EFFECT OF AGE OF ACQUISITION ON CONCEPT MEDIATION IN HERITAGE BILINGUALS

Iyad E. Ghanim

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Linguistics.

Chapel Hill
2016

Approved by:

Misha Becker

Elliott Moreton

Katya Pertsova

ABSTRACT

Iyad E. Ghanim: Effect of Age of Acquisition on Concept Mediation in Heritage Bilinguals
(Under the direction of Misha Becker).

Current models of bilingual lexical systems represent a shared conceptual domain and separate, language-dependent domains. Regarding the second language domain, researchers propose L2 words share a direct connection to the conceptual domain only for fluent bilinguals. Conversely, for non-fluent bilinguals, L2 words lack a direct conceptual connection and instead are connected via L1 translation equivalents. However, previous studies confounded age of acquisition with proficiency as variables that contribute to concept mediation.

The present thesis disentangles these variables' respective effects on developing concept mediation. Thirteen heritage Arabic-English bilinguals are subject to a picture-naming task and a translation task. Heritage speakers' response times match the concept mediation model irrespective of proficiency, with the exception of low proficiency speakers. These results indicate that for individuals who acquired a language at an early age, moderate loss of language proficiency may not remove lexico-conceptual links.

ACKNOWLEDGMENTS

I owe the completion of this work to the community of endless support and encouragement that surrounds me. I am grateful to the faculty of the Department of Linguistics at UNC Chapel Hill for their advice and insight: I am beyond privileged and humbled for the opportunity to have learned from you. Thank you also to Jennifer Arnold of the Psychology and Neuroscience Department at UNC and to Zeina G. Halabi of the Asian Studies department for their incredible assistance.

My gratitude also goes to Akhram Khater and the other members of the Moise A. Khayrallah Center for Lebanese Diaspora Studies, as well as the Triangle Lebanese-American center, whose commitment to research and cultural preservation is inspiring.

Thanks to my friends, whose thoughts and encouragement made the process easier, and to my family, who helped make the notion of earning a degree as a first-generation Arab-American a reality.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Bilingualism has long been a relatively understudied subfield in theoretical linguistics, and for that reason, lacks even a singular definition. Some researchers define bilingualism as the ability to speak and produce two languages simultaneously (Valdés, 2001); others include non-fluent bilinguals in this definition (Edwards, 2006); and yet others stipulate that 'bilingualism' is an umbrella term, covering individuals with linguistic command over three or more languages (Bhatia and Ritchie, 2006).

Much of the linguistic research on bilingualism has been focused on whether the speaker's two languages are connected in one single system or are separated into two language-specific systems (Volterra & Taeschner, 1978; De Houwer, 1998). The single- or dual-system hypotheses have been analyzed in the acquisition of phonology (Bosch and Sebastián-Gallés, 2001; 2003; Deuchar & Quay, 2000), morphosyntax (Yip & Matthews, 2000), among other linguistic subfields. A relatively understudied bilingual process, however, is that of lexical retrieval: the method by which speakers access and retrieve lexical information for production (Kroll & Stewart, 1994).

Current models of bilingual lexical retrieval represent a shared conceptual domain and two separate, language-dependent lexical domains. Of particular interest is how the lexical domain of the second language interacts with the lexical domain of the first language, as well as the conceptual domain, with the currently accepted view being that high L2 proficiency leads to

the development of links between lexical items and their respective concepts (Potter et al., 1984; Kroll & Curley, 1988; Chen & Leung, 1989; Jiang & Forster, 2001).

The model that depicts direct links between lexical items and their respective concepts is known as the concept mediation model, and is associated with high L2 proficiency (Potter et al., 1984; Kroll & Curley, 1988; Chen & Leung, 1989). Like the words of the dominant and native L1, the words of the L2 are understood by a direct connection to the concept they denote, rather than being understood via translation into the L1, which holds the connection to the conceptual representation. These studies compared low proficiency and high proficiency speakers and found that high proficiency speakers were concept mediators. However, in the few studies that did report L2 AoA, the ages at which the L2 was acquired were confounded with proficiency. Because early acquisition of a language all but ensures links from items to concepts, the effect of an early AoA on concept mediation without corresponding high proficiency needs to be studied.

Thusfar, no study has successfully investigated how much an early L2 AoA contributes to the development of lexico-semantic links. To prevent confounding variables, an early AoA would need to be coupled with low proficiency in the second language. This is an atypical trajectory in language acquisition — an early AoA usually ensures high proficiency in that language — but the linguistic circumstances of heritage speakers are ideal for testing the effect of an early AoA without high proficiency in developing lexico-conceptual links. Typically the children of immigrants, heritage speakers are characterized by having learned a family language at home that differs from the language of the society and, as a result of disuse, they have incompletely acquired the heritage language, or have partially forgotten it.

The present study focuses on the lexical organization of heritage Arabic bilinguals with a particular interest on the organization of their weaker language. For the purposes of this study,

heritage speakers provide a unique cross-section that allows research on the effect of a language with an early AoA, but without high proficiency. This work will help provide an understanding of the degree to which early-learned, but non-dominant languages are conceptually mediated, while also providing an understanding of some of the challenges faced by heritage bilinguals.

CHAPTER 2: BACKGROUND

2.1 Bilingualism & the bilingual lexicon

Early research on the structure of bilingual lexicons begins with Weinreich (1953), who proposes three "bilingualism types," which describe ways a given individual may be described as bilingual. The first type of bilingualism, compound bilingualism, refers to a system where the two lexical domains are separate, but each connected to a shared domain that houses conceptual and semantic information. Coordinate bilingualism, on the other hand, refers to a situation where each lexical domain is connected to separate language-dependent conceptual domains. Subordinative bilingualism is the case when the lexical items of the second language are only connected to their translation equivalents in the first language; in other words, the second language is only indirectly connected to the conceptual domain.

By using these models, Weinreich conflates two questions: first, do two language-specific conceptual systems interact with each language-specific lexicon, or conversely, is there only one shared conceptual system? And secondly: what is the extent of the interaction or connection between a given lexicon and conceptual domain, if any interaction exists at all?

Disentangling some of the implications of Weinreich's bilingual types, some studies suggest evidence for independence between two language representation (Brown, Sharma, & Kirsner, 1984; Gerard & Scarborough, 1989; Kirsner, Smith, Lockhart, King, & Jain, 1984; Kolers, 1963; Scarborough, Gerard & Cortese, 1984; from Kroll & Stewart, 1994). Other studies contrast this with evidence for shared a conceptual memory underlying the bilinguals' two

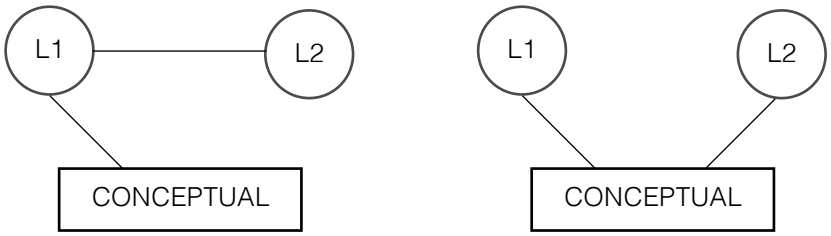
languages (Altarriba, 1990; Chen & Ng, 1989; Glanzer & Duarte, 1971; Meyer & Ruddy, 1974; Schwanenflugel & Rey, 1986; Tzelgov & Henik, 1989). Work by Potter (1979) and Snodgrass (1984), among others to be discussed, propose a single, abstract memory system where concepts are stored and accessed by separate lexical memory systems (Kroll & Stewart, 1994). This model of a shared conceptual domain is the currently-accepted understanding of the bilingual lexicon.

Yet additional research focuses on the organization of the bilingual lexicon; particularly, how the lexicons of a bilingual's two languages interact with each other, if they interact at all. Potter, So, von Eckhardt, and Feldman (1984) develop two models of bilingualism that differ on how the language domains connect to the conceptual domain or to each other: the word association model and the concept mediation model.

The word association model (WAM), similar to the Weinreich's subordinative model, posits that the second language (L2) is only connected to the first language (L1) through translation equivalents (also referred to as "lexical links," or "word-to-word associations") (Kroll & Stewart, 1994; Jiang & Forster, 2001). Critically, the L2 domain is not directly connected to the conceptual domain under this model. Instead, L2 words gain access to concepts only by mediation through the L1.

The second conceptualization of cross-linguistic connection is the concept mediation model which can be compared to Weinreich's compound bilingualism. Unlike the word association model, the concept mediation model (CMM) assumes second language words have direct access to concepts: a bilingual's two languages are connected through shared conceptual representations (Kroll & Stewart, 1994). Figure 1 visually depicts the contrasting word association model and the concept mediation model (adapted from Jiang & Forster, 2001).

FIGURE 1. Word association model (left) & concept mediation model (right).



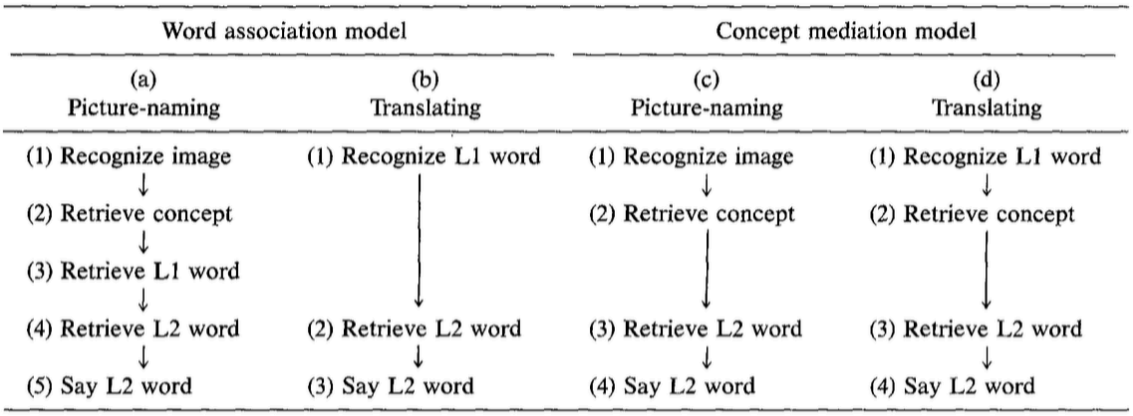
The difference between these two models is where the second language words are associated: either to first language words, as depicted in the word association model, or directly associated to concepts as in concept mediation. To test which of these models more accurately represent bilingual speakers' lexical systems, researchers have compared response latencies in picture naming and translation tasks (Potter et al., 1984). This insight came from prior studies, which have demonstrated that individuals can name words in a first language about 250 ms faster than naming pictures in the first language (Potter & Faulconer, 1975; Smith & Magee, 1980; Theirs & Amrhein, 1989). According to these authors, the reason subjects take longer to name pictures is that picture-naming requires access to concepts; word naming, on the other hand, does not require access to concepts, and therefore takes less time.

