The Critical Period in the Acquisition of L2 Syntax: A Partial Replication of Johnson and Newport (1989)

Heekyung Seol[1]
Teachers College, Columbia University

ABSTRACT

The present study was conducted as an attempt to partially replicate the Johnson and Newport (1989) study (JN89). Given that JN89 yielded the most clear-cut evidence demonstrating the presence of the Critical Period (CP) in the domain of L2 syntax, it has been regarded as one of the landmark CP investigations. This study sought to ascertain whether the evidence provided in JN89 is replicable with a different population of 34 Korean native speakers of English. The results, although not in complete agreement with those of JN89, provide some modest support for the existence of the CP.

INTRODUCTION

Since Lenneberg (1967) put forth the Critical Period Hypothesis that brain maturation constrains human language acquisition, the Critical Period (CP) has been an issue of primary attention in the field of both first language acquisition (FLA) and second language acquisition (SLA). In FLA, research has reached a general consensus as to the existence of the CP on the basis of the observations of abnormal linguistic contexts in which language learners were deprived of the chance to acquire their first language (L1) within the CP (e.g., Genie [Curtiss, 1977]; Chelsea [Curtiss, 1988]; and deaf children using American Sign Language (ASL) [Newport, 1984; Newport & Supalla, 1987]). Despite this consensus in FLA research, there is still plenty of controversy in the field of SLA.

In SLA, proponents of the CP have argued that the general advantage of children over adults, in acquiring a second language (L2), springs from children's innate biological superiority, which is accessible only within the CP, and thus, that those who start their acquisition within the CP can successfully achieve SLA, whereas those outside of it cannot (e.g., Asher & Garcia, 1969; DeKeyser, 2000; Harley & Wang, 1997; Johnson, 1992; Johnson & Newport, 1989; Patkowski, 1980; Shim, 1993). Some researchers, on the other hand, have contended that the age difference prevalent in SLA may result from factors other than maturational constraints, including education (Bialystok & Hakuta, 1999); use of L1 and L2 (Flege, Yeni-Komshian & Liu, 1999); and socio-psychological factors (Schumann, 1978). More strongly, researchers who

[1] Heekyung Seol is a doctoral student in TESOL at Teachers College, Columbia University. Her current research interests include age-related differences in second language acquisition with a focus on the interaction of the Critical Period and L1 influence on L2 ultimate attainment. Correspondence should be sent to Heekyung Seol, 1254 River Rd #12, Edgewater, NJ, 07020. Email: hs246@columbia.edu.

argue against the CP have sought to disprove the existence of maturational constraints by empirically demonstrating ongoing age effects after the closure of the putative CP (e.g., Bialystok & Miller, 1999; Birdsong & Molis, 2001) or nativelike achievement reached by late starters (e.g., Bongaerts, van Summerren, Planken, & Schils, 1997; Flege, 1987; Ioup, 1984; White & Genesee, 1996).

Most notable among a number of theoretical and empirical studies on this controversial topic is the study conducted by Johnson and Newport (1989). JN89 examined the acquisition of L2 syntax by non-native speakers who began their acquisition of English at different ages. Their findings have been considered to be the most clear-cut evidence which demonstrates the presence of maturational constraints in SLA, and thus many subsequent studies have sought to verify or falsify the results of JN89. This line of studies has constituted one major strand of recent CP research. A number of researchers have replicated JN89 (e.g., Bialystok & Miller, 1999; Birdsong & Molis, 2001; DeKeyser, 2000; Johnson, 1992; Johnson & Newport, 1991): some reanalyzed its data by using different statistical procedures (e.g., Bialystok & Hakuta, 1994; Birdsong & Molis, 2001); and some have reviewed its methodology (e.g., Bialystok & Hakuta, 1994; Kellerman, 1995).

The present study is another attempt to replicate JN89. Its primary purpose is to ascertain the existence of maturational constraints in the domain of L2 syntax by examining whether the evidence provided in JN89 is replicable with a different population of 34 Korean native speakers of English. This study employs the same research questions as in JN89 and similar methodologies.[2] Finally, in comparing findings, similarities and dissimilarities between the two studies are identified and discussed, and possible directions for future research are provided.

LITERATURE REVIEW

The examination of the CP in the acquisition of L2 syntax has been prevalent in recent research. Although a number of empirical CP studies have indeed been conducted since the late 1960s, the focus has been predominantly on phonology (e.g., Asher & Garcia, 1969; