2017

Late L2 Acquisition and Phonological Memory Development

Madeline Westerhoff
Butler University

BUTLER UNIVERSITY HONORS PROGRAM

Honors Thesis Certification

Please type all information in this section:

Applicant: Madeline Rose Westerhoff

Thesis title: Late L2 Acquisition and Phonological Memory Development

Intended date of commencement: Saturday, May 6, 2017

Read, approved, and signed by:

Thesis adviser(s) _____ Date: 4-20-17

Reader(s) _____ Date: 4/20/17

Certified by _____(Director, Honors Program) Date:_____

For Honors Program use:

Level of Honors conferred: University _____

Departmental _____

_____

_____

_____

Late L2 Acquisition and Phonological Memory Development

A Thesis

Presented to the Department of Communication Sciences & Disorders

College of Communication

and

The Honors Program

of

Butler University

In Partial Fulfillment

of the Requirements for Graduation Honors

Madeline Rose Westerhoff

April 26th, 2017

Abstract

The development of phonological memory, a subdivision of short term memory, is crucial to learning and analyzing sequences of sounds to form words and phrases. This process utilizes short term representation and rehearsal to allow for the eventual long term representation of language. Individuals learning more than one language must acquire even more of these language sequences than monolinguals. Previous research agrees that early Spanish-English bilinguals have superior phonological memory to monolinguals, however the impact of second language acquisition on phonological memory remains unknown. This study examined three groups of undergraduate participants studying Spanish as a second language including study abroad students (SA), students currently enrolled in a Spanish course at or above the 300-level (CE), and students not currently enrolled in a Spanish course (NE). Participants completed tests of phonological memory including digit span and sentence repetition at the beginning and culmination of an academic semester. Participants also provided demographic data, L2 acquisition information, and their frequency of language use so that changes could be calculated. Results revealed that CE students were superior during pretest, but during posttest SA and CE demonstrated comparable results. The NE students demonstrated consistently lower scores. These findings suggest that foreign language instruction may benefit phonological memory development, even in late acquisition of the second language. The possible role of years of formal language instruction will be discussed.

1. Introduction

At any moment throughout your day, you may encounter the need to remember a series of digits for a phone number. Or, you may need to recall a list of items needed at the grocery store. These tasks, among others, require use of working memory. Working memory can be thought of as a "mental sticky pad" where we place information in our brains until it needs to be used, shortly after placement. According to the Baddeley model of working memory, the limited amounts of verbal information that can be stored are dependent upon a system called the phonological loop (Baddeley, 1986). As vital

information is heard, the sequence enters this phonological loop of working memory where it is held until it needs to be recalled. According to Gathercole (1999), holding information in this way is effortful, attention-demanding, and highly prone to failure. Additionally, one's phonological memory capacity varies with age and experiences.

A widely used mechanism to measure individuals' phonological working memory capacity is the digit span test. During this test, a researcher reads a list of digits that the participant repeats back to the researcher, beginning with two digits and gradually increasing the number of digits within a set with each successful repetition. The expected digit span differs across languages as some words for numbers are phonotactically more complex than others, as is the case for Spanish (i.e. average digit span of 5.8) (Ardila et al., 2000) versus English (i.e. average digit span of 7.0) (Wechsler, 1944). As a result, capacity for digit recall and repetition changes as a function of language. Similarly, bilingualism has been shown to impact digit span (Ardila et al., 2000).

As Ellis (1996) explains, the development of phonological memory is crucial to learning and analyzing sequences of sounds to form words, sequences of words to form phrases, etc. This process utilizes short term representation and rehearsal to allow for the eventual long term representation of the second language. Crucial in this process is the individual's short-term memory capacity, which impacts their ability to recall and repeat phonological sequences. Individuals who are tasked with learning more than one language must acquire even more of these sequences than monolinguals. Therefore, researchers were interested in quantitatively determining differences in phonological memory and subsequent phonological skills, such as digit span, in bilinguals.

Ardila et al. (2000) analyzed digit span in 69 Spanish-English bilinguals with differing ages of acquisition of their second language. They found that their performance in Spanish digit span (6.2) was higher than normal monolingual performance. When these Spanish-English bilinguals were divided into groups according to L2 age of acquisition (AoA), those with an earlier AoA (<12 years) had a higher Spanish performance on digit span than their late acquisition counterparts whose performance corresponded to the average monolingual digit span. Ardila et al. (2000) concluded that bilinguals with earlier

acquisition of a second language (L2) obtained higher digit span scores in their first language (L1). The authors proposed that better working memory was a result of participants' bilingualism.

Another mechanism used to test phonological working memory, although not as widely used as digit span, is word span. Average word span differs greatly and depends on whether words are semantically related, and if they are included in a meaningful sentence (Ardila, 2003). Testing participants' digit and word span allows for researchers to gain a clearer picture of working memory capacities.

Sentence repetition tasks (SRT) have also been previously administered as a measure of phonological working memory (Alloway et al., 2003). In a research study of 663 children which investigated the functional organization of working memory and related cognitive abilities, participants were administered several tests including two SRTs in order to measure working memory. The tasks each consisted of 10 sentences which ranged between 6 and 9 words in length. The experimenter read each sentence aloud, and then ask the participant to repeat the sentence immediately. If a participant repeated the entire sentence correctly, they received a score of 1. If any errors were made, they received a score of 0. Therefore, there was a maximum score of 10 for each set.

Proficiency of language usage has been shown to impact cognition in bilinguals (Perani et al., 1998). For example, the performance of a bilingual who uses both languages daily will pattern differently from a university student whose only L2 exposure is limited to the classroom. In a study of Italian-English bilinguals with differing proficiencies of their L2 and varying AoA, researchers used PET scans of the brain to observe changes in brain activity while participants listened to stories in their L1 and L2. Those with late acquisition of their L2 had acquired English at school after the age of 10. High-proficiency participants had spent one or more years in an English-speaking country (range 1-6 years) and spoke and/or read both languages in daily life. Low-proficiency participants had never spent more than a month in an English-speaking environment. A word-translation task revealed significant group differences, further demonstrating their variation in proficiency. The researchers observed changes in

activation of the temporal lobes and temporoparietal cortex which indicate an individual's active participation in a listening, reading, or continuous speech task. The scans revealed that the participants with high proficiency of their L2 had increased activation of the aforementioned brains areas while listening to stories read aloud in their L2, regardless of their AoA (Perani et al., 1998). Those with low proficiency and late acquisition of their L2 (English) demonstrated less activation in the PET scans when listening to the stories in English. Taken together, L2 proficiency, which is closely related to frequency of exposure and use, plays a more central role in increased cognition in bilinguals than AoA.

Lord (2006) examined phonological memory abilities of students participating in an intensive L2 experience (Lord, 2006). Spanish-speaking university students were tested, before and after their 6-week study abroad experience. In one-on-one interviews, the participants heard 10 sentences read aloud by the researcher. These sentences were between 19 and 26 syllables, with an average of 22.2 syllables. This length was chosen because sentences of that duration are just beyond what the phonological loop of working memory can hold (Lord, 2006). Each sentence also contained one invented, but phonotactically possible nonce word. Data showed no increased ability to repeat nonce words, but an increased ability to repeat longer strings of syllables. Their improved performance on sentence repetition demonstrated gains in one measure of phonological working memory. The participants in Lord's (2006) study served as their own comparison group, pre and post study abroad, which leaves open the question as to whether students in a traditional high level Spanish language course experiences similar gains.

The current study of phonological memory, as measured by a digit span test and sentence repetition, aimed to answer the following research question: What effect does immersion have on the phonological working memory abilities of late L2 learners as compared to learners continuing in L2 learning in a traditional university classroom? This study sought to investigate the role of the learning environment on cognitive gains in late L2 learners. It was hypothesized that increased exposure and use would result in superior

phonological memory. Specifically, it was expected that longer intense exposure (i.e. a full semester abroad), would result in improved production of nonce words, which previously saw no difference. It was also expected that those with low levels of exposure to L2 would demonstrate no changes in phonological memory from pretest to posttest, and lower performance compared to groups exposed to the L2.

Although bilingualism and AoA are known to impact phonological working memory, there is limited data on the influence of an immersion context for late L2 learners. The proposed investigation aims to examine how working memory develops in this understudied group of bilinguals.

2. Methods

2.1 Participants

Participants included three groups of undergraduate students from Butler University with varying degrees of Spanish language experience. All participants were native English speakers, and began Spanish second language acquisition via formal schooling as adolescents. The first group was composed of students that participated in a semester-long study abroad experience in Spain, a second group composed of students enrolled in a 300-level Spanish course or greater, and a third group of students who had discontinued their Spanish studies.

The study abroad (SA; n=4) students were recruited from a set of students who attended a semester-long study abroad trip to Alcalá, a city in Spain. SA participants had an average of 5.5 years of formal Spanish instruction and an average of 2.5 semesters of formal Spanish instruction at the 300-level or above. All SA participants were female.

The second group consisted of students currently enrolled in a Spanish course at the 300-level or above (CE; n=6). CE participants had an average of 7.5 years of formal Spanish instruction and an average of 2.83 semesters of formal Spanish instruction at the 300-level or above. All CE participants were female.

The third group consisted of students who previously took 6 hours of 200-level Spanish courses, but were not currently enrolled in a Spanish course (NE; n=5). NE

participants had an average of 4.4 years of formal Spanish instruction and an average of 0.8 semesters at the 300-level or above. All NE participants were female. See Figure 1 for across group comparison of formal Spanish instruction experience.

Participants from the CE and NE groups were recruited via word of mouth announcements and emails to students enrolled in a 300-level Spanish course. Participants completed all tasks at the start of the semester, and a second time before the semester ended to examine changes in phonological memory. Participants were compensated for their participation in the form of a $10 gift card.

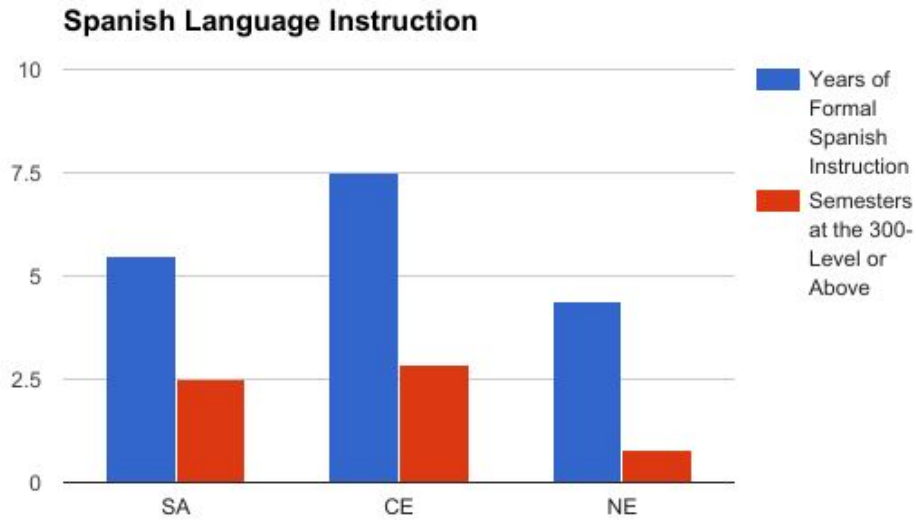

Figure 1: Spanish Language Instruction. Years of Spanish language instruction and number of semesters of Spanish courses at or above the 300-level.

2.2 Phonological Memory Tasks and Analyses

All participants completed two tasks: sentence repetition and digit span. Lord's (2006) twenty sentences (Appendix A) were used for the SRT--10 for the pretest and the remaining 10 sentences for the posttest. The sentences ranged from 19 to 26 syllables, with an average of 22.2 syllables, and contained Spanish real words as well as one meaningless but phonotactically possible nonce word. Sentence repetition analyses

included several measures including accuracy of production of syllables and nonce words.

The first measure for analysis in the SRT was labeled as syllables, and the score was presented as a percentage of the total sentence repetition syllables produced correctly. For example, if the entire pretest SRT consisted of 200 syllables, and the participant accurately produced 150 of those syllables, then they would receive a score of 75%.

The second measure used during analysis of the SRT was accurate production of vocoid segments. Vocoids are the phonetic correspondence to vowels which represent the phonological aspect of language. More specifically, vocoids are categorized as *central oral resonants* because their production causes the passage of air flow through the central oral cavity and does not involve any constriction. Examples of vocoids include [j], [i], and [a]. The score of vocoid segments was presented as a percentage of the total SRT vocoid segments produced correctly. For example, given a total of 80 vocoid segments in the pretest SRT, accurate production of 60 vocoid segments, results in a score of 75%.

The third measure used during analysis of the SRT was accurate production of nonce words. Nonce word production was presented as a raw score, representing the total number of accurately produced whole nonce words. For example, if a participant correctly produced 6 out of the 10 nonce words during the SRT, then they would receive a score of 6.

The fourth measure used during analysis of the SRT was nonce syllables. The score was a percentage of the total syllables within nonce words produced correctly. For example, given 50 total syllables of all the nonce words in the SRT pretest, and correct production of 25 of these syllables, a participant would receive a score of 50%.

The fifth and final measure of SRT was nonce vocoid segments. The was a percentage of the total vocoid segments within nonce words produced correctly. For example, if the total number of vocoid segments for all nonce words was 40, and 20 syllables were produced correctly, the nonce vocoid segment score would be 50%.

In addition to the SRT, participants completed a digit span task in English. In line with Ardila et al. (2000), the digit span task was administered in participants' L1 in order to examine the effects of L2 experience/exposure on L1 digit span. The digit span test was created by a random number generator. For this task, the researcher read aloud sets of series of digits, one set at a time, beginning with the smallest set. The test began with two sets of numbers consisting of a series of two digits each, with subsequent sets increasing by one digit. Participants repeated the series after the researcher read them aloud at a pace of one second per digit. After two failed attempts within a set, the participant failed the task. The digit span score was determined by the number of digits contained in the last series that was repeated correctly. For example, a participant who correctly repeated a series of 7 digits, but incorrectly repeated both series of 8 digits, received a digit span score of 7.

All participants signed an informed consent form for testing and audio-recording of pre- and posttest responses. Participants filled out a questionnaire online at both testing points in which they provided demographic data, information regarding their education setting, language history, and amount of time per week spent using the second language. Responses to the initial questionnaire (Appendix B) also served to confirm group assignment for data analysis. The questionnaire for the end of the semester (Appendix C) served to detect changes in second language use/exposure.

All participants were tested via Skype or FaceTime and were audio recorded for offline scoring and analysis. SA students completed the pretest before leaving the country for testing prior to L2 immersion, and the remaining participants were tested within the first two weeks of class. Both groups completed posttests before the end of the semester.

3. Results

Task analyses included across group comparisons to determine the relationship between varying degrees of L2 exposure and phonological memory. T-tests, with alpha level set at $\alpha=.05$, were performed to evaluate intergroup and intragroup differences at each test

point and across test points. Results for each task measure are presented in the following order: intergroup pretest, intergroup posttest, intra-group pretest to posttest.

3.1 Syllables

3.1.1 Intergroup Pretest

Pretest data revealed that CE scored highest (77%) in average accurate production of overall syllables in the sentence-recall task (Fig. 2). SA scored second highest (61.74%) and NE scored the lowest (54.24%). There was a significant difference between between CE and NE in pretest production of syllables ($p<0.05$). However, there was not a significant difference between SA and CE ($p=0.09$) nor SA and NE ($p=0.40$) in the pretest.

3.1.2 Intergroup Posttest

Posttest data showed that CE scored the highest (69%) (Fig. 2) and SA scored the second highest (68%), but this difference was not significant ($p=0.46$). NE scored the lowest in posttest production of syllables (49%), and this was found to be significantly different than CE posttest production of syllables ($p<0.05$). There was a significant difference between SA and NE in the posttest ($p<0.05$).

3.1.3 Intra group Pretest/Posttest

Although CE scored the highest, SA demonstrated the greatest, and only, improvement from pretest (61.74%) to posttest (68%). However the t-test did not reveal that the increase in SA syllable production from pretest to posttest was significant ($p=0.16$). CE decreased in average total syllables produced correctly from pretest (77%) to posttest (69%), and this decrease was significant ($p<0.05$). NE also decreased in average total syllables produced correctly from pretest (54.25%) to posttest (49%) but this decrease was not significant ($p=0.14$). See Figure 2 for inter and intragroup comparisons of percentage of syllable production for the pretest and posttest SRT.

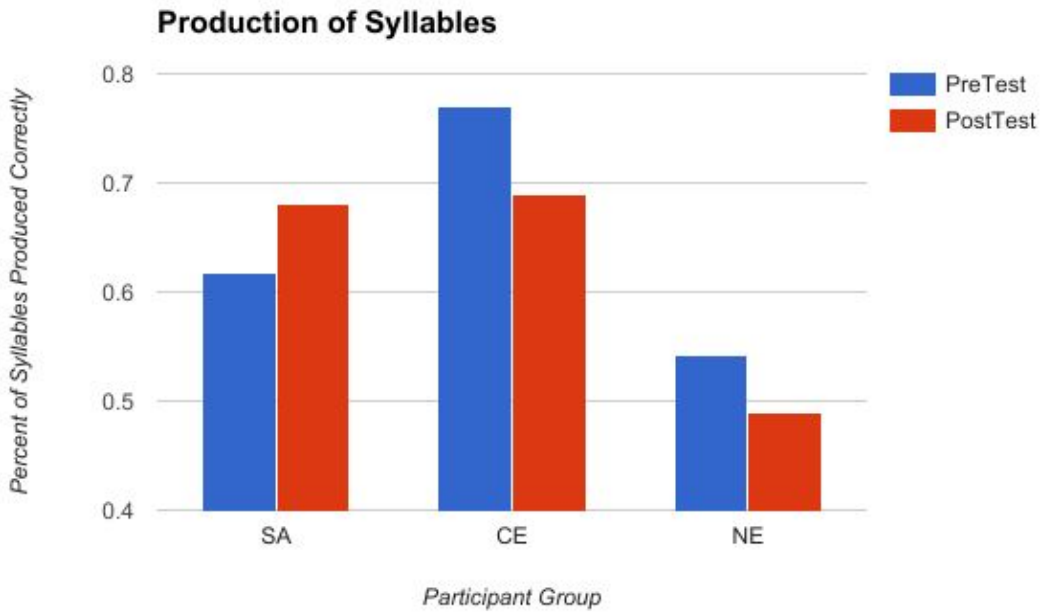

Figure 2: Production of syllables. The average proportion of syllables produced correctly in the sentence repetition task per group at pre- and posttest.

3.2 Vocoid Segments

3.2.1 Intergroup Pretest

Pretest data demonstrated that CE scored highest (81.60%) in average accurate production of overall vocoid segments in the sentence-recall task. NE scored second highest (61.04%), and SA scored the lowest in the pretest (57.07%). There was not a significant difference between SA and CE in pretest production of vocoid segments (p=0.11), nor between SA and NE (p=0.55). However, there was a significant difference between CE and NE (p<0.05).

3.2.2 Intergroup Posttest

Posttest data showed that SA scored the highest (74%), followed by CE (73%), and NE scored the lowest in posttest production of vocoid segments (54%). The difference between SA and CE in posttest production of vocoid segments was not significant (p=0.46), but SA performance was significantly higher than NE in posttest

production of vocoid segments ($p<0.05$). A t-test also revealed that the difference between CE and NE in posttest production of vocoid segments was significant ($p<0.05$).

3.2.3 Intra group Pretest/Posttest

SA saw the greatest, and only, improvement in vocoid production from pretest (57.07%) to posttest (74%), however, the difference was not significant ($p=0.20$). Similarly, NE decreased from pretest (61.04%) to posttest (54%) and this was not significant ($p=0.16$). CE saw a significant decrease from pretest (81.60%) to posttest (73%; $p<0.05$).

3.3 Nonce Words

3.3.1 Intergroup Pretest

Pretest data revealed that CE scored highest (4.67) in accurate production of nonce words in the sentence-repetition task (Fig. 3). NE scored second highest (2) and SA scored the lowest (1.75) in the pretest. There was a significant difference between SA and CE in pretest production of nonce words ($p<0.05$). However, the difference between SA and NE pretest production of nonce words was not significant ($p=0.82$). The difference between CE and NE in the pretest production of nonce words was significant ($p<0.05$).

3.3.2 Intergroup Posttest

Posttest data revealed that CE scored the highest (3.83), however SA scored the second highest (3.25) (Fig. 3). This difference between SA and CE posttest production of nonce words was not significant ($p=0.36$). NE scored the lowest (1.6) in the posttest. There was not a significant difference between NE and SA posttest production of nonce words ($p=0.12$) nor between NE and CE ($p=0.08$).

3.3.3 Intra group Pretest/Posttest

Although CE scored the highest at both points, SA saw the greatest, and only improvement in nonce word production from pretest (1.75) to posttest (3.25). A t-test revealed that the relationship between SA pretest and posttest production of nonce words was significant ($p<0.05$). CE and NE both decreased in their production of nonce words

from pretest to posttest. The difference between CE pretest and posttest production of nonce words was not significant (p=0.13) as well as NE (p=0.48). See Figure 3 for intergroup and intragroup data regarding the production of nonce words.

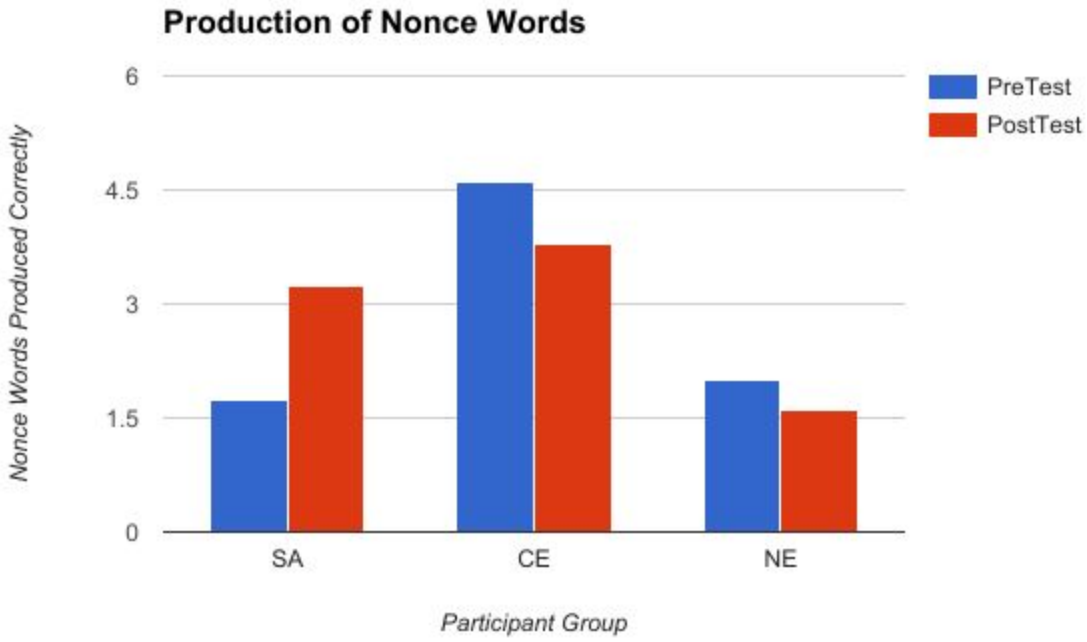

Figure 3: Production of nonce words. The raw number of nonce words produced correctly by group at pre- and posttest.

3.4 Nonce Syllables

3.4.1 Intergroup Pretest

Pretest data revealed that CE scored highest (61%) in accurate production of nonce word syllables in the sentence-recall task (Fig. 4). NE scored second highest (38%) while SA scored the lowest (32%) in the pretest. The t-test revealed that the difference between SA and CE in pretest production of nonce syllables was significant (p<0.05). However, the difference between SA and NE in pretest was not significant (p=0.06). The difference between CE and NE in the pretest was significant (p<0.05).

3.4.2 Intergroup Posttest

Posttest data showed that SA scored the highest (53%) while CE scored the second highest (50%) (Fig. 4). NE scored the lowest (32%) in the posttest. The t-test revealed that no differences concerning posttest production among groups were significant for nonce syllable production.

3.4.3 Intra group Pretest/Posttest

SA saw the greatest, and only, improvement in nonce syllables from pretest (32%) to posttest (53%). This difference was significant ($p<0.05$). CE and NE both saw decreases in production of nonce syllables from pretest to posttest, 61% to 50%, 38% to 32%, respectively. The difference between CE pretest and posttest production of nonce syllables was not significant ($p=0.09$) as well as NE ($p=0.51$). See Figure 4 for intergroup and intragroup data concerning the production of nonce syllables.

