5-2012

Is it really all downhill after puberty?: The Critical Period Hypothesis in Second Language Acquisition - A review of the literature

Katherine Lynn Nelson
knelso13@utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Katherine Lynn Nelson entitled "Is it really all downhill after puberty?: The Critical Period Hypothesis in Second Language Acquisition - A review of the literature." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in French.

Sébastien Dubreil, Major Professor

We have read this thesis and recommend its acceptance:

Harriet W. Bowden, John R. Romeiser

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Is it really all downhill after puberty?: The Critical Period Hypothesis in Second Language Acquisition – A review of the literature

A Thesis Presented for the
Master of Arts
Degree
The University of Tennessee, Knoxville

Katherine Lynn Nelson
May 2012

ACKNOWLEDGEMENTS

This thesis would not have been possible without the help of my co-chairs and my committee. I would like to thank Dr. Dubreil for your countless support and laughs, and thank you for keeping me sane throughout this entire process. I would like to thank Dr. Bowden for her plethora of knowledge in this subject and for all of the reviewing that was needed. I would like to thank Dr. Romeiser for your continuing support despite my scattered brain in and out of class. I would also like to thank Dr. Tracy Parkinson for his enthusiasm for French and French literature and for his belief in me, as well as his support of my desire to go to graduate school.

I would like to thank my family for loving me despite the lack of attention they have received during the writing of this thesis, and David for your support, love, and encouragement of me and for me during this whole process. You are wonderful!

Last, but not least, thank you Rachel Lamance for going through this whole process with me, through the classes, the making of lesson plans, the enormous amount of reading, the even larger amount of complaining, and the endless studying and writing. I would not have wanted to do it without you.

ABSTRACT

The Critical Period Hypothesis in Second Language Acquisition posits that there is a critical period, early childhood until puberty, in which human beings must acquire a second language if they are going to achieve native-like attainment in that language. This thesis is a review of the current state of research in regards to the Critical Period Hypothesis in Second Language Acquisition. While evidence is provided to refute the Critical Period Hypothesis in the studies examined, a general age effect is found in the native-like attainment of a second language both in the acquisition of grammatical features and in the acquisition of phonological system. A look at how to relate these findings to foreign language education in the United States is begun by looking specifically at foreign language immersion programs in the United States and their effectiveness on native-like attainment and proficiency. There are currently 448 foreign language immersion programs in the U.S. with 45% of these programs being Spanish immersion and 22% being French immersion. Research shows that foreign language immersion programs are effective in helping children acquire a second language more effectively, specifically the early immersion programs, as well as acquire more metalinguistic awareness of languages than their monolingual counterparts.

TABLE OF CONTENTS

Chapter Page

CHAPTER I....1
CHAPTER II....8
Statement of purpose....8
Research Questions....8
Significance....8
Delimitations....9
Methodology....9
CHAPTER III....10
Historical Overview of the Critical Period Hypothesis....10
Lenneberg 1967....10
Penfield and Roberts 1959....13
Johnson and Newport 1989....15
Singleton 1989....20
Summary....22
CHAPTER IV....25
The current state of research of the cph in sla....25
Native-like attainment of grammatical features....26
Pronunciation as native-like attainment....36
Summary....46
CHAPTER V....48
Overview of foreign language immersion programs....48
History of Language Immersion Programs....49
What are foreign language immersion programs?....50
CHAPTER VI....60
Conclusions....60
LIST OF REFERENCES....63

LIST OF FIGURES

Figure Page

Figure 1. Stretched Z Function 4
Figure 2.a Stretched 7 Function…….....5
Figure 2.b Stretched 7 Function…….....5
Figure 3 Monotonic Linear Decline……6
Figure 4 Negative Correlation of the Results in JN89……..17
Figure 5 Subjects Arriving Ages 17-39, r=-.16……..19
Figure 6 Seol (2005) Negative Correlation…….30
Figure 7 Percentage of Immersion Programs by language of instruction ...…….56

CHAPTER I
INTRODUCTION

"Imprinting in chicks, acquisition of birdsong, cocoon preference in ants, aggression in mice, vision in cats, sociability in dogs, sexual imprinting in finches, maternal responsiveness in goats, egg recognition in birds and social behavior in monkeys" are all different types of knowledge and skills that must be learned by a certain point in the animal's lives. (Bialystok, 1997, p. 117) This means that after a certain age, these skills can no longer be acquired. Researchers have been conducting studies to see if these same sort of maturational constraints exist in different areas for human beings. This is called the Critical Period Hypothesis (CPH). One area in particular is that of language acquisition. Research has been conducted and is being conducted to look for evidence for or against a critical period in language acquisition in regards to both first language acquisition and second language acquisition (SLA).

More specifically, the CPH is a hypothesis that states that there is a critical period in which certain skills must be attained, acquired, or learned. Birdsong defines a critical period as:

> the temporal span during which an organism displays a heightened sensitivity to certain environmental stimuli, the presence of which is required to trigger a developmental event. Typically, there is an abrupt onset or increase of sensitivity, a plateau of peak sensitivity, followed by a gradual offset or decline, with subsequent flattening of the degree of sensitivity (2005, p. 111)

With regards to language the CPH states that human beings must be exposed to a language during infancy and early childhood, prior to puberty. If this does not happen, then the CPH suggests that one would not learn his/her native language fully. (Lightbrown & Spada, 2006, p. 17) In terms of a second language, it would state that anyone who begins to learn a second language after the critical period has ended should not be able to become native-like in that language. (Reichle, 2010, p. 58)

Right after the idea of the CPH was first proposed, research began to test this putative critical period. This research has generally examined second language learners that have immigrated to a country that primarily uses the target language. Testing the native-like proficiency or attainment of such learners is ideal, as they have had long amounts of exposure to the language. These second language learners normally live in a country where their second language is not only the official language but also the most used language, thus they should have had ample amounts of exposure to the language as well as a long length of exposure, and thus be at asymptotic performance in the language. If it is possible to attain native-like proficiency in a second language, these are the learners who would be the closest to achieving it. In addition, the larger the group of participants, the more valid and reliable the results, and the more accurate the conclusions, will be.

Within the area of the CPH, there are different versions of the CPH. There is the stronger version of the CPH, which claims that even if first language

acquisition begins in childhood before the end of the critical period, it will not continue into adulthood. (Singleton & Ryan, 2004, p. 33) The stronger version of the CPH is similar to the "maturational state hypothesis" that Johnson and Newport (1989) proposed in their seminal study that will be discussed in chapter one. This "maturational state hypothesis" claims that "early in life, humans have a superior capacity for acquiring languages" (Johnson & Newport, 1989, p. 64). Furthermore, the hypothesis that Johnson and Newport (1989) propose claims that this capacity for language learning disappears or declines after the end of the critical period. (Johnson & Newport, 1989, p. 64)

The weaker version of the CPH, as proposed by Singleton (2004), simply claims that language acquisition must begin before the end of the critical period for the capacity for learning languages to be able to continue after the end of the critical period. (Singleton, 2004) "The exercise hypothesis," as proposed by Johnson and Newport (1989) correlates to this weaker version of the CPH. The exercise hypothesis claims that if the language learning capacity that is so strong in childhood is exercised, "further language learning abilities will remain intact throughout life" (Johnson & Newport, 1989, p. 64). Both the stronger version and the weaker version of the CPH suggest a superior capacity for language learning in children.

There are two main geometric shapes that can be produced when looking at the correlations in the results. The geometric shape that provides evidence for a critical period is the stretched Z form. "Regarding the hypothesized temporal features," Birdsong states that, "the period of maximal sensitivity to linguistic

input, with full attainment of grammatical competence assured, extends through early childhood" (Birdsong, 2005, p. 112). The offset of the critical period begins in early childhood and ends at the point "at which full neurocognitive maturation is reached" (Birdsong, 2005, p. 112). There is then a flattening of the line, which indicates the sensitivity to language learning at its lowest level. The end of the critical period is considered to be the point where full neurocognitive maturation is reached.

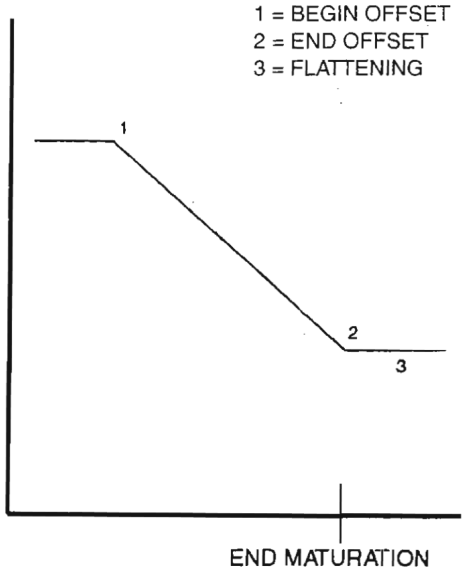

Fig. 1 shows the streteched Z function as described by Birdsong (2005).[1]

In regards to the results and correlations found between age of arrival (AOA) and ultimate performance, this stretched Z indicates that those learners who begin language acquisition very early in childhood, well before the end of the critical period, will reach asymptote, or end-state. There is then a sharp decline in the native-like attainment of those learners who begin in mid to late childhood with

[1] From "Interpreting Age Effects in Second Language Acquisition," by D. Birdsong, 2005. In J.F. Knoll & A.M.B. de Groot (Eds.) *Handbook of Bilingualism: Psycholinguistic Approaches,* 109-127. New York: Oxford University Press.

those who begin after the close of the maturational period only able to attain so much of a language. (Birdsong, 2005, p. 112-3)

Birdsong also discusses three other geometric shapes that are found in the literature, two of which can be considered a stretched 7 geometric shape with different points at which offset begins. The third is a linear decline. Figure 2A represents a stretched 7 where the point at which offset begins is actually the point where neurocognitive maturation is fully reached, and there is no end point for offset. (Birdsong, 2005, p. 113)

Figure 2. The stretched 7 function as described by Birdsong[2]

Figure 2B also represents a stretched 7 shape but in this stretched 7 shape, the point at which offset begins happens prematurationally, earlier in childhood, which is like the stretched Z function. This stretched 7 (Fig. 2), however, does not have an end to offset. Thus, there is no period in which language acquisition can be confined; the stretched 7 function, then, does not correspond to a critical period but to maturational age effects. Figure 3 then represents the linear decline

[2] From "Interpreting Age Effects in Second Language Acquisition," by D. Birdsong, 2005. In J.F. Knoll & A.M.B. de Groot (Eds.) *Handbook of Bilingualism: Psycholinguistic Approaches,* 109-127. New York: Oxford University Press.

that is also sometimes found. This decline indicates that the sensitivity to language learning is at its highest level closer to birth, but there is no point of discontinuity, which means that there is no offset of sensitivity. Both the stretched Z and the stretched 7 functions have points of discontinuity that represent the beginning of offset.

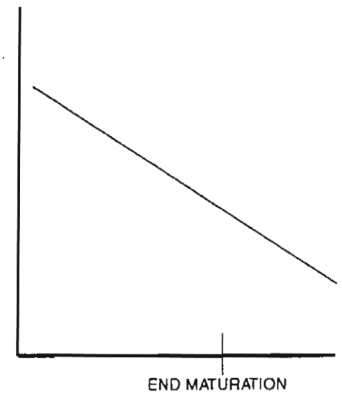

Figure 3 Linear monotonic decline[3]

This monotonic decline indicates that language learning declines as AOA increases, but that it does not stop at the end of maturation. This monotonic linear decline indicates that there are general age effects in language acquisition but provides evidence against a critical period for language learning. (Birdsong, 2005)

The age at which maturation is reached is as much an empirical issue as whether or not there is a critical period. Generally, the upper limit of the critical period is considered to be the early teens, the point where childhood ends and

[3] From "Interpreting Age Effects in Second Language Acquisition," by J.S. D. Birdsong, 2005. In J.F. Knoll & A.M.B. de Groot (Eds.) *Handbook of Bilingualism: Psycholinguistic Approaches,* 109-127. New York: Oxford University Press

adolescence begins with the onset of puberty, roughly ages 12 -14. (Singleton, 2004, p. 40)

Thus, this paper will examine the current state of research on the CPH in SLA, as well as look at a possible response to the implications of the findings in these studies.

CHAPTER II
STATEMENT OF PURPOSE

This paper examines the Critical Period Hypothesis in second language acquisition. Studies assessing learners' nativelikeness in grammatical features as well as phonology are examined for evidence in support of or against a critical period. This paper also glances at foreign language immersion programs in bilingual education in response to the implications of the results of the studies.

Research Questions

The research questions that this paper will examine are the following:

1. What is the current state of research in regards to the Critical Period Hypothesis in Second Language Acquisition?
2. What are foreign language immersion programs? Are they successful and what is the current condition of these programs in the United States?

Significance

A great deal of research has been conducted to test the hypothesis of a critical period in which one must learn a second language before the onset of puberty in order to achieve native-like proficiency. If this putative critical period exists, then the way in which the educational systems are introducing foreign language education needs to be examined. Should bilingual education be the order of the day? This thesis will entail a critical literature review of studies that have been conducted over the past decade, as well as look at foreign language immersion

programs, which could be a solution to the implications of the putative critical period for SLA in the context of foreign language education in the United States.

Methodology

This thesis examines studies that considered the CPH in SLA. This review includes studies found in searches of the “CPH in SLA”, “Maturational constraints in SLA”, and “age effects in SLA” in LLBA, Worldcat, Google Scholar, and JStor with publication dates of 2000 or later. Only the more recent studies were examined as the older studies have all been reviewed and newer studies have been conducted to make up for the limitations of those previous studies. This thesis also includes studies and information on foreign language immersion programs as a means of bilingual education that was found by searching bilingualism, foreign language immersion programs in the U.S., and bilingual education-immersion programs in LLBA, Google Scholar, and JStor.

Delimitations

This paper does not discuss studies older than 2000 unless they are considered seminal studies because studies older than 2000 have been more thoroughly evaluated in the literature, and new research has been conducted to test the results and the limitations of the previous studies. This paper also discusses information on foreign language immersion programs in the area of bilingual

education. There are different types of bilingual education, but foreign language immersion schools are the programs that might be likely proposed in the United States to solve the problem of lack of success in foreign language education attributable to the CPH in SLA.

CHAPTER III
HISTORICAL OVERVIEW OF THE CRITICAL PERIOD HYPOTHESIS

Lenneberg 1967

Lenneberg (1967) is one of the first linguists to label the critical period in regards to first language acquisition. In his book *Biological Foundations of Language* (1967), Lenneberg discusses the reasons why children learn language when they do. He states the fact that mothers do not have a specific time and schedule that they abide by to start language training with their children. Instead he hypothesizes that first languages are acquired through "maturational processes within the individual" (Lenneberg, 1967, p. 125-6). By "maturational processes," whether in behavior or language, Lenneberg is referring to the changes, or processes, that are happening or developing because of changes within the individual, not because of the world around them. (Lenneberg, 1967, p. 125) Lenneberg proposes to use four characteristics of "maturationally controlled emergence of behavior" to discuss the control of maturational processes on the emergence of speech and language. (Lenneberg, 1967, p. 127)

These four characteristics form the basis of Lenneberg's argument. The first, the regularity of onset, is based on the onset, or the "gradual unfolding of capacities" in speech development (Lenneberg, 1967, p. 127). Lenneberg claims that while there are ages at which most children have acquired certain functions of their first language, individual differences between each child must be

considered. Despite these individual differences, normal acquisition of certain language functions will happen between the second and third year of life. (Lenneberg, 1967, p. 127) The second characteristic is the "relation of the environment to the age of onset" (Lenneberg, 1967, p. 135). Through different studies, Lenneberg is able to summarize that even though environments differ amongst children and even change during a child's life, the age of onset of certain speech and language capabilities remains relatively unaffected (Lenneberg, 1967, p. 139).

"The role of utility in the onset of speech" is the third characteristic that Lenneberg discusses in relation to the maturational processes in the emergence of speech and language. Lenneberg argues that children do not begin to acquire language "as a response to an experienced need, as a result of discovery of its practical utility, or as a product of purposive striving toward facilitated verbal communication" (Lenneberg, 1967, p. 139). Lenneberg conducts a study recording the interactions of deaf children born to deaf parents. All of these children vocalize often during their playtime. They also get along very well without the need to communicate verbally. From this study, Lenneberg questions why hearing children bother to learn a language system when they can get along well without it? He argues that it is because it comes naturally. It is not something that they strive for.

The fourth and final characteristic that Lenneberg discusses is "the importance of practice for the onset of speech" (Lenneberg, 1967, p. 140). The cooing and babbling produced by babies does not represent practice stages for

the acquisition of language. Lenneberg references mute children who are learning and are responsive to language, but who just choose not to respond. Either spontaneously or in response to treatment, these mute children will snap out of it and begin talking as fluently as other normal children do at that age level, which suggests that these children have undergone years of training and learning, just without years of practice, and only choose to respond when they feel ready. (Lenneberg, 1967, p. 141) Through these four characteristics, Lenneberg concludes that language acquisition is "primarily dependent upon the maturational development of states of readiness within the child" (Lenneberg, 1967, p. 142).

Lenneberg goes on to discuss the age limitations in first language acquisition by looking at language disorders. Trying to conduct a study where a child is withheld from the natural input that he receives in order to learn a first language in order to observe him cannot be approved. Thus, Lenneberg examines both children and adults with language disorders, i.e. they have lost their ability to communicate and must re-learn the language. Lenneberg cites different studies that examined children and adults who suffered from aphasia (1967, p. 142), as well as their recovery. The prognosis of the recovery of these patients depended greatly upon the age at which the injury occurred. The earlier the age at which the insult to the brain occurred, the more fully the patients recovered. (Lenneberg, 1967) He also looked at the language acquisition of the mentally retarded. (1967) A study done by Lenneberg, Nichols, and Rosenberger (1964) using 54 patients with Down syndrome provides evidence

for these language limitations in relation to age. This study observed that even in the absence of brain lesions, “progress in language learning comes to a standstill after maturity” (Lenneberg, 1967, p. 155). Through the writings of Lenneberg, a putative critical period for first language acquisition can at least be considered, but where did the CPH in SLA derive from?

Penfield and Roberts 1959

In *Speech and Brain Mechanisms*, Penfield and Roberts (1959) write a chapter titled “The Learning of Languages.” Penfield and Roberts were two of the first researchers to suggest that younger children learn languages more easily and efficiently than children in the second or third decade of their lives. In a talk that he gave at Lower Canada College in 1939, Penfield said to the students, “Remember that for the purposes of learning languages, the human brain becomes progressively stiff and rigid after the age of nine” (Penfield & Roberts, 1959, p. 236). Penfield and Roberts also discuss the different methods of learning and teaching second languages. They state that it is much better to learn a second language earlier on in life through the direct method (or the mother’s method) than it is through the “school-time learning of secondary languages in the second decade of life” (Penfield & Roberts, 1959, p. 240). The direct method refers to the natural learning process during first language acquisition that depends on the evolution of the child’s brain (Penfield & Roberts, 1959, p. 239). This process of learning a first language is helped by the mother but is considered inevitable. Penfield and Roberts argue that there are two types

of reasons why the direct method is better than the school-time learning is true: Physiology and Psychology.

Under neurophysiology, Penfield and Roberts state that the reason for success in the direct method is that “a child’s brain has a specialized capacity for learning language – a capacity that decreases with the passage of years” (Penfield & Roberts, 1959, p. 240). This claim is based on the observation that in immigrant families who arrive in a new country, having no prior knowledge of the language of that country, the younger children generally pick up the language within about two years just by being in school or playing with other children; the parents, however, in most cases must take language classes in order to learn the language. It takes the parents longer, and they must be more intentional about their learning. Under the psychological category, Penfield and Roberts discuss the fact that children learn language out of their curiosity about the world. Language is a method through which they learn about their surroundings, their environment, their world.

Penfield and Roberts do not deny that older learners can in fact learn second languages or that they can learn through the direct method of language teaching. Their primary argument is that it is generally easier and quicker for younger learners. For young adults who begin to learn a second language, they begin to learn it through their mother tongue, or through an indirect method. Penfield and Roberts’ remarks regarding the indirect language learning method is not that it is necessarily ineffective, but that there should be an introduction to the second language using the direct method (Penfield & Roberts, 1959, p. 252).

Penfield and Roberts, thus, establish the idea that there could be maturational effects on SLA.

Johnson and Newport 1989

One of the most cited and replicated studies on the CPH in SLA is a study that was conducted by Johnson and Newport (1989). The purpose of this study was to answer the question of whether there is an age-related effect on the acquisition of grammar of a second language, to study the nature of this relationship if there is one, to look at variables that could explain the "effects obtained for age of learning," and to determine what are the most and least problematic areas of grammar for learners of different age groups (Johnson & Newport, 1989, p. 67-8).

Johnson and Newport (1989) used 46-second language learners of English who were native Chinese or Korean speakers. They chose these language backgrounds because of the typological dissimilarity to English. The minimum criteria for all subjects was to have had at least five years of exposure to English and to have lived in the United States for an uninterrupted stay of at least three years prior to the test. Length of residence (LOR) must be accounted for to assure that the second language learners are at asymptote or end-state; this does not mean that they are native-like but that this point is the outcome of acquisition. (Birdsong, 2005, p. 110) Ten years is generally considered to be the LOR at which most participants would have reached end-state. One reason why Johnson and Newport did not find any native-like late learners could be because

these learners had not yet reached asymptote as some of them may have only had 5 years of exposure to the language.

For ensuring the homogeneity of this study, all subjects were chosen from the faculty and student population of the University of Illinois. There was variation in the age of arrival (AOA) to the United States among the 46 subjects; the range of AOAs was 3 to 39. Using their AOAs, these subjects were divided into two groups. The Early Arrivals consisted of 23 subjects with an AOA of younger than 15, and the Late Arrivals group consisted of 23 subjects as well, with an AOA in the United States after age 17. In regards to length of residence in the United States, the average for early and late arrivals is 9.8 and 9.9 years, respectively. Thus, length of residence was matched between the two groups and does not play a significant role in the outcomes of this study. However, the late arrivals group does have a larger range of years in the United States.

The measure that Johnson and Newport (1989) used was a grammaticality judgment test (GJT). The subjects had to judge the grammaticality of 276 spoken English sentences, 140 ungrammatical sentences and 136 sentences were the grammatical counterparts. These sentences covered 12 types of English rules: past tense, plural, third person singular, present progressive, determiners, pronominalization, partical movement, subcategorization, auxiliaries, yes/no questions, wh-questions, and word order. These rules dealt with two different main categories of rules of English, English morphology and English syntax. A native-American female voice was used to record the test sentences. The subjects listened to each sentence twice and

then had to indicate whether that sentence was grammatical or ungrammatical by circling Y (yes) or N (no) on an answer sheet.

In regard to their first question, Johnson and Newport (1989) found that there was a strong relationship between age of arrival in the United States and performance on the grammaticality judgment test. The negative correlation was $r = -.77$, $p < .01$, (Shown in Fig. 4).

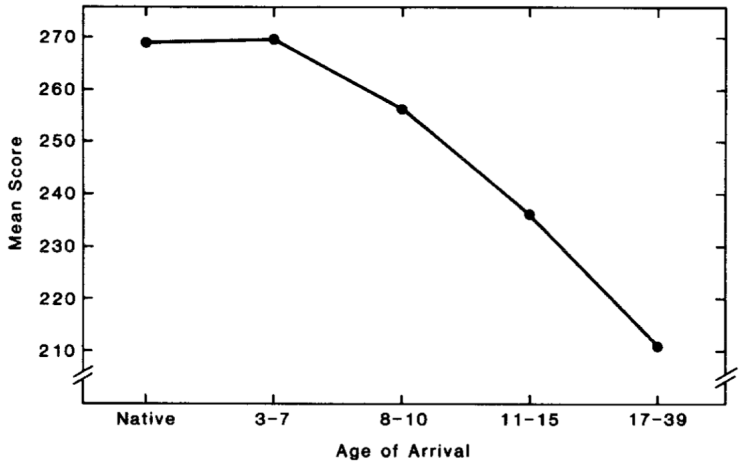

Fig. 4 The relationship between age of arrival in the United States and total score correct on the test of English grammar.[4]

Within the results, Johnson and Newport found that there was no significant difference between the age 3-7 group and the native control group, while all other age groups performed significantly below the native control group. Johnson and Newport conclude from this first finding that if second language immersion happens before the age of 7, native-like attainment of a second language can be achieved. However, if second language immersion happens later, even if soon after the age of 7, there is a significant difference in the level of native-like

[4] From "Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language," by J.S. Johnson and E.L. Newport, 1989, *Cognitive Psychology,* 21, p. 79.

attainment that can be achieved. There was also a significant difference between all of the late groups. Higher scores were obtained by the 8-10 age group than the age 11-15 group, as well as by the age 11-15 group than the age 17-39 group. Thus, there is evidence for a "strong linear relationship between age of exposure to a language and ultimate performance in that language" (Johnson & Newport, 1989, p. 78). One significant finding of this study was that no incidence of nativelikeness among the late learners was found, thus providing evidence against the CPH.

Next Johnson and Newport (1989) divided the subjects into two age groups of age of exposure, as discussed above, ages 3-15 and ages 17-39, to look at the correlations within these two groups between age of exposure and ultimate performance. According to the strong version of the CPH, the older group, ages 17-39, should not produce a significant negative correlation. The results of the younger group, ages 3-15, should be consistent with a stretched 7 function. Overall, the results should be indicative of a stretched Z function. If these functions are found, this would be indicative of a critical period in which two points of discontinuity are found: one during early childhood and one at the end of maturation. In the age 3-15 group, a strong negative correlation was found, $r = -.87$, ($p < .01$), and in the age 17-39 group, there was no significant correlation found, $r = -. 16$, ($p > .05$). Thus providing more evidence for a critical period, because there should be a significant decline in ultimate performance up to and at puberty. However, there should be no significant decline in performance after

puberty, which is what Johnson and Newport found evidence for with these correlations amongst the two age groups.

Johnson and Newport (1989) also found a large variance (See fig. 5) between age of acquisition and ultimate performance among the late learners that created a megaphone shape, whereas among those who were exposed to English at an early age, the variance was very small.

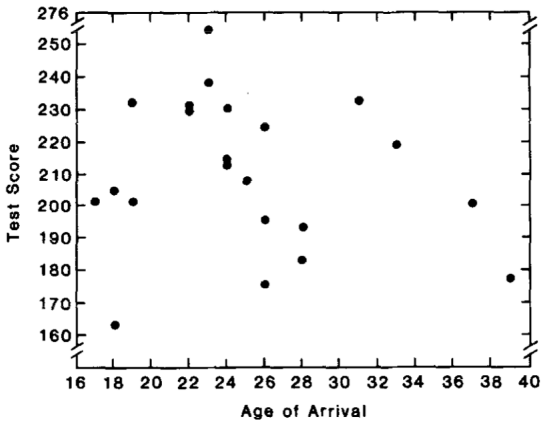

Fig. 5 Scatterplot of test score in relation to age of arrival for subjects arriving in United States after puberty.[5]

This heterogeneity in the variance draws attention to two different points. The first is that for those exposed to a second language before the age of 15, there are very few individual differences in language learning ability. The second is that for adults, one will not become native-like or near native in a second language with a later age of acquisition (Johnson & Newport, 1989, p. 81). Johnson and Newport (1989) also examined whether the initial AOE, initial AOE referring to the age at which some of the learners were first exposed to English

[5] From "Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language," by J.S. Johnson and E.L. Newport, 1989, *Cognitive Psychology,* 21, p. 80.

through formal instruction, would yield similar correlations to AOA. The correlation was -.67, which was not significantly different ($t(43) = 1.26$, $p > .05$) than the negative correlation found between AOA and test score.

Thus, Johnson and Newort (1989) conclude that "there is a gradual decline in language learning skills over the period of on-going maturational growth and a stabilization of language learning skills at a low but variable level of performance at the final mature state" (Johnson and Newport, 1989, p. 97).

Singleton 1989

Singleton has also extensively examined evidence for the CPH. In his book *Language Acquisition: The Age Factor* (1989), Singleton delimits four different positions of the CPH in SLA. The first position is the 'younger = better' position. Singleton cites several different studies that span from the formal education environment to the experience of immigrants to studies that concentrate on both formal education of a second language and the acquisition of immigrants of a second language. All of the studies that Singleton discusses under the 'younger = better' category support this position. These studies do not show that older learners cannot acquire a second language, but just that younger learners have an easier, faster, or more effective learning curve. The study that Singleton labels as the most pertinent one is one done by Yamada *et al*. (1980). Yamada *et al.* used 30 Japanese elementary school children.[6] None of these children had had any previous knowledge of English. This experiment looked at

[6] All of these subjects were of average scholastic achievement. There were 10 first graders, 7 years-old; 10 third graders, 9 years old; and 10 fifth graders, 11 years old.

the subjects' abilities to learn a small selection of English words. In the individual tests, Yamada *et al.* found that the 'mean learning scores decrease with age, i.e. the older the age the lower the score' (qtd. in Singleton, 1989, p. 83).

The second position that Singleton discusses is the opposite of the first. It is the 'older = better' position (Singleton, 1989, p. 94). This position comes from the hypothesis that older learners of a second language are more successful in acquiring the language than younger learners. Most of the evidence for this position comes from studies that are done in the formal education environment.[7] One of the most known short-duration studies is one done by Asher and Price (1967). They used 96 students from the second, fourth, and eighth grades from Blackford School and 37 undergraduate students from San José State College. (Singleton, 1967, p. 95) In this study each older group outperformed their younger counterparts. The adults outperformed all of the adolescents and children. The eighth graders outperformed the fourth graders, and the fourth graders outperformed the second graders. (Singleton, 1967, p. 95)

The third position is the 'younger = better at acquiring accent' position. (Singleton, 1967, p. 107) Some researchers have now taken more differentiated positions in regards to the CPH in SLA. One of these differentiated positions is that younger learners of a second language are more effective in the acquisition of a native-like accent in the second language. (Singleton, 1967, p. 107) A study done by Fathman & Precup (1983) measured the oral proficiency of 2 groups of

[7] I.e. "very short-term experiential research, and studies based on primary school second language teaching projects and second language immersion programmes" (Singleton 95).

Spanish-speakers. Each group had 20 adults and 20 children. The first group, however, was learning English in an informal setting in the United States, while the second group was learning English in more formal settings in Mexico. What was found in this study was that the children scored higher than the adults in terms of English pronunciation. The adults, however, scored better in syntax. (Singleton, 1967, p. 109)

The final and fourth position that Singleton discusses is the 'younger = better in the long run' position. (Singleton, 1967, p. 116) This position is defined by a distinction made by Krashen *et al.* (1979). That is that "acquirers who begin natural exposure to second languages during childhood generally achieve higher second language proficiency than those beginning as adults" (qtd. in Singleton, 1967, p. 117). According to Singleton, the best evidence for this position is found in studies done by Snow and Hoefnagel-Hohle (1978a; 1978b). The first study used 69 subjects who were English-speakers living in the Netherlands. They were tested on their Dutch pronunciation. At first, adult and adolescent beginner learners showed an advantage over the younger learners. However, by the second session (four-and-a-half months later), the younger learners were catching up giving support to the 'younger = better in the long run' position. (Singleton, 1967, p. 118) The second study involved 81 English-speakers who were living in Holland.[8] Again, in the first session, the older subjects had higher

[8] There were 51 beginners, subjects who had just arrived in the Netherlands within the previous 6 months, and there were 30 advanced subjects, who had been in the country at least 18 months. (Singleton 118)

scores on the tests. However, by the second and third sessions the beginners began to catch up with the advanced group. (Singleton, 1967, p. 119)

These last three positions could all be true without necessarily contradicting the others. Singleton (1967) claims that the 'older=better' position is true in formal instruction settings, whereas the 'younger = better in the long run' position, as discussed by Singleton, is true of those learners who begin *natural* exposure to the language in childhood. The 'younger=better at acquiring accent' could also be true, again, for those learners who begin natural exposure to the language in childhood.

Summary

Lenneberg is considered one of the first researchers to suggest and examine the CPH in first language acquisition. Through him, the CPH was established along with the need for further research. Penfield and Roberts suggested that there might be maturational effects on second language learning, specifically because the direct method that is a natural process in first language acquisition is more effective than formal instruction in a second language in both physiological and psychological ways. The Johnson and Newport (1989) study was one of the most important studies to be done at that time to test the CPH in SLA. The results from this study are continually tested and re-tested. Johnson and Newport (1989) found evidence to support the strong version of the CPH, or the maturational state hypothesis, finding that no adult second language learners had achieved native-like attainment. Singleton then proposes four different

positions within the area of age effects on SLA and the CPH in SLA. It is, thus, necessary to examine the more recent research, as the historical research is outdated, and new research and studies have been conducted that have taken into consideration the results and the limitations of any older studies and experiments.

CHAPTER IV
THE CURRENT STATE OF RESEARCH OF THE CPH IN SLA

This chapter examines the research that has been conducted over the past 11 years on the CPH in SLA. First, studies that are able to use large groups of participants can provide more accurate data on the CPH in SLA. Two studies have been conducted using the 1990 and the 2000 U.S. Census data in order to obtain large groups of participants (Hakuta, Bialystok, & Wiley 2003; Chiswick & Miller 2008); however, such studies are also limited by the questions asked by the Census Bureau, and also because they comprise only self-reported information. Hakuta et al. (2003) based their studies on data collected during the 1990 Census. Participants who identified themselves as native speakers of Chinese (324,444) or Spanish (2,016,317) were included in this study. Chiswick and Miller (2008) examined information derived from the 2000 Census included 112,001 non-Mexican immigrant participants and 57, 696 Mexican immigrant participants. Both studies (Hakuta et al. 2003; Chiswick & Miller 2008) used a series of questions from either the 1990 Census or the 2000 Census that pertained to what language was spoken at home and how well the participant spoke English (a form of self-assessment). This sort of self-assessment is not ideal for assessing native-like proficiency, but the results that these studies find show a decline in the speakers' perceptions of their proficiency when correlated with their ages of arrival in the United States.

Although Hakuta et al. (2003) did not find the point of discontinuity that must be evident to provide evidence for the strong version of the critical period,

they did find that “the degree of success in second-language acquisition steadily declines throughout the life span,” as reported by the participants (Hakuta et al., 2003, p. 37). Like the Hakuta et al. (2003) study, Chiswick and Miller (2008) found a similar monotonic decline between proficiency in self-reported spoken English and the age at migration of the participants. (p. 23) These studies provide evidence against the CPH, but they do suggest age effects on second language learning. Because self-assessment of native-like proficiency is not as accurate as testing the learners’ knowledge of grammar and pronunciation, other studies must be reviewed in order to determine the current state of research on the CPH in SLA.

The studies that are reviewed in the main part of this chapter will be divided into the studies that test native-like attainment of grammatical features of a language and the studies that test native-like attainment in phonology, or native-like accent. There is a lot of discussion on which of these two determine native-like attainment. Each of these, studies testing grammar and studies testing pronunciation, test different aspects of the CPH. This review will first look at the studies and results that test grammar as a means for determining native-like attainment.

Native-like Attainment of Grammatical Features

A number of studies conducted in the past 11 years that have looked at the CPH in SLA have used grammaticality judgment tests (GJTs) to measure native-like

attainment in the L2. Several of these studies have found a negative correlation in the age of arrival, age of exposure, or age of onset and the GJT score (DeKeyser, 2000; DeKeyser, Alfi-Shabtay, & Ravid, 2009; Birdsong & Molis, 2001; Reichle, 2010; Abrahamsson & Hyltenstam, 2009; Abrahamsson & Hyltenstam, 2008; van Baxtel, Bongaerts, & Coppen, 2003; Seol, 2005). Four of these studies were replication studies of the seminal study done by Johnson & Newport (1989), which will be referred to as JN89. It is particularly interesting to compare their results across studies as well as to the results of the JN89 study.

These experiments (DeKeyser, 2000; Birdsong & Molis, 2001; DeKeyser, Alfi-Shabtay, & Ravid, 2009; Seol, 2005)[9] replicate the JN89 study to see if the same results will be yielded. All five studies found negative correlations, between the subjects scores on the grammaticality judgment test and age of arrival, that are in line with the negative correlation found by JN89, $r = -.77$, $p < .01$. DeKeyser (2000) found a negative correlation of -. 63 ($p < .001$) among 57 native speakers of Hungarian learning English; these participants had a range of AOAs from 1-40 years and had a length of residence (LOR) of at least 10 years with the average LOR being 34 years. Among 61 native speakers of Spanish learning English, Birdsong and Molis (2001) found a negative correlation of -.77 ($p < .0001$). The participants in the Birdsong and Molis (2001) study had a mean LOR in the United States of 10 years. DeKeyser et al. (2009) used 76 Russian native-speakers learning English as a second language in their first experiment with a minimum LOR of 8 years in Chicago, New York, or Toronto and 62

[9] The DeKeyser et al. (2009) study conducts two different experiments in the same study.

Russian native-speakers learning Hebrew as a second language in their second experiment with a minimum LOR in Israel of 8 years. DeKeyser et al. (2009) found a negative correlation of -.80 ($p < .001$) in the first experiment and a negative correlation of -.79 ($p < .001$) in the second experiment. Seol (2005) found a negative correlation of -.84 ($p < .01$) between 34 native speakers of Korean learning English as a second language; the minimum LOR was five years of an uninterrupted stay in the United States. The Birdsong and Molis (2001) study, the DeKeyser (2000) study, and the Seol (2005) study all found the necessity to use all second language learners from the same language background, i.e. all native Korean speakers learning English as an L2 or all native Spanish speakers learning English as an L2. This was in response to JN89's grouping both native Korean speakers and native Chinese speakers in the same linguistic category. This was found problematic, as a closer examination of potential L1 effects could not be conducted. Thus, the need for homogeneity amongst the participants in their L1 was found. (Seol, 2005, p. 7)

When these newer studies divided the results among age of arrival groups, the negative correlations that they found among these groups were different than what JN89 found when doing the same analysis. JN89 found a significant negative correlation in the age 3-15 group, r= -.87 ($p < .01$). JN89 also found a negative correlation in the age 17-39 group, but it was a weaker correlation, r= -.16 ($p < .05$). The results of the JN89 correlations provide evidence for the CPH, which states that there will be a significant decline in native-like proficiency up to the critical period, at which point the results of native-

like proficiency will flatten out with a non-significant correlation. When looking at these two correlations in the DeKeyser (2000) study, the same sort of decline can be seen, although the negative correlation for the early learners group is weaker than the negative correlation of the early learners group for the JN89 study. DeKeyser (2000) found non-significant correlations among both the early arrivals and the late arrivals. The Birdsong and Molis (2001) study divided the subjects into the same AOA groups as JN89, but had different results in the correlations. Among AOA 3-16 subjects, the negative correlation was -.24 (p = .22), and among AOA 17-44 subjects, the negative correlation was -.69 (p < .0001). Thus, the late learners group in the Birdsong and Molis (2001) study had a stronger negative correlation than the late learners group of the JN89 study as well as the other three experiments. The Seol (2005) study divided the participants into two groups, the early arrival group (AOA ≤ 15) and the late arrival group (AOA ≥ 16). Seol (2005) found strong negative correlations between AOA and performance in both groups, which is different from the other studies. Seol (2005) found a negative correlation of -.68 (p < .01) in the early arrivals group, and a negative correlation of -.66 (p < .01) in the late arrivals group. DeKeyser (2000) did not find a strong negative correlation in either group, and Birdsong and Molis (2001) only found a strong negative correlation in the late learners group.

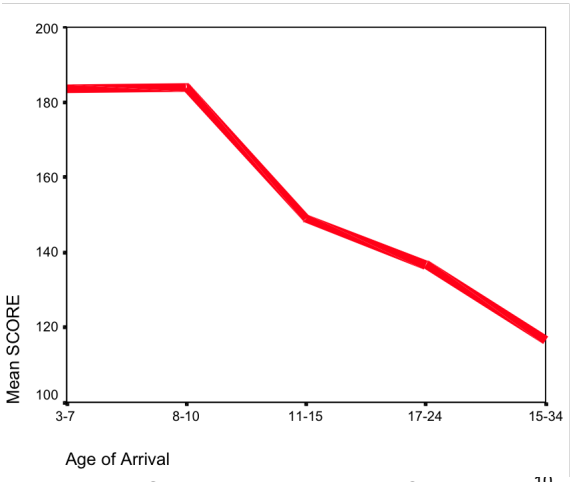

Figure 6 Seol (2005) Negative Correlation[10]

The Seol (2005) study is different from the other three studies as it found that the participants, ages 3-10, hit a ceiling effect. A stronger negative correlation is seen between the ages of 10 and 15, with it tapering off a little more after the age of 15. (See Figure 2) This is different than what JN89 found, as the ceiling effect in that study ended at age 7.

Unlike the previous three studies, the study done by DeKeyser et al. (2009) looked at the correlations within three different ages of arrival groups. In the first experiment (L1 = Hungarian), DeKeyser et al. (2009) found correlations of r= -.69 (p < .01) for AoA < 18, r= -.44 (p < .05) for AoA 18-40, and r= -.27 (*ns)* for AoA > 40. In the second experiment (L1 = Spanish), among the early arrivals, AoA < 18, DeKeyser et al. found a negative correlation of r= -.48 (p = .05). Among the younger of the late learner groups, AoA 18-40, DeKeyser et al. found a weaker negative correlation of r= -.37 (p < .05), and among the older group of

[10] From "The critical period in the acquisition of L2 syntax: a partial replication of Johnson and Newport by H. Seol, (2005), *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics 5(2),* 1-30

late learners, AoA > 40, DeKeyser et al. found a non-significant negative correlation. The two interesting differences in comparing the correlations between the different age groups are the negative correlation that Birdsong & Molis (2001) found among the late learners group and the negative correlation in the late learners group in the second experiment of the DeKeyser et al. (2009) study. The other studies found stronger negative correlations in the early learner groups, whereas Birdsong & Molis (2001) and DeKeyser et al. (2009) found equivalent negative correlations but amongst the late learner groups, indicating a very large decline in the age effects on native-like attainment, which provides evidence against the strong version of the CPH. This strong version, according to JN89, says that there will be a sharp decline in the native-like attainment of learners as a function of age of arrival, exposure, or acquisition before the end of the critical period, and that the negative correlation among late learners, with AOAs after the end of the critical period, would be much weaker than among early learners. This could be because the age at which these different studies are referring to, as the cut-off may need to be adjusted. This could account for some of the differences.

