention, as this is the first study that the author knows of which uses these four tests in a retention study in this manner (i.e. no retraining). A third limitation was the operationalization of proceduralization as the presence of a P600. This was done out of necessity, as the more tell tale LAN was not apparent in any ERP, however, it may be that no participant proceduralized BROCANTO2 to native-like processing levels. Fourth, the sample size used was too small to yield good results. Finally, the retention testing occurred 3 to 6 months following original training. This time span is large and it is possible that differing amounts of time to consolidate BROCANTO2 and the items from the cognitive abilities tests led to differences at retention testing.

Future Directions

Though this study presented a number of limitations in the experimental design, the question behind the study deserves further investigation. It is plausible that proceduralization relates to the threshold and allows cognitive benefits, however both proceduralization and cognitive benefits need to be better operationalized in future studies. These studies should focus on the measures in which benefits for bilinguals have already been attested, such as the Simon task or other measures of control processes. Additionally, the more rigorous biphasic response of LAN and P600 ERP components should be used, as well as other methods of identifying procedural memory

involvement, such as frequency effects. Finally, bilinguals should be measured in two natural languages rather than an artificial language.

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, N.J.: Erlbaum.

Anderson, J. R., Bothell, D., D., B. M., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036-1060.

Babcock, L., Stowe, J., Brovetto, C., & Ullman, M. T. (in prep) Frequency effects, age of arrival and amount of exposure in second language learning: A study of English past-tense production. Unpublished ms. Georgetown University, Washington DC

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology, 24,* 560-567.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York: Cambridge University Press.

Bialystok, E. (2005). Consequences of Bilingualism for Cognitive Development. Kroll, J. & De Groot, A.(eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press, 417-431.

Bransford, J. D. (Ed.). (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academic Press.

Cook, V. (1997) The consequences of bilingualism for cognitive processing. In de Groot & Kroll, *Tutorials in Bilingualism.* Mahwah, NJ: Lawrence Erlbaum.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism, 9,* 1-43.

Cummins, J. (1978). Bilingualism and the development of metalinguistic awareness. *Journal of Cross-Cultural Psychology, 9,* 131-149.

de Groot, A. D. (1965). *Thought and choice in chess* (1st ed.). The Hague: Mouton Publishers.

de Groot, A. & Gobet, F. (1996). *Perception and memory in chess. Heuristics of the professional eye*. Assen: Van Gorcum.

DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125-151). New York, NY: Cambridge University Press.

Diaz, R. M. (1985). The intellectual power of bilingualism. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 7,* 16-22.

Edwards, D., & Christopherson, H. (1988). Bilingualism, literacy, and metalinguistic awareness in preschool children. *British Journal of Development Psychology, 6,* 235-244.

Eviatar, Z., & Ibrahim, R. (2000). Bilingualism is as bilingual does: Metalinguistic abilities of Arabic-speaking children. *Applied Psycholinguistics, 21,* 451-471.

Feldman, C., & Shen, M. (1971). Some language-related cognitive advantages of bilingual 5-year-olds. *Journal of Genetic Psychology, 118,* 235-244.

Foerde, K., R. A. Poldrack, et al. (in press). "Secondary task effects on classification learning." *Memory and Cognition.*

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language, 50*, 259-281.

Friederici, A. D., Gunter, T., Hahne, A., & Mauth, K. (2004). The relative timing of syntactic and semantic processes in sentence comprehension. *NeuroReport, 15*(1), 165-169.

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *NeuroImage, 20 Suppl 1*, S8-S17.

Friederici, A. D., Kotz, S. A., Werheid, K., Hein, G., & von Cramon, D. Y. (2003). Syntactic comprehension in Parkinson's disease: investigating early automatic and late integrational processes using event-related brain potentials. *Neuropsychology, 17*(1), 133-142.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1*(3), 183-192.

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*(1), 529-534.

Friederici, A. D., von Cramon, D. Y., & Kotz, S. A. (1999). Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. *Brain, 122*(Pt 6), 1033-1047.

Gathercole, V. C. M (1997). The linguistic mass/count distinction as an indicator of referent categorization in monolingual and bilingual children. *Child Development*, 68, 832-842.

Gleason, J. B., & Ratner, N. B. (Eds.). (1998). *Psycholinguistics* (2 ed.). Fort Worth: Harcourt Brace College Publishers.

Green, G.A., & Gilhooly, J.K., 1992. Empirical advances in expertise research. In: Keane, M.T.,

Heller, J. I. & Reif, F. (1984). Prescribing effective human problem solving processes: Problem descriptions in physics. *Cognition and Instruction, 1* (2), 177-216.