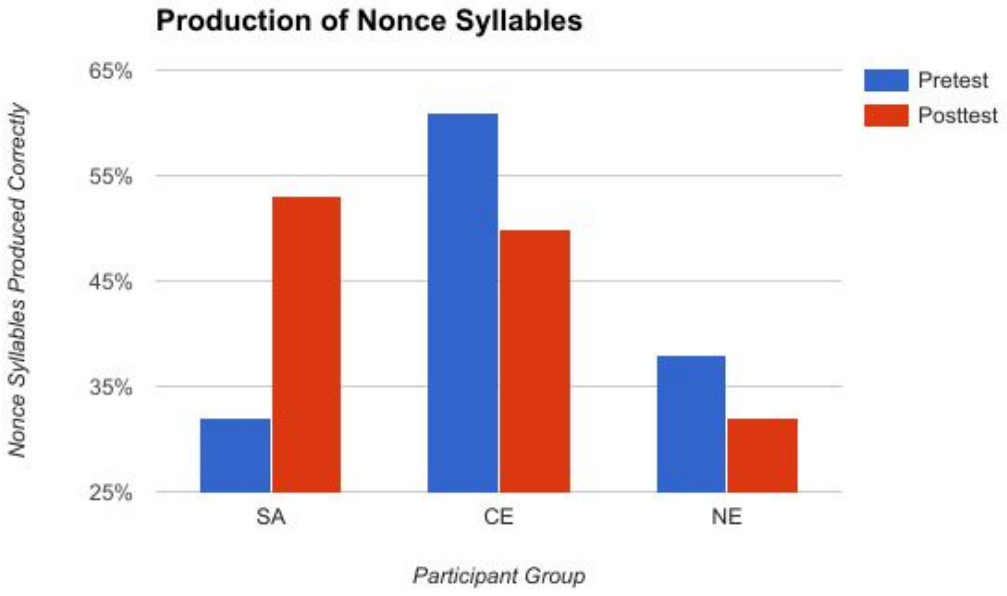

Figure 4: Production of Nonce Syllables. The proportion of nonce syllables produced correctly in the sentence repetition task by group at pre- and posttest.

3.5 Nonce Vocoid Segments

3.5.1 Intergroup Pretest

Pretest data revealed that CE scored highest (73%) in accurate production of nonce vocoid segments in the sentence-recall task. NE scored second highest (49%), and SA scored the lowest (25%). There were no significant differences in pretest production of nonce vocoid segments across groups.

3.5.2 Intergroup Posttest

Posttest data showed that SA scored the highest (63%), CE scored the second highest (58%), and NE scored the lowest (43%). These differences were not significant.

3.5.3 Intra group Pretest/Posttest

SA saw the greatest and only improvement in accurate production of nonce vocoid segment from pretest (25%) to posttest (63%), and CE and NE decreased in production of nonce vocoid segments from pretest to posttest. However, these within group changes were not significant.

3.6 Digit Span

3.6.1 Intergroup Pretest

Pretest data revealed that NE scored highest (7.8) in the digit span task. CE scored second highest (7.17), and SA scored the lowest (6.25) in the pretest (Fig. 5). There were no significant differences concerning pretest digit span across the groups.

3.6.2 Intergroup Posttest

Posttest data showed that NE once again scored highest (7.4) in the digit span task. CE and SA had the same score (7) for the posttest digit span task (Fig. 5). No significant differences were found among groups for posttest digit span scores.

3.6.3 Intra group Pretest/Posttest

Although NE exhibited the highest digit span score at both data points, SA saw the greatest, and only improvement in digit span score from pretest (6.25) to posttest (7). However, this difference was not significant (p=0.11). CE and NE both saw decreases in

digit span from pretest to posttest, and the t-test revealed that these differences were not significant for CE (p=0.40) nor NE (p=0.59). See Figure 5 for across group comparisons of pretest and posttest digit span scores.

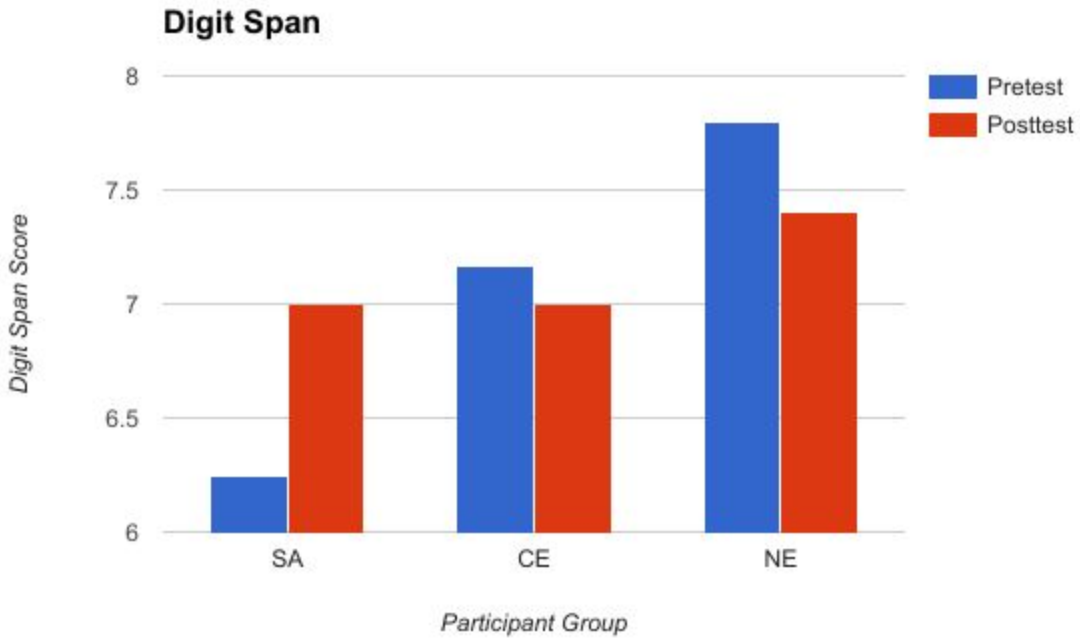

Figure 5: Digit Span. Digit span scores per group at pre- and posttest.