Despite these differences among the negative correlations within the different age groups, these studies all used a version of a GJT of morphosyntactic constructions used by JN89. Both the DeKeyser (2000) study and the DeKeyser et al. (2009) study shortened and adapted the grammaticality judgment test used in the JN89 study. DeKeyser (2000) shortened the test from 276 items to 200 items. The number of items per subcategory was decreased

from 6 to 8. Other changes that were made included: the deleting of some of the subcategories and practice items added to the beginning to ensure that the participants understood the procedure of the test. Changes made to the grammaticality judgment test in the DeKeyser, et al. (2009) study included: shortening the number of items from 276 to 204, adding a few extra items to the definite article category, and adding some training problems. The Seol (2005) study used the modified grammaticality judgment test that DeKeyser (2000) developed. The Birdsong and Molis (2001) study used the same exact voice recordings that the JN89 study used. Birdsong and Molis (2001) did eliminate two of the items because of an ungrammatical variant in one of the pairs.

Four of these studies provided results that were very similar to the JN89 study, providing evidence for the critical period. While the Birdsong and Molis (2001) study's overall negative correlation fell within the same range as the JN89 study's overall negative correlation, when split into age of arrival groups, this study provided counterevidence to the critical period, by showing a stronger negative correlation in the late learners category than in the early learners category. Like the Birdsong and Molis (2001) study, the overall negative correlation that was found by Seol (2005) was in line with the overall negative correlation of the JN89 study. There was, however, a difference when split into the early and late arrival groups. The Seol (2005) study found a strong negative correlation in both the early arrivals group and the late arrivals group, providing evidence that the decline in native-like proficiency as AOA increases does not

flatten out after the critical period, again providing evidence against the strong version of the CPH.

These five studies (DeKeyser 2000; Birdsong & Molis 2001; DeKeyser, Alfi-Shabtay, & Ravid 2009; Seol 2005) all provide evidence against the strong version of the CPH in SLA because none of the results produced a stretched Z function when correlated; however, all of these studies provide evidence for a general effect of age on second language learning.

Both Abrahamsson and Hyltenstam (2009) and Reichle (2010) found results indicative of a stretched 7 geometry. This stretched 7 function indicates a very slight decline, if any at all, in cognitive performance before the critical period followed by a much more significant decline post critical period. Abrahamsson and Hyltenstam (2009) found a negative correlation between the score of scrutinized nativelikeness (SN) and the age of onset (AO) of -.38 ($p < .02$), using 41-second language learners of Swedish with a minimum LOR in Sweden of 10 years (mean LOR = 25 years). Reichle (2010) conducted two different experiments in his study. In the first experiment, a negative correlation of -.46 ($p < .063$) was found, which was considered a weaker correlation, between AOA and percent correct on task scores, using 26 native-English speakers of French with AOAs ranging from 1 to 34 years old and LOR's ranging from 4 to 32 years (4 years being the minimum LOR). The second experiment was conducted to "determine if L2 speakers at asymptote perform comparably to native speakers on IS [information structure] judgment tasks" (Reichle, 2010, p. 70-2). This second experiment was also conducted to determine if the first experiment was

too easy for native-like second language speakers. In the second experiment, Reichle (2010) found that the L1 control native speakers judged 47.3% of the anomalies correctly, the low-proficiency L2 learners judged 24.9% accurately, and the high-proficiency L2 learners judged 53% accurately. The 24 L2 learners all had AOE's around the end of or after the end of the putative critical period. This high level of native-like performance in late learners provides evidence against the CPH, as the L2 learners judged more of the anomalies correctly than the L1 control native speakers.

Although the results of the Abrahamsson and Hyltenstam (2009) and the Reichle (2010) studies were relatively similar, the tests were quite different. Reichle (2010) tested information structure in French using a GJT. The test that Abrahamsson and Hyltenstam (2009) used was much broader. Specifically, the aim of the test was to judge the nativelikeness of each subject. This assessment included parts testing production and perception of voice onset time, speech perception in noise, grammaticality judgment, grammatical, lexical, and semantic inferencing, and formulaic language.

At first glance, it would seem as if Abrahamsson and Hyltenstam (2009) provide evidence to support the CPH in SLA, as they did not find any late learners who scored within the range of the native speakers. However, in the correlation between AO and SN, a stretched 7 function is produced providing evidence against the CPH. Reichle (2010), however, did find incidence of nativelikeness in the scores of the participants. Ceiling-level or near-ceiling-level performances were found among late learners on the GJT.

The final two studies discussed here that found evidence against the strong version of the CPH were studies done by Abrahamsson and Hyltenstam (2008) and van Boxtel et al. (2003). In the study by Abrahamsson and Hyltenstam (2008), 4 of the 11 (i.e. 36%) late learners scored at or above the native-speaker range on the GJT, which provides evidence against the CPH. These 11 late learners had earlier been perceived as native-like speakers in ordinary oral communication and had a mean LOR in Sweden of 25 years (range: 12-42); all 11 participants had an L1 of Spanish. Abrahamsson and Hyltenstam (2008) also found that 13 out of the 31 early learners scored below the lowest scoring native-speaker. Thus, they conclude that "when faced with a rather demanding linguistic task," 42% of the early learners failed to score native-like. (Abrahamsson & Hyltenstam, 2008, p. 496)

The van Boxtel et al. (2003) study found that 13 of the 32 (41%) of L2 late learners of Dutch scored within native-speaker range; the 32 late learners had a range of LOR from 4 to 51 years with two different language backgrounds: German and French. This percentage of L2 late learners who scored within native-speaker range is in line with what Abrahamsson and Hyltenstam (2003) found, 36% incidence of nativelikeness among late learners. These two studies found similar results despite the differences in their assessments. Abrahamsson and Hyltenstam (2008) used a GJT that specifically focused on four morphosyntactic structures of Swedish grammar that were known to be specifically difficult for L2 learners of Swedish. This GJT was administered in two versions, both auditory and in written form. The van Boxtel et al. (2003) study

focused on dummy constructions in the GJT that are also known to be difficult for second language learners of Dutch.

All of these studies that assessed grammar attainment as a means of investigating native-like attainment in second languages all provide evidence to refute the strong version of the CPH. They do, however, lend evidence towards a general age effect on native-like attainment. That is to say, all of these studies found an overall decline in native-like attainment the later the age of arrival, age of onset, or age of exposure.

Pronunciation as native-like attainment

Another common aspect that is assessed in looking at native-like attainment is second language pronunciation. Several studies have looked at native-like accent and found negative correlations between AOA, AO, or AOE and perceived native-like accent (Mackay, Flege, & Imai, 2006; Abrahamsson & Hyltenstam, 2009; Rasinger 2007). There have been a number of studies conducted that looked to see if there is a critical period for becoming native-like in the phonetics and phonology of a second language.

Three of the studies that assessed phonology found negative correlations between the AOA or AO and the scores of native-like pronunciation that are in line with the negative correlations found in studies assessing native-like proficiency in grammatical features of a second language. Studies by Mackay et al. (2006), Abrahamsson and Hyltenstam (2009), and Rasinger (2007) found negative correlations of -.75, -.72, and -.63 respectively. Mackay et al. (2006)

and Abrahamsson and Hyltenstam (2009) used larger groups of participants, n=138 and n=195 respectively, whereas Rasinger (2007) used a smaller set of participants, n=12. Mackay et al. (2006) found a negative correlation of -.75 between AOA and the degree of L2 foreign accent found among 138 native Italian speakers learning English as a second language. Their AOAs ranged from 7 to 36 years of age. The Abrahamsson and Hyltenstam (2009) study found a negative correlation of -.72 (p <.001) between AO and the score of perceived nativelikeness (PN) among 195 L2 speakers of Swedish, 107 had an AO before 12 years old and 88 of the L2 speakers of Swedish had an AO at or after 12 years of age. Rasinger (2007) found a negative correlation of -.63 (p=.01) between AOA and the Overall Proficiency Score (OPS) among 12 Bangladeshi migrants to East London. Two of these migrants had AOAs of six and eight, while the rest had late AOAs.

What differed between the Abrahamsson and Hyltenstam (2009) study and the Mackay et al. (2006) study was the type of speech sample that they took from the participants. Abrahamsson and Hyltenstam (2009) conducted interviews with the participants over the phone. The speech samples that were judged were taken from a part of the interview in which the participants were asked to talk freely for a minute about a famous Swedish children's author, Astrid Lindgren. Thus, Abrahamsson and Hyltenstam used "more or less spontaneous speech" as their samples. The Mackay et al. (2006) study used an imitation task as the speech sample. The participants listened to a recording of Native English speakers repeating a series of "questions" and "answers." The participants

listened to a question, then an answer, and then the same question through a loudspeaker. The participants were then asked to repeat the answer to the question. These participants were also allowed to listen to their recording, and if they were not satisfied with their production, they were allowed to repeat the test sentence.

Most of the studies that examine native-like attainment in a second language by assessing the second language accent use native speakers of the target language to judge whether or not the speech samples of the participants in the study are from native or non-native speakers. The Rasinger (2007) study, however, used a "slightly adapted version" of the OPS to assess the native-like proficiency of the participants in the local vernacular rather than their competence in the Standard language (Rasinger, 2007, p. 535). The OPS is a "comparatively reliable solution for measuring L2ers' spoken performance" (Rasinger, 2007, p. 535). The OPS takes into account the mean length of utterance (MLU) and a target-like score (TLS)[11]. Only 33% of the utterances of the participants are acceptable in the target vernacular. Rasinger (2007) also found a strong correlation between length of residence (LoR) and performance. Performance increases the longer the participants have lived in London. Also the coefficients between LoR and performance are stronger than the coefficients between AoA and performance. There is also a strong correlation between LoR and AoA, r = -.83. One of the most surprising findings is the low impact that AoA had on the performance of the second language learners.

[11] OPS= modMLU + TLS. modMLU refers to a modified version of the MLU.

All three of these studies found a decline in their results between the age of onset/arrival and the different scores. Abrahamsson and Hyltenstam (2009) and Mackay et al. (2006) found very similar results despite using different types of speech samples, imitation task samples and spontaneous speech samples. All three of these studies provide evidence against the strong version of the CPH, as none found a point of discontinuity in the results (Abrahamsson & Hyltenstam, 2009; Mackay et al., 2006; Rasinger 2007). They do, however, provide evidence for age effects, that is that they show a general decline in native-like proficiency of a phonological system of a second language the later one begins to acquire this phonological system.

A study by Flege, Birdsong, Bialystok, Mack, Sung, and Tsukada (2006) also looked at several correlations when assessing the degree of foreign accent in 62 native Korean speakers learning English as a second language; these participants were split into four groups: children with a LOR of 3 years, children with a LOR of 5 years, adults with a LOR of 3 years, and adults with a LOR of 5 years. Overall this study found that "native Korean children… were judged to produce English sentences with milder foreign accents than the native-Korean adults" (Flege et al., 2006, p. 168). A control group of 36 native English speakers was also included in this study. Flege et al. (2006) found a correlation of -.52 (p = .01) between the degree of foreign accent of the adult group (n=36), with a range of ages between 23 and 41, and the chronological age of the participants. A negative correlation of -.55 (p = .01) was again found in the adult group between the degree of foreign accent and AoA.