Hualde, J. I., & Ortiz de Urbina, J. (Eds.). (2003). *A grammar of Basque*. Berlin: Mouton de Gruyter

Lado, B. (2006) The Role of Bilingualism, Type of Feedback and Cognitive Capacity in the Acquisition of Non-primary Languages: A Computer-generated Study. PhD dissertation, Georgetown University.

Lambert, W. E., Tucker, G. R., & d'Anglejan, A. (1973). Cognitive and attitudinal consequences of bilingual schooling. *Journal of Educational Psychology, 85,* 141-159.

Landry, R. G. (1974) A comparison of second language learners and monolinguals on divergent thinking tasks at the elementary school level. Modern Language Journal, 58, 10-15.

Mallison, G. (1986). *Rumanian*. Dover, New Hampshire: Croon Helm.

Morgan-Short, K. (2007) A Neurolinguistic Investigation of Late-learned Second Language Knowledge: The Effects of Explicit and Implicit Conditions. PhD dissertation, Georgetown University.

Nayak N., Hansen, N., Krueger, N., & McLaughlin B. (1990). Language-learning strategies in monolingual and multilingual adults. *Language Learning, 40*(2), 221-244.

Neville, H. J. (1991). Neurobiology of cognitive and language processing: Effects of early experience. In K. R. Gibson & A. C. Peterson (Eds.), *Brain maturation and cognitive development: Comparative and crosscultural perspectives* (pp. 355-380): de Gruyter.

Oldfield, R. C. (1971). "The assessment and analysis of handedness: The Edinburgh inventory." *Neuropsychologia* **9**(1): 97-113.

Opitz, B., & Friederici, A. (2002). Artificial language acquisition: Changes in brain activity during the course of learning. *Journal of Cognitive Neuroscience, Supplement*, 35.

Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language–like rules. *NeuroImage, 19*(4), 1730-1737.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785-806.

Osterhout, L., & Holcomb, P. J. (1995). Event-related potentials and language comprehension. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of Mind: Event-related brain potentials and cognition* (Vol. 25, pp. 171-215). Oxford: Oxford University Press.

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. Psychological Monographs: General and Applied, 76, 27, 1-23.

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia, 41*(3), 245-251.

Ramsay, R. M. G. (1980). Language-learning approach styles of adult multilinguals and successful language learners. In V. Teller & S. J. White (Eds.), *Studies in child language and multilingualism*, (pp. 73-96). New York: New York Academy of Sciences.

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode. Journal of Memory and Language, 33, 390-405.

Ricciardelli, L. A. (1992). Bilingualism and cognitive development in relation to threshold theory. Journal of Psycholinguistic Research, 21, 301-316.

Rosler, F., Putz, P., Friederici, A., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience, 5*(3), 345-362.

Sanz, C. (2000). Bilingual education enhances third language acquisition: Evidence from Catalonia. *Applied Psycholinguistics, 21,* pp.23-44.

Sanz, C. (2007) The role of bilingual literacy in the acquisition of a third language. In Perez-Vidal, Bel, and Garau (Eds.) A Portrait of the Young in the New Multilingual Spain, Clevedon (UK): Multilingual Matters.

Segalowitz, N., & Frenkiel-Fishman, S. (2005). Attention control and ability level in a complex cognitive skill: Attention shifting and second-language proficiency. *Memory & Cognition, 2005, 33, 4*, 644-653.

Ullman, M. T. (2001a). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition, 4*(1), 105-122.

Ullman, M. T. (2001b). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30*(1), 37-69.

Ullman, M. T. (2001c). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience, 2*, 717-726.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*(1-2), 231-270.

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., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., et al. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience, 9*(2), 266-276.

Appendix I

Table III-3. *The BROCANTO2 Lexicon*

Lexical category	BROCANTO2 word	Corresponding symbols and meanings
Nouns		
	pleck$_m$	
	neep$_m$	
	blom$_f$	
	vode$_f$	
Adjectives	trois(e$_m$/o$_f$)	round
	neim(e$_m$/o$_f$)	square
Determiners	li$_m$/lu$_f$	the
Verbs	klin$_{intran}$	move
	nim$_{tran/intran}$	capture
	yab$_{tran/intran}$	release
	praz$_{tran}$	swap
Adverbs	noyka	vertically
	zayma	horizontally

Note. Subscript m = masculine; subscript f = feminine; tran = transitive; intran = intransitive.

From Morgan-Short (2007)