These models make different predictions about the speeds of picture naming and translation for bilinguals, as shown in Figure 1.2 (Potter et al., 1984). Under the word association model, translation into the second language is a shorter sequence than picture naming in the second language, therefore taking less time. In a picture naming task, the image prompt requires retrieval of the concept before retrieving the L1 word, and finally, retrieval and production of the L2 target word. Compared to a translation task, however, an L1 prompt bypasses conceptual retrieval and can directly access the L2 word for production. Because a translation task omits two

steps in the sequence, it takes a shorter length of time to translate into the L2 than name a picture in the L2 under a word-association model.

The concept mediation model, however, predicts little or no difference between picture prompts and L1-word prompts when the task is to produce the L2. Unlike the WAM, the producing an L2 word after viewing a picture prompt has the same number of steps as the process of translating from the L1 into the L2 under the concept mediation model. However, while the pathways are similar, the first step is not identical.

FIGURE 1.2. L2 picture-naming & L1 to L2 translation in the two models (Potter et al., 1984:26).



Given these models, Potter et al. (1984) compared translation and picture naming response times in a group of highly-fluent Chinese-English bilinguals. Potter et al. (1984) hypothesized that if the time to translate into the L2 was faster than picture naming in the L2, then a participant's L2 system relied on lexical links to the L1, and did not have conceptual access (WAM). Conversely, the concept mediation model would be indicated by speeds of translation into L2 that were similar to picture naming into the L2, because both processes require conceptual access to retrieve the L2 word.

The results of Potter et al. (1984) clearly supported the concept mediation model, with the bilinguals taking about the same amount of time to translate words from L1 to L2 as naming pictures in L2. Potter et al. (1984) interpreted these results to indicate that both the translation and picture naming tasks followed a trajectory through a shared conceptual system: therefore, both the L1 and L2 lexical systems were conceptually mediated. To see if L2 fluency was the factor that determines the form of a bilingual's connection, Potter et al. (1984) also tested a group of less-fluent English-French bilinguals. Surprisingly, the results for the less-fluent group also supported the concept mediation model.

The conclusions offered by Potter et al. (1984) suggested that concepts mediated translation equivalents for all bilinguals in any given level of fluency. Kroll & Curley (1988) challenged this conclusion, suggesting that the responses of lower-proficiency bilinguals should be different from more fluent bilinguals. They suggested that the low-proficiency bilinguals in the Potter et al. study (1984) had already passed the word-associative point of second language development, where lexical links mediate the processing of second language words. Speculating that the English-French bilinguals were already *too* fluent, Kroll & Curley used a wider range of bilingual subjects, including some with less than 2 years of language experience (1988). Following a similar procedure as in the Potter et al. (1984) study, Kroll & Curley (1988) demonstrated that subjects who studied an L2 for less than two years had results consistent with word-association predictions. In other words, for truly non-proficient second-language learners, translation into an L2 was faster than picture naming in an L2 (Kroll & Curley, 1988; Jiang & Forster, 2001). These conclusions provided support for a developmental hypothesis (also referred to as "intermediate hypothesis") which states that a non-proficient bilingual's L2 lexicon relies on word-to-word connections to the L1 during early stages of L2 acquisition (Gekoski, 1980; Opoku, 1983, Chen & Leung, 1989). According to these authors, after achieving a certain level

of proficiency, words of the L2 system shift from relying on word-to-word L1 connections to relying on direct connection to their respective concepts.

However, these studies confound age of acquisition with proficiency. High proficiency in a language is often associated with an early AoA in that language, while low proficiency is associated with a later AoA, when considering the typical acquisition trajectory (Deuchar & Quay, 2000; Jiang, 2000). Therefore, given the results of these studies, the links from the lexicon to the conceptual space (as in the CMM) may develop not with proficiency, but instead, as a factor of the age at which a language is acquired.

To resolve this confound, a study by Chen & Leung (1989) replicated the Potter et al. (1984) study with the same low-proficiency and high-proficiency groups. The researchers also included a third group of participants, child beginners, whose low proficiency was coupled with an early age of acquisition. Their rationale is clear: any differences in the performance of child beginners to adult beginners is attributable to age of acquisition of the L2 rather than proficiency. However, significant differences between the child and adult L2 learners rendered their results indiscernible. Not only were the ages of L2 acquisition different between the child and adult learners, but the children's ages, cognitive abilities, and length of experience with the language differed significantly from the adults' as well.

The present thesis instead uses adult heritage speakers to decouple the effects of AoA and proficiency rather than child learners of a second language. Characterized by an early age of acquisition paired with low proficiency, the response rates of heritage speakers can be compared with that of late AoA/low-proficiency speakers and early AoA/high-proficiency speakers to observe the effects of each variable independently. Further, using adult heritage speakers reduces

the variable effects of age, cognition, and length of language experience that made a direct comparison impossible in Chen & Leung's study (1989).

One study by Silverberg & Samuel (2004) also investigated the effect of L2 AoA on lexical access in bilinguals using a lexical priming experiment. The rationale for conducting a priming experiment is similar to the picture-naming and translation studies: an L1 target word that is semantically-related to the L2 prime word should be accessed faster and more accurately if the lexicons are conceptually mediated. The word association model predicts no effect of a prime on a target. Using three groups of L1 English, L2 Spanish bilinguals (early L2 learners of Spanish, late L2 learners with high proficiency, and late L2 learners with low proficiency), Silverberg & Samuel (2004) found that only the group with an early AoA demonstrated semantic priming effects, as explained with the conceptual-mediation model. These results seem to suggest that an early age of acquisition may be associated with a conceptual connection, for which late beginners simply exhibit no evidence.

As further rationale for why early acquisition may lead to the development of lexico-semantic links, I consider the case of sequential childhood bilinguals, who acquired one language exclusively for a period of time before beginning exposure to a second language during early childhood. For these individuals, it is safe to assume that early exposure to the first language results in the development of direct links to the conceptual domain. This follows the trajectory of monolingual language acquisition, where words are directly associated with their semantic and conceptual meaning as they are acquired by children.

Two scenarios are theoretically possible for the L2, which was acquired later in childhood: the L2 words can be associated to the translation equivalents of first language, as in the word-association model. This is plausible, given the fact that it is the second language

acquired and, in theory, is the non-dominant language for the speaker. Conversely, though, it is also possible that the words of the second language can earn a direct connection to the conceptual store if it becomes the more dominant language. This illustrates the trajectory of a heritage speaker.

For heritage speakers, acquisition of the first language either halts or regresses, and the L2 becomes the dominant language (Valdés, 2000). For that reason, despite acquiring it at an early age, they demonstrate minimal proficiency in their L1. Their linguistic circumstances serve as a testbed for the effects of an early AoA without high proficiency inasmuch as it relates to developing lexico-semantic links (Scontras et al., 2015). I will proceed to discuss the particular circumstances that lead to heritage bilingualism, what constitutes a heritage language, and the defining characteristics of its speakers.

2.2 Heritage Bilingualism

Heritage speakers are individuals who were raised in places where the home language differs from the language that is spoken in the dominant community (Valdés, 2000; Rothman, 2009). As expected, exposure to both the home language (termed the "heritage language") and the dominant community language grants them some degree of bilingualism. They have normal and natural exposure to a home language that should, by all typical accounts, lead to complete acquisition and full command of a language. What distinguishes heritage speakers from other bilinguals is the fact that, despite early exposure to a language, heritage speakers have a critically low proficiency in the heritage language which often makes it difficult to produce the heritage language at all (Valdés, 2000; Polinsky, 2006).

Most frequently, children of immigrant families grow up to be heritage speakers, irrespective of whether exposure to the societal language occurs concurrently with, or after, the heritage language (Ammerlaan, 1996; Hulsen, 2000; Polinsky, 2005). In either case, whether exposure to what will become the stronger language occurs simultaneously or successively, the loss of dominance in the heritage language results (Scontras et al., 2015).

The aspects that contribute to such a low level of grammatical competence for otherwise typical simultaneous bilinguals are not just linguistic, but inherently social, cultural, and psychological in nature. A child bilingual who begins to socialize with speakers of the majority language necessarily decreases in their use of the minority language; as use decreases, productive abilities in the heritage language structurally and functionally degrade (de Bot, 1996; Scontras et al., 2015). As the child responds more frequently in the majority language, the caretakers' use of the heritage language may decrease as well, deteriorating the child's receptive competence in the language (Scontras, 2015; Rothman, 2009).

The literature originally introduced this pattern as "incomplete acquisition" (Polinsky, 2006; Montrul, 2008; Benmamoun et al., 2013). Several researchers have challenged the use of this term, writing that its use entails several questionable assumptions. For example, Pascual y Cabo and Rothman (2009) point out that a heritage speaker's trajectory of acquisition is "incomplete" only relative to that of monolinguals. In fact, the researchers argue that because a heritage speaker's input is inherently different from that of monolinguals, a comparison cannot be justifiably made. No longitudinal studies show a heritage speaker's path of acquisition, up until the point of exposure to a majority language, to be in perfect synchrony with the path of a monolingual (Pascual y Cabo and Rothman, 2009). Following Scontras et al., (2015), I will

henceforth refer to the phenomenon where heritage bilinguals receive divergent linguistic input as "divergent attainment."

Alongside divergent attainment, attrition is a second reason why heritage speakers' grammatical competency is lacking. Attrition implies that full mastery of a language or grammatical structure had been reached before a change in the relative use of one language causes the weakening or full erosion of grammatical structures (Seliger, 1996; Polinsky, 2006; Rothman, 2009). Several factors affect the likelihood that attrition will occur, the rate of language loss, and which grammatical features may be attrited. One factor is the age of onset of bilingualism, which is inversely related to the extent of attrition; if a child is to acquire an L2 at a later age, the less likely attrition is to occur (Pallier, 2007; Montrul, 2008). This is a logical correlation, given the knowledge that language skills tend to stabilize and solidify with an individual's age. As a result, prepubescent children who have experienced a language shift tend to lose language skills at a greater rate and to a greater extent than individuals who have gone through the experience as adults (Ammerlaan, 1996; Hulsen, 2000; Pavlenko & Malt, 2011).

Whether it is by divergent attainment (systematic disuse of a language) or attrition (language loss), the language experience of heritage speakers at some point diverges from that of their bilingual peers (Pascual y Cabo & Rothman, 2009). What results is systematic degradation of linguistic structures and a loss of dominance in the heritage language.

What is specifically meant by "loss of language skills" is a relatively understudied area of research, perhaps owing to the range of proficiency levels that vary by individual. The developmental trajectories of heritage bilinguals, it is clear enough, differ substantially from fully "bi-fluent" and balanced bilinguals, as well as from a native monolingual in the heritage language. Case studies on heritage bilinguals reveal these deficits are largely morphosyntactic in

nature, specifically a loss of grammatical contrast, omissions of function words and morphemes, and reduction in allomorphic variation (Andersen, 1982; Bullock & Toribio, 2006; Hagen & de Bot, 1990; Maher, 1991; Polinsky, 2006a, b; Schmid, 2002; from Sherkina-Lieber 2011). Data on heritage-Russian speakers in the United States revealed a shift in preference toward rigid word order, slower speech rate, and greater length and frequency of pauses, in addition to general misuse of words (Polinsky, 2006).

Lexical access is a language skill that relies heavily on experience and consistent use (de Bot, 1998; Weltens & Grendel, 1993). As a result, low-levels of language activation — productively and/or receptively — put the lexical retrieval facilities at an early and increased risk of erosion or complete attrition. Studies that have observed heritage speakers have found lexical inaccessibility and retrieval difficulties manifested in high error rates and slowed processing, especially with lower-frequency words (Hulsen, 2000). Polinsky's analysis of heritage speakers of Russian in the United States reveals significantly slower speech rate, extensive code switching, and pauses even between elements of the same constituent (2006). Both Hulsen (2000) and Polinsky (2006) cite these patterns as evidence of lexical challenges due to disuse.

The fact that heritage speakers may be experiencing lexical retrieval difficulties is interesting, but also unaccounted for given an understanding of lexical retrieval processes. If we follow the previous studies' understanding of proficiency as it relates to developing concept mediation, we can predict the lexicon of heritage speakers (who demonstrate markedly low proficiency) to be word-associative in nature (Kroll & Curley, 1988; Chen & Leung, 1989). However, this is improbable considering what is expected from childhood first-language learning. During acquisition, speakers ought to have developed lexico-semantic links as the words of the heritage language were acquired.

Few studies used heritage speakers to investigate the effects of AoA on developing links between lexical items and their respective concepts. One such study by Montrul & Foote (2012) looked at the individual and interactive effects of word-level AoA in addition to language AoA (typically called "global AoA"; Montrul & Foote, 2012). In other words, rather than just observing the age at which an individual is first exposed to a language, this study also looked at the age at which each word is first learned and its individual conceptual connection. The two groups of the Montrul & Foote (2012) study were English-Spanish bilinguals who differed in their global AoA of Spanish, but who were both dominant in English (i.e., L2 Spanish learners and heritage Spanish speakers). Each group conducted a lexical decision task in Spanish, as well as an English-Spanish translation decision task. The researchers found that the L2 learners responded just as accurately, but far more quickly than the heritage speakers.

Relevant to the present study, these results cannot inform which model — word association or concept mediation — better characterizes the lexical organization of L2 learners compared to heritage speakers. Only the results of the English to Spanish translation decision task indicated any difference between the heritage speakers and that of the L2 learners. It is impossible to determine whether this is attributable to the heritage speakers' lexical retrieval failures, or some other grammatical failure. Even still, Montrul & Foote's (2012) results do indicate a difference between speakers who acquired their language in childhood and L2 learners. The operative variable between these two speakers is global AoA.

Research shows the advantageous effects of early learning on some aspects of phonology (Au et al., 2002) and lexical semantics (Montrul, 2005). What is missing, however, is research regarding the effects of early acquisition on lexical access. Current research that deals with this

gap either does not use heritage bilinguals, or otherwise, produces results inapplicable to the developmental hypothesis.

The goal of this current study is to distinguish the effects of AoA from proficiency by approximating the procedures of the Chen & Leung study (1989). Chen & Leung (1989) compared the response latencies of a picture naming and translation task with young bilinguals. Similarly to Montrul (2012), I will use heritage bilinguals as participants in a picture-naming and translation study. I predict that heritage speakers' response times will match the predictions of the concept mediation model, with L2 picture-naming taking about the same amount of time as translation to L2. This would suggest that even for individuals who have lost proficiency in a language, the fact that a language is acquired during an early age ensures lexico-conceptual links.

I acknowledge that the use of the labels L1 and L2 to reference the order of the languages' acquisition may unintentionally imply language dominance. This may be misleading in the case of heritage speaker, where the "L1" becomes the weaker language and the "L2" becomes the dominant language. For bilinguals who have acquired two languages simultaneously, the use of these labels is even more misleading. Therefore, I will henceforth refer to the heritage language as the "weaker language" (WL) for clarity, and the societal language — which is second-acquired for successive bilinguals — as the language that became the "dominant language" (DL).

CHAPTER 3: EXPERIMENTAL DESIGN

3.1 Participants

The participants were eleven college-aged individuals who learned Arabic as a heritage language in United States communities where the societal language is English. Seventeen subjects were recruited for the study, with thirteen being eligible or continuing the study to completion. These subjects ranged from very low-proficiency speakers to very high-proficiency, fully-balanced bilinguals.

Preference was placed on selecting participants who were exposed to a Levantine Arabic dialect (*al-Shami*). In order to avoid dialect variation as a confounding variable, the lexical tokens involved in the study were chosen to be dialect non-specific.

Participants were recruited from Arabic-speaking University groups or clubs and Arabic culture clubs from the University of North Carolina at Chapel Hill, North Carolina State University, or Duke University. Further, access to some individuals was provided through local cultural and language organizations; these organizations include the Moise A. Kharyallah Center for Lebanese Diaspora Studies (NC State); the Triangle Lebanese-American Center, and the UNC Mideast Center.

In addition to the response times, the biographical information of the participants was also analyzed to look for interesting trends. Four individuals (107, 109, 110, 114) learned Arabic

and English concurrently in the United States, with exposure to both languages beginning simultaneously in infancy. Seven participants reported exposure to English beginning sometime after exposure to Arabic between the ages of 0-5 years old, one of whom reported exposure to English beginning between the ages of 5-10 years old.

TABLE 3.1. Participant biographical data.

Participant	*Arabic proficiency*	*Began English exposure*	*Ended Arabic exposure*
114	(1/10)	birth	continuing
103	(3/10)	0-5 y/o	continuing
115	(4/10)	0-5 y/o	0-5 y/o
101	(4/10)	0-5 y/o	continuing
102	(5/10)	0-5 y/o	0-5 y/o
113	(5/10)	0-5 y/o	continuing
110	(6/10)	birth	continuing
105	(7/10)	0-5 y/o	continuing
109	(8/10)	birth	continuing
111	(9/10)	birth	continuing
117	(10/10)	5-10 y/o	continuing

One of the three participants (117) who reported simultaneous language exposure was born in the United States, but spent the entirety of his childhood (infancy to age 16) in an Arabic-speaking country where he learned Arabic and English simultaneously. At the time of the study, he had been living in the United States for over a year, and reported dual dominance in both Arabic and English. In this case, while English is by definition a heritage language for this participant, he identified as and was considered a highly-proficient and balanced heritage bilingual.

3.2 Stimuli

Picture stimuli and spoken-word stimuli were used for the experiment. Picture stimuli consisted of a total of 46 images representing nouns from the Khawaileh et al. (2014) database of

Arabic nouns with corresponding images in the Snodgrass & Vanderwart (1980) norms, as described by Sholl et al. (1995). Of the total spoken-word stimuli, 24 were from the Arabic database with corresponding images in the picture naming norms; the remaining 22 did not have corresponding images in the picture naming norms.

Extracting only the tokens from the Khawaileh database with the Snodgrass & Vanderwart norms for the picture naming tasks resulted in consistent line-drawn images of household objects, animals, and other nouns (a full list is provided in the appendix). Further, use of the Arabic database allowed tokens in all tasks to be stabilized in aspects such as the age at which the word is acquired (AoA), participant agreeability on the objects name (name agreement), conceptual and visual complexity of the item being named (visual complexity/naming latency), etc. (Khwaileh, et al. 2014). This was designed to reduce extraneous variables that could affect the duration of retrieval in the tasks.

3.3 Design

The experiment involved four total tasks: picture-naming in English, picture-naming in Arabic, translation into English, and translation into Arabic. These tasks were split across two blocks, with one block containing the picture naming tasks, and the other block containing the translation tasks. The task type was blocked, and the order of the blocks were counterbalanced. Crucially, this means a participant completed one type of task in both language directions before continuing to the next task type.

Each block of the picture naming phase consisted of 22 line drawings of standard objects presented in succession. Each phase also included two "practice" items to demonstrate the task and were not included in the analysis. At the onset of the presentation of each line drawing

stimulus, a 400 Hz tone was heard for 400 ms. Participants named the item out loud into a recording device, and then pressed the space-bar to continue to the next token. The key press triggered an inter-stimulus interval of 500 ms before the presentation of the next token.

In the Translation phase, one block consisted of 22 DL items that the participant translated into the WL. The other block consisted of 22 WL items that the participant translated into the DL. Because heritage language speakers have inconsistent language reading abilities, tokens in the translation phase were auditory prompts recorded by an American college-aged male who was a fully-balanced native speaker of Lebanese Arabic. The rationale for the auditory stimuli is that heritage speakers often demonstrate inconsistent language reading abilities, if any at all (Polinsky, 2006). Difficulties reading the script would add time to the response latencies, which are only intended to measure cognitive processing.

Blocks and phases were separated by short filler paragraphs in the DL with one or two comprehension questions for a total of three filler tasks. This served to "reset" the participant to their dominant language and reduce order and practice effects. The total of four blocks and fillers were presented to participants on a 14.7 inch laptop monitor using PsychoPy 2 experimental display software (Pierce, 2007; 2009). The study was conducted in one of two locations: an office space at UNC at Chapel Hill, or an office space at the Triangle Lebanese-American cultural center.

3.4 Method

Prior to the experimental phase, participants completed a short language history survey to validate and compare their language circumstances. The purpose of this survey was three-fold: first, to identify the participants as eligible heritage speakers; second, to collect descriptive

information about the participants' language input and circumstances; and third, to provide participants with an opportunity to self-report their language proficiency (Albirini, 2013; De Houwer, 2007; Godson, 2003, from Polinsky, 2008). This information was critical in not only identifying heritage speakers, but also to be able to connect biographical factors to linguistic effects (De Houwer, 2007; Polinsky, 2006).

After giving informed consent and asking questions, participants read the instructions provided and began the experiment, consisting of two picture naming tasks and the two translation tasks. For all phases, responses were recorded via voice-recording software on a mobile device and were then uploaded into Praat for response latency measurement (Boersma et al., 2013). Response latencies for the picture naming tasks were measured from the onset of the tone, which was shown concurrently with the picture, to the onset of the participant's spoken response of the correct item. The participant's correct response does not include fillers such as "um" or "ahh"; as such, these articulations were ignored and counted as latency. False starts, where a participant begins an articulation, halts, and begins the word again, were also counted as latency and not included as a response onset. Accuracy of response was also recorded and measured as either a failed response (an "I don't know," "pass," or "I don't remember") or an erroneous (non-target) response, in addition to response latency. Failed and erroneous responses were excluded from analysis.

Following the experiment, participants completed a formal proficiency assessment consisting of 10 grammaticality judgments. Participants were presented with 10 randomly ordered sentences, 5 of which were grammatical, and 5 of which were ungrammatical. These sentences captured the participants' ability to detect linguistic errors in a range of grammar constructions and difficulties. These included: errors in verb number or gender agreement with

its subject; errors in adjective gender agreement with a noun; errors in the formation of plurals (i.e., the incorrect use of a plural suffix for an irregular noun); errors in which preposition to use for a verb; and errors in the use of a definite article in a nominal construction (see Appendix for complete list). Participants' accuracies (correct judgments out of ten total judgments) were recorded.

The biographical information reported in the survey confirmed the participants' status as heritage bilinguals. An analysis of their actual responses during the experimental tasks confirmed this, being in line with observations set forth by Polinsky (2001, 2006), Sherkina-Lieber (2011), and others. Phonologically, many of the extremely low-proficiency speakers were not able to produce segments not present in English, such as the pharyngeal fricatives in words as "apple" (/təfaħ/) or "eye" (/ʕɛjn/), though the intermediate and higher proficiency subjects were able to do so easily. Participants often deleted them and compensated with a long vowel, glottal fricative *h*, or stop (i.e., /təfa:/, /ʔɛjn/). In terms of the lexicon, all but the highest proficiency heritage speakers produced false starts and tip-of-the-tongue experiences, as well as indications that they could not recall the target word. These patterns appeared to correlate with proficiency, with lower proficiency speakers having longer and more frequent recall difficulties.

CHAPTER 4: RESULTS & DISCUSSION

Two dependent variables were measured for the purposes of this study: first, the length of time until the speaker's response, and second, the accuracy of their response. I present the reaction time results followed by the accuracy results; finally, I present a discussion of the outcomes.

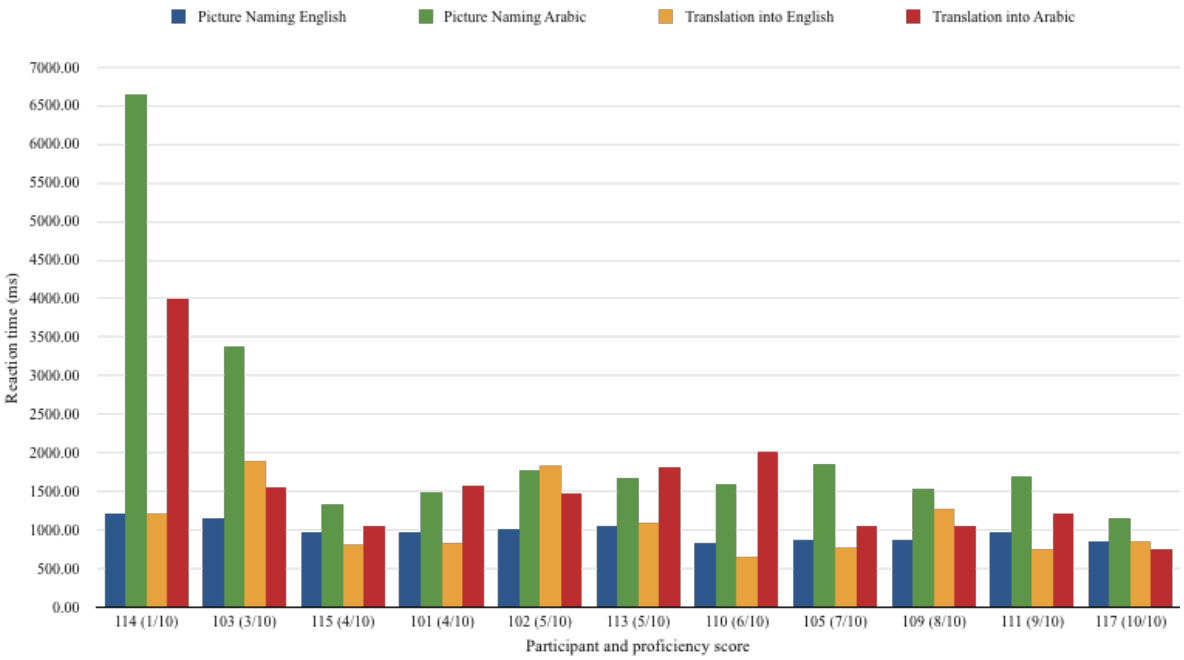
4.1 Reaction time

The hypothesis of this study was that heritage speakers' reaction time will match the predictions of the concept mediation model, with WL picture-naming taking about the same time as translation from the DL to the WL. The reaction times of each participant are presented in Table 4.1 and Figure 4.1.

TABLE 4.1. Raw reaction time (s) for each participant.

Participant	*Proficiency*	*Picture Naming English*	*Picture Naming Arabic*	*Translation into English*	*Translation into Arabic*
114	1/10	1.211	6.648	1.207	3.994
103	3/10	1.145	3.371	1.885	1.555
115	4/10	0.978	1.334	0.816	1.058
101	4/10	0.976	1.489	0.824	1.582
102	5/10	1.017	1.777	1.829	1.473
113	5/10	1.044	1.667	1.093	1.812
110	6/10	0.829	1.597	0.659	2.014
105	7/10	0.869	1.849	0.764	1.061
109	8/10	0.872	1.533	1.272	1.049
111	9/10	0.965	1.698	0.745	1.221
117	10/10	0.853	1.161	0.845	0.741
Raw mean	**--**	**0.978**	**2.193**	**1.085**	**1.597**

FIGURE 4.1. Response latencies for each task (ms).



Response times to the picture naming in English task (PN-E) were low, with the mean response rate attenuated for missingness in the data being the lowest of the tasks at 973 ms. This falls in line with response rates of monolinguals naming pictures, or proficient bilinguals naming pictures in their dominant language as presented in previous studies (Jiang & Forster, 2001; Kroll & Stewart, 1994; Sholl et al., 1995). In Table 4.1.2, I present the mean response rates of each task across all participants.

Mean response times to the picture naming in Arabic task (PN-A) were, conversely, the highest among the four tasks at 1881 ms, which is almost double the response rate of PN-E. This is also as expected, given the difficulties with the WL that the heritage speakers experience.

Mean response rate for the translation into English task (Tr-E) and the translation into Arabic task (Tr-A) were similar, at 1109 ms and 1468 ms respectively. Similarly to these results,

the results of translation tasks in previous research show translation into the WL taking longer than the DL, if the two translation directions are not equivalent in speed (Kroll & Stewart, 1994).

TABLE 4.1.2. Mean reaction times (s) attenuated for missingess per task

Task	*Estimate (s)*	*Standard Error*
Picture Naming Arabic	1.8811	0.1245
Picture Naming English	0.9736	0.1150
Translation into Arabic	1.4687	0.1245
Translation into English	1.1098	0.1203

Because of the large variance between the subjects' responses, I investigated low proficiency speakers, intermediate proficiency speakers, and high proficiency speakers separately to identify a main effect of proficiency level. Individuals were categorized in low, intermediate, or high proficiency based on their scores in the grammaticality judgments in the post-hoc proficiency assessment. Low proficiency speakers scored between 1 and 3 out of 10 (participants 114 and 103); intermediate proficiency speakers scored between 4 and 6 out of 10 (participants 101, 102, 113, and 110); and high proficiency speakers scored 7 to 10 out of 10 (105, 109, 111, and 117). The reaction times for each task per group is presented in Table 4.1.3.

TABLE 4.1.3. Raw scores for each participant group (seconds)

Participant Group	*Included Proficiencies*	*Picture Naming English*	*Picture Naming Arabic*	*Translation into English*	*Translation into Arabic*
Low	(1-3/10)	1.178	4.603	1.546	2.299
Intermediate	(4-6/10)	0.969	1.580	1.044	1.614
High	(7-10/10)	0.890	1.569	0.907	1.018

Picture naming in English took the shortest amount of time for all participant groups. This is expected given that it is participants' dominant language. For our analysis, picture

naming in English served as a baseline against which I compare the reaction times of the other tasks.

The picture naming in Arabic task was 3425 ms slower than the picture naming in English task for low proficiency speakers, 611 ms slower for intermediate proficiency speakers, and 679 ms slower for high proficiency speakers. PN-A for lowest proficiency speakers took the longest amount of time among all tasks and for all proficiency groups. This seems to indicate that low proficiency speakers have trouble accessing a WL word when presented with a picture prompt.

The translation into English task for high proficiency speakers appeared to be the fastest task compared to any task for any proficiency level. The fact that English is the dominant language of these speakers, paired with their high proficiency in Arabic, the short speeds for L2 to L1 translation is expected and concurs with previous research (Kroll & Stewart, 1994; Jiang & Forster, 2001).

Translation into Arabic took longer than translation into English for all participants (low: 1230 ms longer; intermediate: 544 ms longer; high: 112 ms longer). Because of their status as heritage speakers (i.e., Arabic is their non-dominant language), it is expected that translating from the stronger into the weaker language would be slower than the reverse. Further, the difference in translation speeds appears to correlate with proficiency: the additional length of time to translate into the WL decreases — approaching the time to translate into the DL — as proficiency increases.

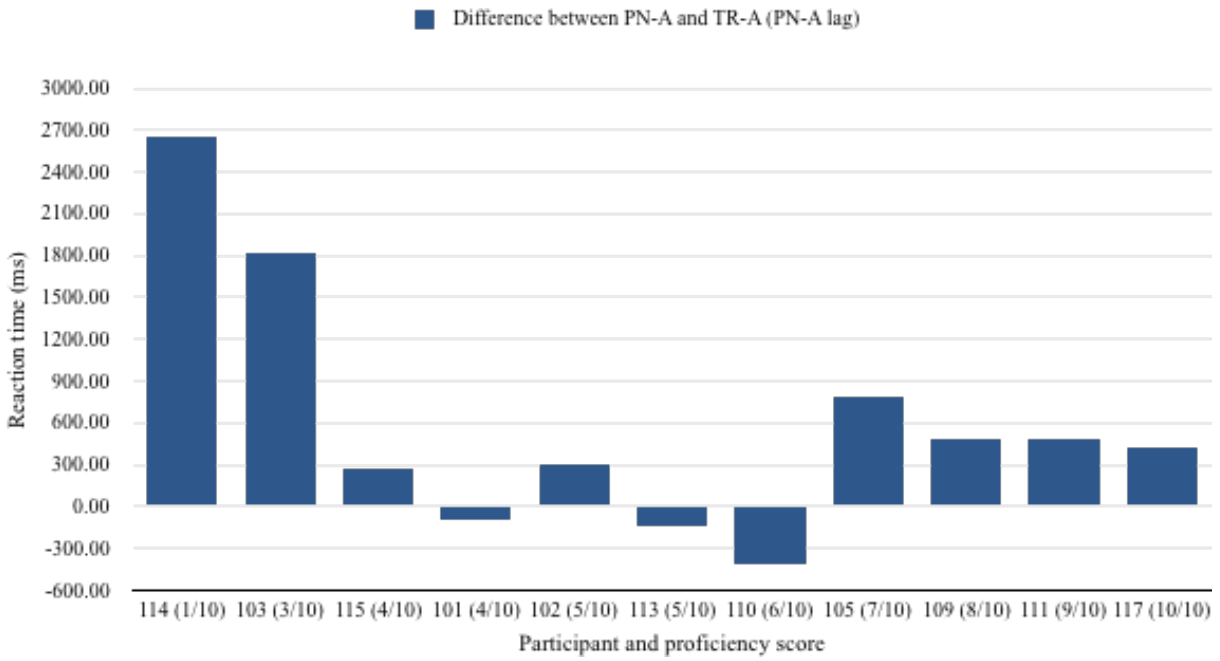
To test if heritage speakers show evidence of concept mediation, I investigated the difference between the PN-A and Tr-A tasks, which is key to indicating which lexical model

describes each participant. While PN-A took an average 412.4 ms longer, the variance in the subjects was large, with some participants demonstrating an additional 2653 ms lag on the PN-A task. The means for the two Arabic tasks per proficiency group, extracted from Table 4.1.3, are presented in Table 4.1.4 and Figure 4.1.4 alongside the difference in reaction time between the two tasks.

TABLE 4.1.4. Group means & picture naming lag

Proficiency	*PN-A (s)*	*Tr-A (s)*	*Difference*
Low	4.6032	2.2989	2.3043
Intermediate	1.5796	1.6143	-0.0347
High	1.5690	1.0181	0.5508

FIGURE 4.1.4. Picture naming lag for each participant



For the lowest proficiency speakers, the difference between the picture naming in Arabic task and the translation into Arabic task is larger than both other groups. Surprisingly, the absolute mean difference in the task speeds for intermediate proficiency speakers (34.7 ms) is

closer to zero than the means of the other groups (2304.3 ms for low proficiency speakers, and 550.8 ms for high proficiency speakers).

The fact that intermediate speakers PN-A lag is close to zero can indicate that a concept mediation model is likely characteristic of the intermediate proficiency speakers. Conversely, the large lag in the PN-A task for the low proficiency speakers is accounted for in a word-associative model, in which the picture prompt must be mediated through the DL system first before production in the WL. For analysis, I subjected these data to a 2 (task type) X 2 (language) X 3 (proficiency) mixed effects model to evaluate the hypothesis regarding the main effect of proficiency on reaction time, with random effects for multiple observations within the subject, and conducted a two-tailed *T*-test.

For the lowest proficiency participants, the reaction time of the Arabic picture naming task (μ=4.6032 s, *SE*=0.2767) was significantly slower than the reaction time for the translation into Arabic task (μ=2.2989 s, *SE*=0.2265); $t(739)=7.05$, $p<.0001$. For the subjects with intermediate proficiency, the reaction time of the picture naming task (μ=1.5796 s, *SE*=0.1235) was not significantly different from the translation task (μ=1.6143 s, *SE*=0.1244), indicating a concept-mediation model; $t(739)=-0.23$, $p=0.8175$. However, for the highest proficiency subjects, the reaction time of the picture naming task (μ=1.5690 s, *SE*=0.1188) was significantly slower than that of the translation task (μ=1.0181 s, *SE*=0.1224); $t(739)=3.98$, $p<.0001$. The results of the task contrasts per group are presented in Table 4.1.5.

TABLE 4.1.5: Analysis of within-group means (s) using a mixed effects model

Mean	*Estimate*	*Standard Error*	*DF*	*T-Value*	*Pr > |t|*
PN-A vs. Tr-A (low prof.)	2.3043	0.3270	739	7.05	<.0001
PN-A vs. Tr-A (intermediate prof.)	-0.0347	0.1503	739	-0.23	0.8175
PN-A vs. Tr-A (high prof.)	0.5508	0.1383	739	3.98	<.0001

Because of the unexpected results of the high-proficiency group in support of word-association, I also analyzed the difference in the mean picture-naming task response rate and the mean translation task response rate between proficiency groups.

As presented in Table 4.1.6, the lower proficiency speakers' difference in the picture-naming task and the translation task (i.e., the picture-naming lag) (μ=2.3043 s, *SE*=0.3270) was significantly different from that of both the intermediate proficiency speakers ($t(739)=6.50$, $p<.0001$) and the high proficiency speakers ($t(739)=4.94$, $p<.0001$). Additionally, the differences in the picture-naming task and the translation task for the intermediate proficiency speakers (μ=-0.03470 s, *SE*=0.1503) compared to that of the highest proficiency speakers (μ=0.5508 s, *SE*=0.1383) (105, 109, 111, and 117) was also significant; $t(739)=-2.87$, $p=0.0043$.

TABLE 4.1.6: Analysis of between-group means (s) using a mixed effects model

Mean	*Estimate*	*Standard Error*	*DF*	*T-Value*	*Pr > |t|*
Difference in low proficiency vs. intermediate prof. PN-A lag	2.3390	0.3599	739	6.50	<.0001
Difference in intermediate prof. vs. high proficiency PN-A lag	-0.5855	0.2042	739	-2.87	0.0043
Difference in low proficiency vs. high proficiency PN-A lag	1.7534	0.3551	739	4.94	<.0001

4.1.2 Discussion

Analyzing the difference between the responses of the PN-A task and the Tr-A task was key to determining participants' lexical organizations. For concept mediators, this difference should be very close to zero, because these individuals use the same path through the conceptual domain for both translation and for picture naming tasks. In other words, both tasks should take

just about the same length of time. For word associative models, PN-A should take a significantly longer time than translation into Arabic.

The results of the intermediate-proficiency speakers were consistent with the concept-mediation model, which predicts the times for the tasks to be approximately equal to each other. The fact that heritage speakers achieved concept mediation even with only intermediate proficiency provides evidence that age of acquisition may have an effect on WL structure. This differs from the results of previous studies, in which only speakers with a conversational proficiency showed evidence of concept mediation (Potter et al., 1984; Kroll & Curley, 1988; Chen & Leung, 1989).

However, for the very lowest proficiency speakers (114, 103), picture naming took far longer than translation, providing support for Arabic lexical items to be word-associated to English lexical items. The fact that the lower-proficiency heritage speakers appeared to be word-associating their WL words shows that age of acquisition alone does not necessarily grant conceptual mediation. These results could, however, suggest an interactive effect between proficiency and AoA in developing conceptual mediation. It is likely that for a language acquired during childhood, the level of proficiency that must be reached before developing conceptual mediation is different than for a language learned in adulthood matched for proficiency. This would explain the difference in the results between the lower proficiency heritage speakers compared to the intermediate proficiency speakers.

An interactive model as described does not fully explain the response rates of the higher proficiency speakers which more-closely resembles those of the lower-proficiency speakers. Interestingly, the results of the highest proficiency subjects (105, 109, 111, and 117) yielded a statistical significance in the difference between the picture naming and the translation tasks, in

that picture naming took significantly longer than translation. This is an unexpected result, because it seems to suggest a word-association model for a group of speakers who, by all means, should most closely approximate fully-fluent bilinguals of all the groups. In fact, the previous studies that attributed proficiency to concept mediation would predict these speakers to be concept mediators, which is not what this statistical analysis suggests.

On the other hand, the mean difference in picture-naming and translation for the highest proficiency speakers (μ= 0.5508 s) is not as large as that for the lowest proficiency speakers (μ=2.304 s). Further, additional analyses did confirm a significant difference between the means of the low and high proficiency speakers as well as between the means of the intermediate and the high proficiency group (Table 4.1.6). Because of their results, it is highly probable that the response rates of the high proficiency heritage speakers are anomalous. Additional research is required to investigate and confirm this result more closely.