Flege et al. (2006) found little difference between the results of the two groups of native Korean (NK) adults (LOR of 3 years and LOR of 5 years), as well as of the NK children groups (LOR of 3 years and LOR of 5 years). This study assessed the accents of the participants at two different times, T1 and T2, which were a year apart. Flege et al. (2006) did not find a significant difference in the results in all four groups between T1 and T2. In this study, a year did not have a significant effect on the participants accent. (Flege et al., 2006, p. 169)

It is noteworthy that the native Korean children did, in fact, receive lower scores than the native English children from the control group. This is inconsistent with the CPH, as some of the native Korean children should have been judged with no detectable foreign accent (Flege et al., 2006, p. 169). Also in contrast to the CPH is the finding of the strong negative correlation in the adult group between the degree of foreign accent and the AOA, but not in the children's group. As Flege et al. (2006) said:

> if foreign accents arise from the passing of a maturationally defined critical period for L2 learning, then foreign accents should be increasingly mild as the age of first exposure to the L2 (AOA) nears the end of the critical period, but not as AOA extends beyond the critical period (p. 169)

In this study, the decline that should exist before the end of the critical period does not show until after the critical period has ended. Thus, providing evidence against the strong version of the CPH, but providing more evidence to the general decline that is being found between native-like attainment in second language accent and AOA.

Studies by Aburabia and Kehat (2004), Nikolov (2000), and Bongaerts, Mennen, & van der Silk (2000) used a smaller group of participants than the Abrahamsson and Hyltenstam (2009) study or the Mackey et al. (2006) study. These studies all found similar results: a lower number of late learners who were identified as native speakers of their second language. Arurabia and Kehat (2004) conducted four different tasks: a free speech task, in which the participants were asked to discuss a trip they had taken or to describe a favorite recipe; a paragraph task, which consisted of the participants reading a short passage from a Hebrew book of newspaper; a sentence task, where the participants read aloud a few sentences in Hebrew; and a word task, in which the participants read a list of 33 Hebrew words. The word task was conducted in order to cover the different phonemes in the Hebrew language. For the free speech task, Arurabria and Kehat (2004) found that 5 out of 10 non-native speakers were judged native by at least 2 judges. On the paragraph task, 4 judges judged only 1 non-native speaker as a native, and the same non-native speaker that was judged as a native speaker in the paragraph task was also the only participant to be judged as a native speaker by 3 judges on the sentence task. For the word task, 3 non-native speakers were judged as native speakers by at least 2 judges. Overall, there was only 1 non-native speaker out of 10 non-native speakers that was judged as a native speaker by at least 3 judges in each of the tasks, i.e. only 10% of the non-native speakers were fairly consistently judged as native speakers. Abrurabi and Kehat (2004) used 10 second language learners of Hebrew that were considered very native-like by the author Kehat.

This study also looks at the language learning background of each learner. Ilana, the 1 non-native speaker who was nearly consistently judged as a native speaker of Hebrew, was a native Romanian speaker who arrived in Israel at the age of 10 ½. She had been in the country for 40 years, and was married to a Hebrew native speaker. She was currently a teacher of Hebrew, science, and Arabic. She knew several languages and was often thought to have an oriental or eastern European accent. (Abrurabi & Kehat, 2004) Thus, the only participant judged as a native speaker began acquisition, or exposure, before the age at which the critical period ends. In this study Abrurabi and Kehat (2004), as most other studies, consider age 12 to be when puberty begins, or the age at which the critical period ends. Three of the five participants who were judged in the free speech task as a native speaker all began acquisition of, or were exposed to, Hebrew after the age of 12. These participants were all exposed to Hebrew, however, by the age of 16. There was one exception of a non-native speaker with an AOE of 20 who was judged as a native speaker in the word task by two judges. This was, however, the only task on which this participant was judged as a native speaker. (Abrurabi & Kehat, 2004) This study lends evidence against the strong version of the CPH, as all of the other studies have so far, but it also lends evidence to the idea stated by Birdsong that the critical period must be extended past puberty; “if the performance of exceptional learners is to be accounted for in biological terms, then the hypothesized end of the critical period must be pushed well past puberty, or the ‘window of opportunity’ for language learning must be extended and made flexible” (Birdsong, 1992, p. 742).

The Nikolov (2000) study conducted two experiments using 20 learners of Hungarian with varying native languages and 13 native Hungarian speakers learning English as a foreign language, all of whom began SLA or foreign language acquisition (FLA) at the age of 15 or later, thus after puberty. All of the native Hungarian speakers learning English as an L2 had studied abroad at some point in the US. These subjects were interviewed and were asked first to tell a story and then to read a passage aloud. 58 native speakers of Hungarian were asked to judge the speech samples of the first experiment and 36 native English speakers were asked to judge the speech samples of the second experiment by selecting whether the participants in the recordings were native speakers of Hungarian or non-native speakers. Out of the 20 participants in the first experiment, five of the participants were judged by at least 55% of all the judges as native speakers of Hungarian. (Nikolov, 2000) Two participants in the Nikolov (2000) study were judged as native speakers of Hungarian by 97% and 98% of the judges, respectively. The other three participants were judged as native speakers of Hungarian by 71%, 60%, and 55% of the judges. In the second experiment, only 1 of the participants was judged by 89% of the judges as a native speaker of English. There were four other participants who did relatively well, but the percentage of judges that evaluated them as native speakers was still lower than in the first experiment. Participants 8, 9, 10, and 11 were judged as native speakers of English by 47%, 50%, 42%, and 56% of the judges, respectively. As Nikolov (2000) stated, the judges in the second

experiment were less impressed with these foreign language learners of English, therefore they received much lower scores.

One main difference in the results of these two experiments is the fact that one judged pronunciation of SLA and the other judged pronunciation of FLA. Thus, length, amount, and type of exposure play a big role in the results. However, this study does provide evidence against the strong version of the CPH, but lends evidence to a non-maturational decline in native-like attainment of a phonological system in a second language.

Bongaerts et al. (2000) also found late learners who were judged as native speakers of a second language in their pronunciation, thus providing more evidence against the strong version of the CPH. Bongaerts et al. (2000) used 30 advanced learners of Dutch as an L2, who had arrived in the Netherlands between the ages of 11 and 34. All of the participants read aloud ten sentences, "which contained multiple examples of all but the most marginal Dutch phones" (Bongaerts et al. 2000). Eleven experienced judges and ten inexperienced native speaker judges judged these speech samples. Experienced meant the judges were teachers of Dutch as an L2, and inexperienced meant these judges had no formal training with Dutch as an L2 or with linguistics or phonetics. Because of the standard deviation difference between the ratings of the experienced judges and the ratings of the inexperienced judges, Bongaerts et al. (2000) decided to analyze these two data sets separately. The experienced judges judged four of the participants as native speakers with a standard deviation of < 2, according to Flege et al.'s criterion of nativelikeness. The

inexperienced judges judged only two of the participants as native speakers. Therefore, if the ratings from the experienced and inexperienced judges are combined, two of the 30 late L2 learners of Dutch were judged as native speakers of Dutch,

Stefanik (2001) conducted a study to "verify the validity of the CPH in the Slovak language." Ten-second language learners of Slovak were used as participants, as well as ten native speakers of Slovak. All of the second language learners had an AOA after 16 years old. These subjects were asked to read a short text and write a short essay. This study did not just assess the perceived nativelikeness of second language accent, but also assessed the perceived nativelikeness in a written text. Looking at the results of the recordings, 55.7% of the non-natives were judged correctly, which means the judges thought that 44.3% of the non-native speakers were native speakers of the Slovak language. The judges also, however, only judged 69.7% of the native speakers correctly. The percentage of correct judgments on the written test was lower for both the native speakers and the non-native speakers. Only 45.5% of the non-native speakers were judged correctly, which again means that the judges thought 54.5% of the non-native speakers were native speakers. Again, one must take into account that the judges only judged 65.4% of the native speakers of Slovak correctly. (Stefanik, 2001) Because of the high percentage of misidentifying the non-native speakers as native speakers, these results then provide evidence against the strong version of the CPH.

Summary

The research reviewed in this chapter provides evidence against the stronger version of the CPH in SLA; the stronger version of the CPH claims that second language acquisition will not happen outside of this critical period. Evidence of native-like attainment in second language learners who began SLA after puberty refutes the CPH, and that is what was found in this review. The studies testing both grammatical features and pronunciation provided evidence against the stronger version of the CPH as native-like incidence was found in most of the studies.

This review does provide evidence for a general age effect on SLA. Several of the studies found a linear monotonic decline in their negative correlations between AOA, AO, or AOE and the scores received on performance on GJTs or on perceived nativelikeness exercises. This linear decline suggests that there is a general age effect in SLA, meaning that native-like attainment is not limited to just those learners who acquire a second language before puberty, but that it is generally easier for children to acquire a second language. For those learners who begin after childhood, these studies provide evidence that it is possible. Thus starting SLA before puberty, in childhood will increase the language learning capacity and allow it to expand past puberty.

The results do support three of the four positions that Singleton labels in his book *Language Acquisition: The Age Factor* (1989). The “younger = better” position claims that younger learners learn or acquire a second language easier, faster, and more effectively than older learners (Singleton, 1967, p. 61). The

studies that examined both early and late learners of a second language had more early learners receive higher scores or be judged as native-like than late learners. The second position that is supported by these studies is the "younger = better at acquiring accent" position (Singleton, 1967, p. 107). This position holds that young learners are more effective in acquiring the phonological system of a second language. The studies that examined the perceived nativelikeness of second language learners in relation to AOA provided results that support that younger learners have a greater capacity in acquiring a native-like accent in a second language. The third position that is supported from this chapter is the "younger = better in the long run" (Singelton, 1967, p. 116). This position claims that those learners exposed to SLA in childhood achieve a higher proficiency in a second language than those learners who are exposed to SLA for the first time in adulthood. Native-like incidence was found among adult or late learners in several of the studies in this chapter. However, more of the early learners achieved native-like attainment than the adult learners.

Overall, this review provides evidence that there is not a critical period in which one must acquire a second language before the end in order to attain that language fully. This review does, however, confirm that there is a general age effect in SLA, and that younger learners have a higher or greater capacity in language learning than adults, but that it is not impossible after puberty.