4. Discussion

The overall trends in performance revealed greater phonological memory for CE during the pretest, and comparable phonological memory for SA and CE in the posttest. SA exhibited some statistically significant gains in phonological memory measures from pretest to posttest while CE demonstrated some statistically significant decreases in phonological memory task performance. As predicted, NE demonstrated consistently lower scores.

4.1 Syllables and Vocoid Segments

In consideration of the first two SRT analyses, syllables produced correctly and vocoid segments produced correctly, CE decreased from pretest to posttest, and this difference was statistically significant. However, despite this significant decrease, SA and

CE were comparable in their performance of syllable and vocoid segment production at both testing points. These results differ from previous research (Lord, 2006) in which participants that were studying abroad demonstrated an increased ability to repeat longer strings of syllables at the culmination of their L2 immersion experience.

4.2 Nonce Words

Now we will consider the final three analyses of the SRT which all examined the production accuracy of nonce words, including number of nonce words produced correctly, percentage of nonce syllables produced correctly, and percentage of nonce vocoid segments produced correctly. In the pretest, CE was significantly higher than SA and NE in production of nonce words and nonce syllables. SA exhibited significant increases in production accuracy across the majority of nonce word measures. The posttest production of nonce words, nonce syllables, and nonce vocoid segments was comparable between SA, CE and NE. These results once again differ from previous research (Lord, 2006) in which participants demonstrated no changes in their ability to repeat nonce words after their study abroad experience. The data from this research study suggest that the longer L2 immersion experience increases an individual's ability to repeat nonce words in an SRT, produced in their late L2 of Spanish.

It was predicted that pretest performance for SA and CE would be comparable across all measures of phonological memory since both groups were in continuation of their L2 studies at the university level. However, CE demonstrated superior pretest scores as compared to SA. The superior scores of CE in the pretest was possibly a result of more years of formal instruction, or more years of exposure and overall experience with their L2, Spanish. On average, CE participants had two more years of formal Spanish instruction than those in SA, but only 0.3 more semesters of Spanish courses at the 300-level or above. The difference between SA and CE in terms of years of formal Spanish instruction was significant, but the number of semesters at the 300-level or above was not (p=0.77). It is unclear whether significantly lower performance for NE in the pretest was a result of absence of Spanish instruction/exposure, fewer years of formal Spanish instruction (4.4), and/or fewer semesters of formal Spanish instruction at the

300-level or above (0.8) as compared to SA and CE. Another possibility is a self-selection bias for the CE group. Individuals were informed during recruitment that testing would involve tasks requiring Spanish-speaking, so those who were less confident in their speaking skills may have chosen to abstain. Therefore, the CE group may not be an accurate representation of the average student in Spanish classes at the 300-level or above. Additionally, one might speculate that SA may have performed significantly lower than CE in the pretest because those who chose to study abroad sought an opportunity to improve their Spanish. Overall, these findings suggest that an immersion experience, such as that provided by a study abroad program, may result in improved phonological memory as measured by production of nonce words, nonce syllables, and nonce vocoid segments in an SRT.

4.3 Digit Span

Results of the digit span task revealed no significant differences. An explanation for the lack significant differences between groups in the digit span task includes the possibility that this digit span test was not sensitive enough to capture differences between groups, or that the a larger N is required to observe such anticipated group trends.

5. Conclusion

The goal of this research was to fill the gap in the current literature concerning phonological memory development for bilingual individuals, specifically how phonological memory is impacted in those with late acquisition of their second language. Through the examination of phonological memory development among three different groups of late-acquisition Spanish bilinguals over the course of a semester, several conclusions about the impact of varying forms and degrees of Spanish instruction and exposure can be made. The data suggest that participation in a semester-long study abroad program in a Spanish-speaking country provides the best opportunity to improve phonological memory as compared to other forms and degrees of Spanish language exposure, or a lack of foreign language exposure. The results here also suggest that the

positive impact of a number of years of formal Spanish instruction on phonological memory may be achieved through participation in a semester-long study abroad experience.

Although this research demonstrated several statistically significant increases in phonological memory for SA participants, future research should consider modifications to independent variables in the form of matching in consideration of L2 AoA, years of formal Spanish instruction, and/or amount of daily L2 exposure/use. Future studies may include a self-rating of Spanish skills for all participants during pretest in order to determine whether performance is related to self-perceived skills. Finally, as usual, a larger number of participants to increase statistical power. Even though several measures demonstrated increases in phonological memory for SA, these increases were not significant. A larger *n* may reveal statistically significant differences in future research.