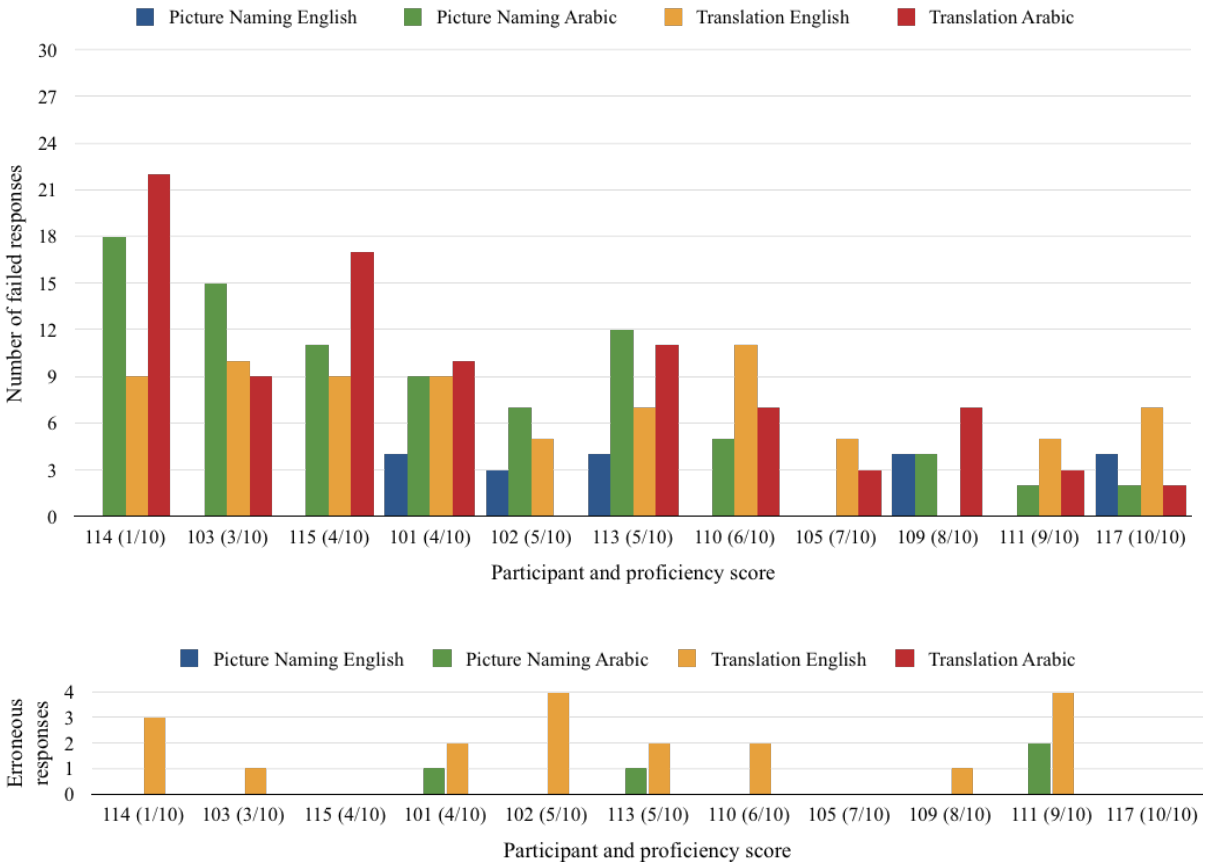
4.2 Response accuracy

To quantify the accuracy of participants' responses, I distinguish between two types of non-target responses: failed responses, which is a "don't know" or a "pass" that indicates a failure to retrieve the word; and an erroneous response, which is a response that deviates from the target. The rates of non-target responses are presented in Table 4.2 and Figure 4.2 for each task and for all participants. Throughout experimentation, participants produced a total of 272 failed responses and 23 erroneous responses. A list of failed responses and erroneous responses for each participant can be found in the appendix.

TABLE 4.2. Failed and erroneous response percentages per task

	PN English	*PN Arabic*	*Tr. English*	*Tr. Arabic*
Failed responses (272 total)	6.6%	31.3%	28.3%	33.5%
Erroneous responses (23 total)	0%	17.4%	82.6%	0%

FIGURE 4.2. Failed and erroneous response counts per task



4.2.1 Failed responses

Failing to provide a response for a certain token of the experimental tasks represents an inability to retrieve either the conceptual information for that token (picture naming tasks) or the translation equivalent for that token (translation tasks). As shown in Table 4.2, the task with the highest number of failed responses was the Tr-A task (91, or 33.5% of all failed responses); and

the task with the second highest number of failed responses was the PN-A task (85, or 31.3% of all failed responses). English tasks (PN-E and Tr-E) were relatively consistent between subjects, with very low rates of response failure in the PN-E task (μ=0.6) and a mean of 2.57 failures in the Tr-E task.

In relation to subject proficiency, I predicted high rates of failures for low-proficiency participants, especially for the tasks that require access to the WL lexicon (PN-A and Tr-A). The failed response and erroneous response rates out of the total responses of each group are presented in Table 4.2.1.

TABLE 4.2.1. Percent failed and erroneous responses of all responses per participant group.

	Low Proficiency				Medium Proficiency				High Proficiency			
	PN-E	PN-A	TR-E	TR-A	PN-E	PN-A	TR-E	TR-A	PN-E	PN-A	TR-E	TR-A
Total No. of Responses	44	44	44	44	110	110	110	110	88	88	88	88
% Failed Responses	0%	75%	43%	70%	10%	36%	29%	40%	9.1%	9.1%	19%	17%
% Erroneous Responses	0%	0%	9%	0%	0.9%	1.8%	9%	0%	0%	2.3%	5.7%	0%
% Target Responses	100%	25%	48%	30%	89%	62%	62%	60%	91%	89%	75%	83%

The fact that the failed response rate for the PN-E task was greater than zero for medium and high proficiency speakers is surprising, as I expected a value closer to zero for dominant language picture-naming failures. While producing any errors at all in a dominant-language picture-naming task is unexpected, the 18 errors in the PN-E task were a result of individuals' inabilities to identify or name the line drawings. This could be due to the nature of the line-drawings, the small number of practice items, or other issues with the visual presentation of the experiment. However, the other tasks — the picture naming in Arabic task, the translation into English task, and the translation into Arabic task — appear to each show a correlation with

proficiency (in that higher proficiency participants produce less failed responses than low proficiency participants).

I used a standard two-tailed *T*-test in a mixed effects model to test the interactive effects of proficiency and failed error rate. The failed response count made by subjects with lower proficiency was significantly greater than zero (μ=41.5, *SE*=5.4618). While intermediate subjects made fewer failed responses, the number of their failed responses was also significantly greater than zero (μ=28.2, *SE*=3.454). In addition, the number of failed responses made by high proficiency subjects was significantly greater than zero (μ=12.0, *SE*=3.8621). These results are presented in Table 4.2.2.

TABLE 4.2.2. Mean failed responses by proficiency group using a mixed effects model

Parameter	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Low proficiency speakers	41.500	5.4618	7.60	<.0001
Intermediate prof. speakers	28.200	3.4543	8.16	<.0001
High proficiency speakers	12.000	3.8621	3.11	0.0145

4.2.2 Erroneous responses

Rather than the absent lexico-semantic connections that are indicated by failed responses, the number of erroneous responses indicate misdirected or weak connections to either the conceptual space (picture naming tasks) or to translation equivalents (translation tasks). In looking at erroneous responses, the Tr-E task had a far greater number of erroneous responses with 82.6% of all erroneous responses (19 errors), while the Tr-A task had no erroneous responses. For the picture naming tasks, the PN-E task produced no erroneous responses, while

the PN-A task accounted for only 17.4% of erroneous responses (4 errors), as shown in Table 4.2 above.

While the Tr-E task had a far greater number of erroneous responses, there was no correlation between erroneous responses and participant proficiency. The mean of the erroneous responses of low (μ=2.00) and high proficiency speakers (μ=1.7500) were non-significant compared to zero, as represented in Table 4.2.3. However, the mean of the erroneous responses of intermediate proficiency speakers (μ=2.400) were significantly different from zero.

TABLE 4.2.3. Mean erroneous responses by proficiency group using a mixed effects model

Parameter	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Low proficiency speakers	2.0000	1.4990	1.33	0.2188
Intermediate prof. speakers	2.4000	0.9480	2.53	0.0352
High proficiency speakers	1.7500	1.0599	1.65	0.1373

The result is that the erroneous response rates are generally low and are not significantly different from zero for low and high proficiency groups, but are significantly greater than zero for the intermediate proficiency group. This, however, should take into account that the intermediate proficiency speakers produced a greater number of responses in general, and so the likelihood they would produce errors was high as well. The highest proficiency speakers, as expected, had the lowest mean error rate, indicating more robust connections.

The fact that the majority of errors is occurring in the (Tr-E task, and decidedly not in the Tr-A task) indicates that WL words are activating inaccurate DL translation equivalents. These results are also logical considering the limited WL knowledge of heritage speakers.

4.2.3 Discussion

The result of the failed response measurements shows that lower proficiency participants more frequently fail to respond to the experimental prompts. The high rates of response failure in tasks that require access to the WL lexicon (PN-A and Tr-A) is expected for low proficiency subjects, given these subjects' low proficiency in Arabic. As a result, it is generally expected that proficiency and failed response rate are correlative.

Specifically, a high failed response rate for picture naming in Arabic tasks indicates no conceptual connection between the WL and the conceptual domain, as is represented in the WAM. Conversely, an approximately-equal failure rate in the translation into Arabic task can be attributable to poor proficiency in the WL. The low-proficiency speakers' failed response results are in line with both the reaction time results and the results of previous studies in suggesting a WAM.

Similarly, a significantly smaller PN-A failed response rate for the intermediate proficiency speakers suggests a CMM. The result of the accuracy measurements for the intermediate speakers is also in line with the reaction time results for the intermediate speakers.

By the same rationale, the low failed response rates of the high proficiency speakers indicates a CMM, which counters the results of the reaction time results. This also counters the results of the previous studies, which attribute a conceptual mediation model to high proficiency speakers. The discord between the accuracy results and the reaction time results cite the need for further investigation on high-proficiency subjects.

The aggregated analysis of the erroneous and failed responses per task has other interesting implications. Primarily, the fact that participants produced very few errors in the PN-E task is very likely due to their high fluency and dominance in English. This result suggests that

the connections between DL words and their respective concepts are robust, as is expected for a dominant language.