CHAPTER V
OVERVIEW OF FOREIGN LANGUAGE IMMERSION PROGRAMS

The results from the studies of the CPH in SLA all provide evidence to a general decline in native-like proficiency in a second language as the age of arrival, exposure, or acquisition increases. Therefore, one can acknowledge the decision that starting foreign language education earlier in schools could be more beneficial for the students in becoming more proficient in second languages. This begs the question of why do parents who wish for their children to be more proficient in a second language not raise their children bilingual? Bilingualism is becoming more and more popular and necessary to compete in this world:

> As the world becomes more interconnected, it is increasingly apparent that bilingualism is the rule and not the exception. Not only do some countries support bilingual populations because of cultural and linguistic diversity within its citizenry, but also increased global mobility has enlarged the number of people who have become bilingual at all levels of society (Bialystok, 2009, p. 89)

For some people, however, this is not an option as maybe they do not speak a second language themselves or are not proficient enough in more than one language to try to raise their children as bilingual speakers. Out of this desire for children to become more proficient in second languages arose the topic of bilingual education. There are several different types of bilingual education programs. The type of bilingual education that this chapter will focus on is what is called a language immersion program. In the discussion portion of their

seminal study, Johnson and Newport (1989) concluded from their results that it may be beneficial for earlier classroom exposure to a second or foreign language. They suggested that the classroom variables may be more significant in the research if these classrooms were specifically immersion classrooms, i.e. the research may yield higher incidence of native-like attainment and proficiency in younger children. (Johnson & Newport, 1989, pp. 83)

History of Language Immersion Prorams

In the past 40 to 50 years, the term “language immersion program” has taken on several different meanings. There are many different types of immersion programs that are now offered all over the world. The most common are probably those that are offered in a country that speaks the target language. Students travel to these countries to live for a set amount of time immersed in the target language. While living in the country of the target language may be ideal, it is not always possible, especially if one takes into account that the CPH in SLA suggests starting SLA earlier in life. Thus, the program that initially developed the term “immersion” will be the focus of this chapter.

This new immersion program began in 1965 in Canada. These programs arose out of the concerns of English-speaking parents in St. Lambert, Quebec, an English-speaking community. The children of these parents went to a school where the instruction was in English, but they were receiving formal instruction in a course in French as a second language. These parents were concerned that their English-speaking children were not proficient enough to compete in the

growing French-speaking community. They lobbied their school board to figure out how to better the teaching of French as a second language. The change of moving all instruction during the first three to four years in school to all French was proposed and accepted. (Johnson & Swain, 1997, pp. 2)

What are foreign language immersion programs?

There are numerous immersion programs that have been started in the world, including in Canada and in the U.S. These programs seek to aid students in becoming bilingual by teaching all subjects for the first three to four years in the second language. Thus, the target language becomes the medium through which children will learn all of the subjects and does not constitute a subject in itself. Consequently, this feature enables children enrolled in immersion programs to learn the target language in a more naturalistic setting, through a more direct method, which Penfield and Roberts (1969) suggested yielded better outcomes. The students implicitly learn the target language while they are explicitly learning math, history, or any other subject. As the students get older, the percentage of courses taught in the second language lessens, and curriculum is divided relatively evenly between the first and second languages.

In their book *Immersion Education: International Perspectives*, Johnson and Swain (1997) discuss these immersion programs and state that all of these programs usually contain the following eight core features:

- The L2 is a medium of instruction
- The immersion curriculum parallels the local L1 curriculum

- Overt support exists for the L1.
- The program aims for additive bilingualism
- Exposure to the L2 is largely confined to the classroom
- Students enter with similar (and limited) levels of L2 proficiency
- The teachers are bilingual
- The classroom culture is that of the local L1 community (p. 8-9)

Most of these eight features must be present in a program in order to refer to itself as being an "immersion" program (Johnson & Swain 1997, p. 8). Other features may vary among programs. One such feature is the age at which students enter immersion programs. There are early, middle, and late immersion programs. Early immersion programs are where the students begin their formal education in the target language, at the age of four or five. Middle immersion usually begins in fourth or fifth grade, and late immersion programs begin in grades six and seven. Another feature is the "extent of immersion" (Johnson & Swain 1997, p. 9). Some programs are full immersion programs, i.e. all instruction is in the target language, and others are partial immersion programs, in which, the instruction is split 50/50 between the first language of the learners and the targeted second language. (Johnson & Swain 1997, p. 9)

Three main questions then arise about these language immersion programs. Which type of immersion program is better or yields better results? What, if any, are the advantages of these immersion programs in regards to second language learning? And finally, what is the current state of language immersion programs in the U.S., and are they successful?

Which type of immersion program yields better results?

The next question that developed after the initial start of the immersion programs was between which was not only more effective but better for the students: early immersion programs, mid immersion programs, or late immersion programs? A study done by Turnbull et al. (1998) studied the time on task and the proficiency of graduates of immersion French programs. Turnbull et al. (1998) looked at the amount of exposure that students who complete French immersion programs in Canada have in the target language. For early immersion (EI) programs, students usually received around 6000 total hours of French exposure by the end of grade eight. Students in middle immersion (MI) and late immersion (LI) programs received between 1200 and 2000 hours of French exposure. Students in all three, EI, MI, and LI, programs also end up acquiring 1000 to 1500 hours of French exposure in high school courses that were taught in French. (Turnbull et al., 1998, p. 32)

Turnbull et al. (1998) tested 1160 seniors who were graduating from EI, MI, and LI programs. The *Senior French Proficiency Test Package for French Immersion* was used in this study. This test covers the four skill areas. In the listening area, there are two tests to measure comprehension of spoken French. The reading section includes three passages that the students must read and then answer several multiple-choice comprehension questions. There are two writing tasks in the writing section of the test: a cloze test and a free writing task, in which the students are asked to express an opinion and support it with examples. The speaking test consists of two tasks as well. The first is a

sentence repetition or imitation task, and the second is an oral opinion measure – the students are asked to express their opinion orally on a given subject.

Turnbull et al. (1998) found that the EI students scored significantly higher on eight of the test measures involving listening and speaking on the sentence-repetition task and the cloze test than the MI students. The EI students also significantly outperformed the LI students on the listening and speaking measures of the sentence-repetition task. (Turnbull et al., 1998, p. 39) Turnbull et al. (1998) did not, however, find any significant difference on the listening total score, the oral and written opinion scores, and the reading test score between EI students and both LI and MI students. The scores on all of the tests between the MI and LI students were compared, and no significant differences were found.

Turnbull et al. (1998) also wanted to look at "to what degree the test score differences across programs are proportional to total accumulated instructional hours in French" (p. 41). In other words, is there a correlation between the length of exposure to French and the students' test scores? Turnbull et al. (1998) expected the EI students to score higher on the tests than the MI and LI students because the EI students had accumulated 2.3 and 2.5 times as many hours of French exposure as either MI students or LI students. The interesting results are that the EI students did score better on a few of the tests, but it was restricted to mostly the speaking area. (Tunrbull et al., 1998, p. 41)

What are the advantages of language immersion programs?

Empirical studies reveal that immersion programs primarily present two types of advantages with regard to second language learning and proficiency. The first is the overall native-like proficiency that students gain through immersion instruction. The second is the metalinguistic awareness that students acquire because of becoming bilingual at such a young age.

One specific study that is included in Johnson and Swain's book, *Immersion Education: International Perspectives* (1997), is a study by Duff that looks at three Hungarian-English dual language (DL) programs. Dual language here refers to the same type of immersion programs that have been described previously. Duff used a 150-item test that was taken from Forms 1 and 2 of the Educational Testing Service's Secondary Level English Proficiency (SLEP) to assess the proficiency of the DL students in listening and reading comprehension. (Duff, 1997, p. 33) This study also used two cloze tests, questionnaires, graded writing samples, and structured oral interviews that were administered, along with the SLEP, at both the beginning and the end of the year. Overall, this study found that students who had only been in the program a year or two demonstrated near-ceiling performance on the cloze tests and on the SLEP test. There was, however, room for improvement in the students' scores on the writing samples. One interesting note is that these students in the DL program after the first year performed within comparable range of a group of native English speakers who had taken the test earlier. (Duff, 1997, p. 34)

Two studies done by Bialystok and Barac (2011) show significant differences in the metalinguistic awareness of bilingual children that participated in language immersion programs than in monolingual children. These studies were conducted in schools in Canada. The first study used 100 children in grades 2 and 3, who were enrolled in a school where Hebrew was the language of instruction. 65 of these children reported speaking only English at home, and the rest of the 35 children spoke primarily Hebrew or Russian at home. The second study had 80 children in Grade 2 who spoke primarily English at home but were enrolled in a French immersion school.

The first study used a metalinguistic task, which tested the ability of the children to apply morphological rules of English to unfamiliar forms, and a nonverbal executive control task. Non-verbal executive control performance consisted of two tasks, a flanker task, which consisted of children indicating the direction of an arrow that appeared in the middle of the screen by clicking on either the left or right side of the screen, and a task switching test, in which students paired opposite pairs[12] by either color or shape, depending on what was indicated on the screen. (Bialystok & Barac, 2011, p. 68) The second study used a metalinguistic task, where the children listened to sentences that were grammatical and ungrammatical, and had to decide if the sentence was said correctly or not. (Bialystok & Barac, 2012)

[12] A blue horse and a red cow would appear at the top of the screen and a red horse and a blue cow would appear at the bottom center of the screen.

Both studies found that “metalinguistic performance improved with increased knowledge of the language of testing and executive control performance improved with increased experience in a bilingual education environment” (Bialystok & Barac, 2011, p. 71).

Have these programs been implemented into the United States, and have they succeeded?

The final question to be addressed in this initial look into language immersion programs was: have these types of programs been implemented in the U.S.? And have they succeeded? The answer to both of these questions is yes. According to Johnson and Swain (1997), as of 1997, there were 187 elementary immersion programs in the U.S. (Johnson & Swain, 1997, p. 243) About 40% of these programs were early total immersion, meaning the programs started in kindergarten and all instruction was in the target language, and about 60% of these programs were early partial immersion, which started in kindergarten as well, but only a portion of the instruction was in the target language. (Met & Lorenz, 1997, p. 243) As of 1997, there had not been a lot of research done to assess the language proficiency of the students in these immersion programs in the U.S. (Met & Lorenz, 1997, p. 256)

Genesee (1985) points out that these immersion programs in the U.S. were developed for different reasons than those developed in Canada. Genesee suggests that these purposes are: “(a) as linguistic, cultural, and general educational enrichment; (b) as magnet schools to bring about a more balanced ratio of ethnolinguistic groups; and/or (c) as a means of achieving some degree

of two-way bilingualism in communities with large populations of non-English speaking residents" (Genesee, 1985, p. 544). Genesee relates these purposes to the differences in programs both between the programs in Canada and the programs in the U.S. and between the programs in the U.S.