6. Acknowledgements

This project was supported by an Undergraduate Research Thesis Grant awarded by the Center for High Achievement and Scholarly Engagement at Butler University.

References

Allen, H. W., & Herron, C. (2003). A mixed‐methodology investigation of the linguistic and affective outcomes of summer study abroad. *Foreign Language Annals, 36(3)*, 370-385.

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of experimental child psychology*, *87*(2), 85-106.

Ardila, A., Rosselli, M., Ostrosky-Solis, F., Marcos, J., Granda, G., & Soto, M. (2000). Syntactic comprehension verbal memory and calculation abilities in Spanish-English bilinguals. *Applied Neuropsychology, 7*, 3–16.

Ardila, A. (2003). Language representation and working memory with bilinguals. *Journal of Communication Disorders, 36(3)*, 233-240.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review, 105(1)*, 158.

Ellis, N. C. (1996). Sequencing in SLA. *Studies in second language acquisition, 18(01)*, 91-126.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in cognitive sciences, 3(11)*, 410-419.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105(19)*, 6829-6833.

Li, P., Sepanski, S., & ZhAo, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior research methods*, 38(2), 202-210.

Lord, G. (2006). Defining the indefinable: Study abroad and phonological memory abilities. In *Selected proceedings of the 7th conference on the acquisition Spanish and Portuguese as first and second languages* (pp. 40-46).

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, Cappa, V., S.F., Fazio, F., & Mehler, J. (1998). The bilingual brain. *Proficiency and Age of acquisition of the second language. Brain, 121*(10), 1841-1852.

Wechsler, D. (1944). *The Measurement of Adult Intelligence*. Williams & Wilkins.

Appendices

Appendix A. Phonological Memory Task, Sentences (Lord, 2006)

Students will hear each of these sentences read out loud by the researcher, and will be asked to repeat them one by one. Each sentence contains an invented word that does not exist in Spanish; therefore, the sentences are essentially meaningless. All sentences are between 10 and 15 syllables in length.

1. Mi mamá es una persona pabira y amable.
2. Se escucha el tabar de los obreros.
3. La duración del safilo es breve.
4. El hombre le da el maile al bueno chico.
5. El pequeño raito se metió en la cama.
6. Nunca sabes lo que puede pasar en un palaldo.
7. Encima del palcrue había una piedra.
8. Ella se sentía cada vez más maquel.
9. Ese pobre talatrei está loco.
10. En el degalo siempre hay muchas personas.
11. Si quieres probar un buen metaco, ven aquí.
12. Una dilosa vale más ahora que antes.
13. Con la llegada del fimesol todo se resolvió.
14. Para ser blarucio es necesario practicar mucho.
15. Es importante utilizar la bleimora aquí.
16. Con un sólo nobilu se enteró de todo.
17. “Voy a darte un zumil,” me prometió.
18. La familia es famosa por el cruate de antes.
19. Se la tovara funciona hoy, estamos bien.
20. La profesora es tan craitela que me gusta.

Appendix B. Participant Questionnaire (Pretest)

1. What is your gender?
    a. Male
    b. Female
2. What year are you in school?
    a. Freshman
    b. Sophomore
    c. Junior
    d. Senior
3. What option best describes you?
    a. Completed 6 hours of Spanish at the 200 level or above, not presently enrolled in a Spanish course
    b. Presently enrolled in a Spanish course at the 300 level or above
    c. Enrolled for Fall 2016 Study Abroad in Spain
4. Including your time at Butler, how many years of formal Spanish instruction have you received?
    a. 1 year
    b. 2 years
    c. 3 years
    d. 4 years
    e. 5 years
    f. 6 years
    g. 7 years or more
5. How many semesters of Spanish have you completed at the 300 level or above?
    a. 0
    b. 1
    c. 2

   d. 3
   e. 4
   f. 5
   g. 6
   h. 7

6. How long ago did you complete your last Spanish course at University?
   a. Last semester
   b. Two semesters ago
   c. Three semesters ago
   d. Four semesters ago
   e. Five semesters ago
   f. Six semesters ago
7. On average, during the past semester, how many hours a week were you exposed to/in contact with Spanish speakers? (This includes formal Spanish instruction, watching/reading/listening to Spanish media, and speaking with Spanish speakers)
   a. 0-1 hours
   b. 1-3 hours
   c. 3-5 hours
   d. 5-8 hours
   e. 8 hours or more
8. What languages are spoken in your home on a regular basis?
   a. English
   b. Spanish
   c. Other
9. Please provide three dates/times you would be available for a short video chat (5-10 minutes).
10. Please provide a Skype username or FaceTime number that I will be able to contact you with.

Appendix C. Participant Questionnaire (Posttest)

1. On average, during the past semester, how many hours a week were you exposed to/in contact with Spanish speakers? (This includes formal Spanish instruction, watching/reading/listening to Spanish media, and speaking with Spanish speakers.)
    a. 0-1 hours
    b. 1-3 hours
    c. 3-5 hours
    d. 5-8 hours
    e. 8 hours or more
2. How many years of formal Spanish instruction have you completed (not including this semester)?
3. Please provide three dates/times you would be available for a short video chat (5-10 minutes).
4. Please provide your randomized participant identification number given by the researcher.