Further, it is notable that participants produced no erroneous responses during the Tr-A task, but a high failed response rate. This indicates that individuals are more inclined to fail to respond rather than to respond erroneously when responding in the weaker language. In isolation, this could indicate a lack of connections between the semantic store and the weaker language for those specific lexical items, rather than existent but weak connections. Given the reaction times, however, I still argue that the domain has access to/is connected to the conceptual space; just not all words have a connection. Therefore, this result could evidence a lack of confidence in the WL, anxiety about producing the WL incorrectly, etc.

The results of the WL tasks (PN-A and Tr-A) are expected for low proficiency heritage speakers with reduced information in WL lexical entries. The fact that lower WL proficiency correlates with failed responses confirms this, as individuals with lower proficiency also have a more limited WL lexicon.

A general overview of the participants' errors seems to agree with the findings above. As an example, Participant (115) produced responses that appeared to be semantically similar to the target. During the translation into Arabic task, the spoken English word prompt was "wallet," (target: /*dʒɪzdan*/) and she produced the Arabic for the word "notebook" (i.e, /*daftar*/). Notably, this implicates the semanticity of both folding objects moreso than the limited phonological relatedness of the words. During the translation into English task, this same participant (115) responded to the Arabic prompt /*bənaja*/ (target: "building") with the English word "apartment," for which the Arabic is /*ʃaqa*/. Similarly of note, these words are far more semantically related than they are phonologically related. This appears to evidence a conceptual mediation for the

translation task. For this participant, Arabic picture naming appears to take almost the same length of time as translation into Arabic (on average only about .200 ms faster), providing more evidence of conceptual mediation.

As a counterexample, Participant (116) produced erroneous responses that were phonologically — not semantically — similar to the target. During the translation into English task, this participant responded to the Arabic prompt /*barʔanə*/ (target: "orange," as in the fruit) with "curtain," the Arabic for which is /*bərdajə*/. Yet others (110, 114) responded with the word "cold" (Arabic: /*bardanə*/) rather than the target "orange." In the case of participant 116, phonological errors are evidence for a lack of conceptual mediation. While the data of participant of 116 was disqualified on the basis of incompletion, had the participant completed the study, we would predict a word-association model because of the status of his errors.

Crucially, no participant ever made both a semantically-related error and a phonologically related error throughout the course of their experimentation, and largely, these errors appeared to correlate to the subject's proficiency, with higher proficiency subjects making errors based on semanticity.

Interestingly, during the Tr-E task, translation of the Arabic word for *car* into English took the shortest amount of time across all tasks for all participants, at a mean of 384 ms. It is possible that this is due to the three-syllable length of the Arabic prompt causing a response prior to the end of the word, in addition to a high use frequency of the word.

As an informal observation, participants' actions and vocalizations as they attempted to recall the lexical item proved interesting and potentially informative. During the picture naming in Arabic task, some participants gesticulated semantic information with the objects using their

hands, such as fanning the fingers outwards from the hips to recall the Arabic word for “skirt”, or gazing at and rubbing their knuckle for “ring.” This could indicate a subconscious attempt to surface additional semantic information to make the lexical item more salient. However, this did not occur during the translation task.

CHAPTER 5: GENERAL DISCUSSION

The present study found evidence that some heritage language bilinguals show evidence of concept mediation, even without high proficiency. Past studies showed that only proficiency was a factor that led to conceptual mediation. Given the data presented, an additional exploration of the lexical retrieval processes of heritage bilinguals as it relates to CMM or WAM is required.

While the reaction time results of the intermediate proficiency speakers fell in line with predictions made by the concept mediation model, the results of the highest proficiency and lowest proficiency heritage speakers did not.

In isolation, the results of the intermediate group may suggest that age of acquisition has an effect in developing conceptual mediation, even without language proficiency. However, considering the results of the low proficiency speakers, it is critical to consider a model that allows proficiency and AoA to interact (Montrul & Foote, 2012). Given the results, it appears as though despite having acquired Arabic at an early age, losing the language skills leaves the low proficiency speakers dependent mainly on word-to-word connections. It seems that an early age of acquisition can facilitate lexico-semantic links only while having reached a certain level of proficiency. However, without reaching a certain level of proficiency, early AoA seems to have no effect on strengthening the lexico-semantic connections.

The reaction times of the high proficiency group appears to support the word association model, and counters the results of the intermediate proficiency speakers. Further, the reaction time results of the high proficiency speakers counters their low error and failure rates. This could

be a result of the small sample size, or indicate a need for a model that takes language dominance and frequency of use into account. These results validate the need for further investigation on high proficiency bilinguals, as well as a comparison to second language learners of Arabic.

There are several ways to expand upon this study. First, designing a study comparing the response times of low proficiency, low AoA heritage speakers with the response times of low proficiency, late AoA second language learners would also be helpful in determining an effect of the AoA variable without proficiency. Additionally, additional participants would help to stabilize the results and demonstrate a more definitive trend. Data from a larger number of participants would be able to withstand more rigorous statistical analysis, and perhaps indicate more definitive results. In effect, running additional participants would also require a finer gradient of proficiency testing. For example, rather than a grammaticality judgment score of 1-10, a larger number of judgments, or another type of proficiency assessment would be required. This would operationalize language proficiency on a continuous scale. The combination of running more participants and organizing participant data by continuous proficiency measures would allow a trend to be more acutely visible.

Additionally, a wider range of stimuli and a repeated-measures design that re-used stimuli in different tasks would support the validity of the study and show, perhaps, that picture naming was faster than translation even for the same exact concepts. This would reduce the effect that the semantic classification of words, frequency in childhood, phonological properties, and word-level AoA has on recall speed.

To confirm these results, it would be beneficial to conduct a lexical judgment task, a translation judgment task, or a priming experiment with heritage bilinguals. Other studies conducted one of these other tasks to demonstrate the effect of mediation through the conceptual

domain. Conducting such tasks with heritage speakers in a range of proficiencies might illuminate additional details regarding what lexico-semantic connections might have once existed during childhood, and what connections still remain. This also has the possibility of uncovering what mechanisms account for the productive deficiencies of heritage speakers, and how their language skills can be maintained and improved.

Research dealing with heritage speakers is a great benefit to work in bilingual acquisition, owing to their unique trajectories and the resulting linguistic outcomes. What can trigger successful recall of lost or inaccessible lexical items remains to be studied, as well as how heritage speakers cognitively manage their two languages (such as choosing between them in code-mixing situations). With significantly asymmetric language abilities, heritage bilinguals remain to be more extensively studied in future research.

APPENDIX

TABLE 1.1. List of stimuli.

	Picture #	Lexical Item	Arabic text	Practice	Picture Naming		Translation	
					L1	L2	to L1	L2
1	7	arm	يد	1				
2	12	axe	فاس	1				
3	38	brush	فرشاية	1				
4	49	cat	قطة	1				
5	65	comb	مشط	1				
6	69	crown	تاج	1				
7	70	cup	فنجان	1				
8	78	dress	ثوب	1				
9	106	glove	كفة	0				
10	116	hanger	عُلاقة	1				
11	153	necklace	سلسال	0				
12		saucepan	طنجرة	1				
13	2	airplane	طيارة		0	1		
14	6	apple	تفاحة		0	1		
15	14	ball	طابة		1	0		
16	16	banana	موزة		0	1		
17	25	bell	جرس		1	0		
18	40	butterfly	فراشة		0	1		
19	43	camel	جمل		1	0		
20	44	candle	شمعة		0	1		
21	55	chicken	جاجة		1	0		
22	66	corn	ذرة		0	1		
23	68	cow	بقرة		1	0		
24	73	dog	كلب		1	0		

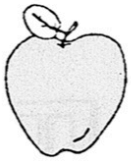
25	75	donkey	حمار	1	0
26	76	door	باب	0	1
27	84	elephant	فيل	1	0
28	88	finger	اصبع	0	1
29	89	fish	سمكة	1	0
30	90	flag	علم	0	1
31	91	flower	وردة	1	0
32	97	fork	شوكة	0	1
33	100	frog	ضفدع	1	0
34	104	glass	كاسة	1	0
35	114	hammer	شاكوش	1	0
36	118	hat	طاقيّة	1	0
37	121	horse	حصان	0	1
38	122	house	بيت	1	0
39	130	knife	سكينة	0	1
40	145	monkey	قرد	0	1
41	149	mouse	فار	0	1
42	167	pen	قلم	1	0
43	185	refrigerator	ثلاجة	1	0
44	187	ring	خاتم	0	1
45	197	scissors	مقص	0	1
46	198	screw	برغي	1	0
47	203	shirt	قميص	1	0
48	205	skirt	تنورة	0	1
49	209	snake	حيّة	1	0
50	212	spider	عنكبوت	1	0
51	215	spoon	معلقق	0	1
52	217	star	نجمة	1	0
53	226	table	طاولة	0	1
54	241	tree	شجرة	0	1

55	245	umbrella	شمسية	0	1		
56	252	watermelon	بطيخة	0	1		
57	5	ant	نملة			0	1
58		baby	طفل			1	0
59		bag	شنطة			0	1
60		bathroom	حمام			1	0
61	22	bed	تخت			0	1
62	28	bird	عصفور			1	0
63		bone	عظمة			0	1
64	30	book	كتاب			0	1
65	32	bottle	قنينة			1	0
66		boy	ولد			1	0
67	36	bread	خبز			0	1
68		building	عمارة			1	0
69	47	car	سيارة			1	0
70		egg	بيضة			0	1
71	86	eye	عين			1	0
72		feather	ريشة			1	0
73		girl	بنت			0	1
74	127	kettle	ابريق			0	1
75	128	key	مفتاح			0	1
76	131	ladder	سِلم			0	1
77		lighter	قداحة			1	0
78	140	lion	اسد			0	1
79		man	زلمة			0	1
80		mirror	مراي			0	1
81	154	needle	إبرة			1	0
82	158	orange	بردقانة			1	0
83	172	pig	خنزير			0	1
84		pigeon	حمامة			1	0

85		plate	صحن					0	1
86		pocket	جيبة					1	0
87		police	شرطي					0	1
88		queen	ملكة					1	0
89	182	rabbit	ارنب					1	0
90		scorpion	عقربة					0	1
91	199	screwdriver	مفك					1	0
92		ship	سفينة					1	0
93		soldier	جندي					1	0
94		tray	صينية					0	1
95		wallet	جزدان					0	1
96		washing machine	غسّالة					1	0
97	250	watch	ساعة					0	1
98	254	wheel	عجل					1	0
99	257	window	شباك					0	1
			Totals per	10	22	22	22		22
			Totals	**10**	44		44		