Looking at the statistics from the Center for Applied Linguistics, a significant jump over the past five years can be seen. In 2006, there were 263 language immersion programs in the U.S in 33 states, while in 2011, there were 448 language immersion schools in the U.S. in 38 states. (Center for Applied Linguistics, 2011) Of these 448 schools, 97 are pre-school immersion schools, 337 are elementary immersion schools, 128 are middle school immersion programs, and 41 are high school immersion programs. This significant jump in the amount of language immersion programs is around 70.3%, with an average annual growth rate of 14%. In the 35 years prior, the average annual growth rate was only 2.8%. Of the 448 language immersion programs that exist in the U.S. presently, Spanish and French immersion programs make up the majority of these programs, 45% and 22%, respectively. (Center for Applied Linguistics, 2011)

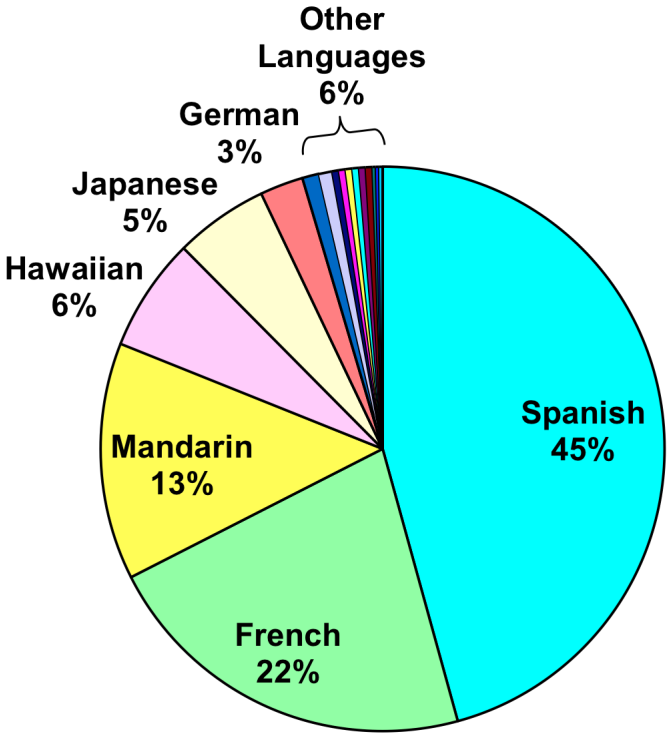

Figure 7 Percentage of Immersion Programs by language of instruction[13]

This significant jump in the number of programs could be a result of a number of reasons. One very likely reason is the need for young Americans who wish to make their way in this ever-changing world to compete at the international level in business, school, and politics. As technology continues to connect countries and cities with their counterparts on the other side of the planet, the need to be able to communicate well and effectively in a second language continues to rise. Taking into account the research that has been done on the CPH in SLA, the general age effect that has been found on native-like

[13] From "Center for Applied Linguistics. (2011). *Directory of foreign language immersion programs in U.S. schools.*

attainment suggests that the earlier a child can start to learn a second language, the better. And taking into account that the majority of the research on the CPH in SLA tests immigrants who have learned their second language in this immersion type setting, it is necessary to assume that these foreign language immersion programs are about as close to the type of immersion these immigrants have experienced as one can get without actually moving to a different country. Therefore, these language immersion schools are an effective way to promote and incorporate immersion in childhood.

CHAPTER VI
CONCLUSIONS

The research on the CPH in SLA currently provides evidence against the stronger version of the CPH as set out by both Singleton and Ryan (2004) and Johnson and Newport (1989). The stronger version claims that before puberty, human beings have a strong capacity to learn and acquire languages and that this capacity declines, or perhaps even, disappears after the onset of puberty. The studies reviewed in chapter four of this thesis provide evidence against the stronger version by presenting late adult learners of a second language who received scores on grammaticality judgment tests and on scrutinized nativelikeness phonological exercises that were within the native speakers scores. These studies also provide evidence against the stronger version of the CPH by not producing a stretched Z geometric shape in the correlations of the test scores and age of arrival, exposure, or onset.

The general monotonic decline that is found in the correlations does, however, provide evidence of a general age effect on native-like attainment in a second language. The general age effect on language acquisition does suggest that children have a larger capacity for learning languages before puberty, which is similar to the weaker version of the CPH as proposed by Singleton and Ryan (2004). This does not suggest that a second language cannot be acquired after puberty but that before puberty, humans have a larger capacity for it. Thus, making SLA easier and perhaps less extenuating for children.

Foreign language immersion programs were developed, initially in Canada, because of parents' disappointment at their children's' native-like proficiency in a second language. These immersion programs have spread throughout the United States and throughout the world. Currently, there are 448 language immersion schools in the United States in 38 different states. These programs all differ in early, mid, or late immersion as well as partial or full immersion, but they are working to increase the proficiency and attainment of children in a second language. These foreign language immersion programs are similar to the direct method, or the mother's method, that Penfield and Roberts (1959) claimed was more effective than formal instruction in a language, because children acquire the language easier through implicit learning and input rather than explicit instruction. Children in foreign language immersion programs acquire a second language implicitly through their learning of other subjects. Formal instruction may be used to help enforce or to clarify specific aspects, but most of the learning happens implicitly. Thus, foreign language immersion programs are the answer to this general age effect on native-like attainment of a second language in the realm of foreign language education.

This review was limited in the amount of information that pertained to foreign language immersion programs, to the number of articles or studies that have been written on their curriculum and to the number of studies conducted to test their success in helping children in the programs achieve native-like attainment.

More research needs to be conducted on the success of the foreign language immersion programs. There are several studies that focus on their proficiency in the subjects that are being taught through the second language, but more research is needed on the effectiveness of these programs on native-like language attainment and proficiency. Comparisons between those children who grow up bilingual (specifically immigrant children), children that have gone through foreign language immersion programs, and students who learn a foreign language through formal instruction that normally begins in late elementary or middle school would be interesting to look at. The comparison between immigrant bilingual children and those children in the foreign language immersion programs is needed to again test the effectiveness of these programs on native-like language attainment and proficiency.

LIST OF REFERENCES

Abrahamsson, N. & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *SSLA* 30, 481-509.

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

Abu-rabia, S. & Kehat, S. (2004). The critical period for second language pronunciations: is there such a thing? *Educational Psychology: An International Journal of Experimental Educational Psychology* 24, 77-97.

Bialystok, E. (1997). The structure of age: in search of barriers to second language acquisition. *Second Language Research* 13(2), 116-137.

Bialystok, E. (2009). Bilingual Minds. *Psychological Science in the Public Interest* 10 (3), 89-129.

Bialystok, E. & Barac, R. (2011). Emerging bilingualism: dissociating advantages for metalinguistic awareness and executive control. *Cognition* 122, 67-73.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language* 68, 706-755.

Birdsong, D. (2005). Interpreting age effects in second language acquisition. In J. F. Knoll and A. M. B. de Groot (Eds), *Handbook of bilingualism: Psycholinguistic approaches,* 109-127. New York: Oxford University Press.

Birdsong, D. & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language,* 44, 235-249.

Bongaerts, T., Mennen, S., & van der Silk, F. (2000). Authenticity of pronunciation in naturalistic second language acquisition: the case of very advanced late learners of Dutch as a second language. *Studia Linguistics* 54(2), 298-308.

Center for Applied Linguistics. (2011). *Directory of foreign language immersion programs in U.S. schools.* Retrieved from http://www.cal.org/resources/immersion/

Chiswick, B. & Miller, P. (2008) A Test of the Critical Period Hypothesis for Langauge Learning. *Journal of Multilingual and Multicultural Development,* 29 (1), 16-29.

DeKeyser, R. M. (2000) The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition,* 22, 499-533.

DeKeyser, R., Alfi-Shabtay, I. & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics* 31(3), 413-438.

Duff, P.A. (1997). Immersion in Hungary: an EFL experiment. In R.K. Johnson & M. Swain (Eds.), *Immersion Education: International Perspective,* 19-43. New York: Cambridge University Press.

Flege, J.E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics* 34*,* 153-175.

Genessee, F. (1985). Second language learning through immersion: a review of U.S. programs. *Review of Educational Research* 55 (4), 541-561.

Hakuta, K., Bialystok, E. & Wiley, E. (2003) Critical Evidence: A test of the critical period hypothesis for second-language acquisition. *Psychological Science* 14 (1), 31-38.

Johnson, J.S. & Newport, E.L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21, 60-99.

Johnson, R.K. & Swain, M. (1997). *Immersion Education: International Perspectives*. New York: Cambridge University Press.

Lenneberg, E. (1967). *Biological Foundations of Language.* New York: John Wiley & Sons, Inc.

Lightbown, P. & Spada, N. (2006). *How Languages are Learned* (3rd Ed.) Oxford, UK: Oxford University Press.

Mackay, I.R.A., Flege, J.E., & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics* 27, 157-183.

Met, M., Lorenz, E.B. (1997). Lessons from U.S. immersion programs: two decades of experience. In R.K. Johnson & M. Swain (Eds.), *Immersion*

Education: International Perspective, 243-264. New York: Cambridge University Press.

Nikolov, M. (2000). The critical period hypothesis reconsidered: successful adult learners of Hungarian and English. *IRAL* 38, 109-124.

Penfield, W. & Roberts, L. (1959) *Speech and Brain-Mechanisms.* Princeton: Princeton University Press.

Reichle, R.V. (2010) Judgments of information structure in L2 French: nativelike performance and the critical period hypothesis. *International Review of Applied Linguistics in Language Teaching* 48, 53-85.

Rasinger, S. (2007). Revisiting the critical period: age of onset or length of exposure. *Proceedings of the Annual Boston University Conference on Language Development* 31 (2), 532-542.

Seol, H. (2005). The critical period in the acquisition of L2 syntax: a partial replication of Johnson and Newport (1989). *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics* 5(2), 1-30.

Singleton, D. (1989) *Language Acquisition: The Age Factor*. Philadelphia: Multilingual Matters LTD.

Stefanik, J. (2001). The critical period hypothesis and the Slovak language. *Applied Language Learning* 12 (2), 161-176.

Turnbull, M., Lapkin, S., Hart, D., & Swain, M. (1998). Time on task and immersion graduates' French proficiency. In S. Lapkin (ed.), *French Second Language Education in Canada: Empirical Studies*, 31-55. Toronto: University of Toronto Press.

Van Boxtel, S., Bongaerts, T., & Coppen, P.-A. (2003). Native-like attainment in L2 syntax. *EUROSLA Yearbook* 3, 157-181.

VITA

Katherine Nelson was born on April 11, 1988 in Radford, Virginia. She is the second daughter of Robert Nelson and Vanessa Nelson. After completing her work at Johnson County High School in Mountain City, Tennessee, Katie went on to study French, Youth Ministry, and Mathematics at King College in Bristol, Tennessee, as well as to play softball for the lady tornadoes. In May of 2010, she completed a Bachelor of Arts from King College and began her graduate studies of French Literature at the University of Tennessee in Knoxville, Tennessee in the Fall of 2010.