TABLE 2. List of picture-naming norms used (Snodgrass & Vanderwart, 1980)


002_plane.gif


006_apple.gif

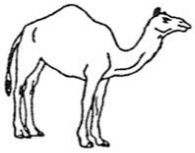

007_arm.gif

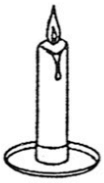

014_ball.gif


016_banana.gif


025_bell.gif

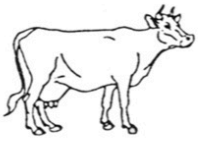

040_butterfly.gif

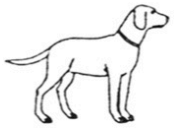

043_camel.gif

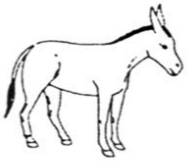

044_candle.gif

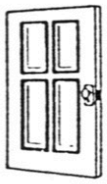

055_chicken.gif

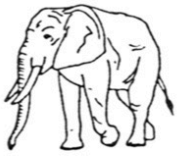

066_corn.gif


068_cow.gif


073_dog.gif


075_donkey.gif


076_door.gif


084_elephant.gif


088_finger.gif

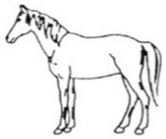

089_fish.gif


090_flag.gif


104_glass*.gif


114_hammer.gif


118_hat.gif


121_horse.gif

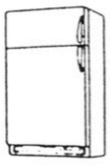

122_house.gif


130_knife.gif


145_monkey.gif


149_mouse.gif


167_pen.gif


185_refridgerator.gif


187_ring.gif

197_scissors.gif

198_screw.gif

203_shirt.gif

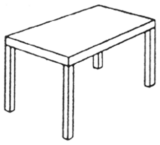
205_skirt.gif

209_snake.gif

212_spider.gif

215_spoon.gif

217-star.gif



226_table.gif

241_tree.gif

245_umbrella.gif

252_watermelon.gif

TABLE 3. Grammaticality judgment task.

Grammar Construction	Ex	Ungrammatical	Grammatical	Translation
Verb agreement	A	الولد درسوا اللغة العربية *al-walad dɛrɛsu al-loɣa al-ʕrabijə*	الاولاد درسوا اللغة العربية *al-awla:d dɛrɛsu al-loɣa al-ʕrabijə*	The boys (*they/he) studied the Arabic language.
	B	إشترت الرجّل سيارة حمرة و حبها. *iftarat arrajol sijara ħamra wa ħabha.*	إشترى الرجّل سيارة حمرة وحبها. *iftara arrajjol sijara ħamra wa ħabha.*	The man$_i$ bought a red car and (*she$_i$/he$_i$) loved it.
Gender Agreement	A	المرأة الطويل إشترت خضرة من الدكان *al-imraʔ alTawil iftarat xudra min ad-dukan.*	المرأ الطويلة إشطرت خضرة من الدكان. *al-mara alTawilɛ iftarat xudra min ad-dukan.*	The tall (*m/f) woman bought vegetables from the store.
	B	خاف الولد من البنت، فا هرب منه. *xaf alwalad men albint fa harab meno.*	خاف الولد من البنت فا هرب منها. *xaf alwalad men albint fa harab minha.*	The boy was afraid of the girl, so he ran away from (*him/her).
Formation of Plurals	A	انا عايز قلمات للمدرسة. *ana ʕajiz qalamat lal madrasa*	انا عايز اقلام للمدرسة. *ana ʕajiz aqla:m lal madrasa*	I need (*pens/pens) for school.
	B	هي اكلت بِكل المطعمات. *hije akalat bi kol almatʕamat.*	هي اكلت بِكل المطاعِم. *hije akalat bi kol almataʕim.*	She ate in all the (*restaurants/restuarants)
Preposition	A	بدوّر للكتاب *bidawwir lil kitab.*	بيدوّر على الكتاب. *bidawwir ʕla al-kitab.*	He is looking (*for/on) the book.
	B	أبي بيشتغل مع السيارات. *abbi bi-jiftaɣl maʕ al-sijara:t*	أبي بيشتغل بِالسيارات *abbi bi-jiftaɣl fi al-sijara:t*	My father works (*with/in) cars
Nominal Construction	A	الدفتر الولد ما كان بِبيته. *al-daftar alwalad ma kan bi beto.*	دفتر الولد ما كان بِبيته. *daftar alwalad ma kan bi beto.*	(*The/∅) notebook of the boy was not at home.
	B	الولد دشّر اللعبة بِالبيت البنت *Al walad tərək al loʕbɜ bi albet albinit.*	الولد ترك اللعبة بِبيت البنت. *Al walad tərək al loʕbɜ bi bet al binit.*	The boy left the toy at (*the/∅) house of the girl.

REFERENCES

Abrahamsson, N., and Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning* 59(2): 249-306.

Albirini, A. (2013). Toward understanding the variability in the language proficiencies of Arabic heritage speakers. *International Journal of Bilingualism*: 1367006912472404.

Albirini, A., and Benmamoun, E. (2015). Factors affecting the retention of sentential negation in heritage Egyptian Arabic. *Bilingualism: Language and cognition* 18(03): 470-489.

Andersen, C. (2004). Literacy and language revitalization in Labrador Inuttitut. Unpublished ms. Nain: Torngâsok Cultural Centre.

Bhatia, T. K. and Ritchie, W. C. (2006) Bilingualism in the Global Media and Advertising, in *The Handbook of Bilingualism* (eds T. K. Bhatia and W. C. Ritchie), Blackwell Publishing Ltd, Oxford, UK.

Bosch, L., & Sebastían-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy,* 2, 29-49.

Bosch, L., and Sebastían-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language- specific vowel contrast in the first year of life. *Language and Speech, 46*, 217-243.

Bullock, B. E., & Toribio, A. J. (2006). Intra-system variability and change in nominal and verbal morphology. In R. S. Gess & D. Arteaga (eds.), *Historical Romance linguistics: Retrospective and perspectives*, pp. 305–325. Amsterdam: John Benjamins.

Chen, H., and Leung, Y. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15.2: 316.

Chen, H., and Ng, M. (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals." *Memory & Cognition* 17.4: 454-462.

De Houwer, A. (1998). By way of introduction: Methods in studies of bilingual first language acquisition. *The International Journal of Bilingualism* 2(3).

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. Applied Psycholinguistics, 28(3), 413.

Deuchar, M., & Quay, S. (2000). Bilingual acquisition: Theoretical implications of a case study. New York: Oxford University Press.

Edwards, J. (2006). Foundations of bilingualism. In T. Bhatia & W. Ritchie (Eds.), *The handbook of bilingualism* (pp. 5-29). Oxford, UK: Blackwell.

Jiang, N. (2000) Lexical representation and development in a second language. *Applied linguistics* 21.1: 47-77.

Jiang, N. and Forster, K. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language* 44.1: 32-51.

Khwaileh, T., Body, R., & Herbert, R. (2014). A normative database and determinants of lexical retrieval for 186 Arabic nouns: Effects of psycholinguistic and morphosyntactic variables on naming latency. *Journal of psycholinguistic research* 43.6: 749-769.

Kroll, J. F., and Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. *Practical aspects of memory* 2.389-395.

Kroll, J. F., and Dijkstra, A. (2002). The bilingual lexicon. In R. Kaplan (Ed.), *Handbook of applied linguistics* (pp. 301–321). Oxford, UK: Oxford University Press.

Kroll, J. F., and De Groot, M. B. (2005). *Handbook of bilingualism*. Oxford, UK: Oxford University Press.

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of memory and language* 33(2):149-174.

Maher, J. (1991). A crosslinguistic study of language contact and language attrition. In H. W. Seliger & R. M. Vago (eds.), *First language attrition*, pp. 67–84. Cambridge: Cambridge University Press.

Montrul, S, and Foote, R. (2012). Age of acquisition interactions in bilingual lexical access: A study of the weaker language of L2 learners and heritage speakers. *International Journal of Bilingualism* 18(3):274-303.

Pavlenko, A. and Malt, B. (2011). Kitchen Russian: Cross-linguistic differences in first-language object naming by Russian-English bilinguals. *Bilingualism: Language & Cognition* 14:19-45.

Pierce, J.W. (2007). Psychophysics software in Python. *Journal of Neuroscience Methods.*162(1-2):8-13.

Polinsky, M. (1994). Structural dimensions of first language loss. *Papers from the 30th Regional Meeting of the Chicago Lingusitic Society Vol. 2.* 257-77

Polinsky, M. (1997). Cross-linguistic parallels in first language loss. *Southwest Journal of Linguistics*, 14, 87–123.

Polinsky, M. (2005). Word class distinctions in an incomplete grammar. *In D. Ravid & H. Bat-Zeev Shyldkrot (Eds.), Perspectives on language and language development* (pp. 419–436). Dordrecht, the Netherlands: Kluwer.

Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics, 14*, 191–262.

Polinsky, M. (2008). "Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization." *Heritage Language Journal* 6.1: 40-71.

Polinsky, M. (2008b). Heritage language narratives. In D. Brinton, O. Kagan, & S. Bauckus (Eds.), Heritage language education: A new field emerging (pp. 149–164). London: Routledge.

Polinsky, M., and Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass* 1.5: 368-395.

Potter, M. C., So, K. F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of verbal learning and verbal behavior*, *23*(1), 23-38.

Sholl, A., Sankaranarayanan, A. & Kroll, J. F. (1995). Transfer between picture naming and translation: A test of asymmetries in bilingual memory. *Psychological Science* 6(1), 45-49.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13.2 (2009): 155-163.

Sherkina-Lieber, M., Pérez-Leroux, A., & Johns, A. (2011). Grammar without speech production. *Bilingualism: Language & Cognition 14(3),* 301-317.

Silverberg, S., & Samuel, A. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51*, 381–398.

Snodgrass, J. G., and Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory* 6.2: 174.

Valdés, G. (2000). Introduction. In Spanish for native speakers: AATSP professional development series *handbook for teachers K-16* (Vol. I). New York, NY: Harcourt College.

Valdés, G. (2001). Learning and not learning English: Latino students in American schools. New York: teachers College Press

Volterra, V., and Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language* 5.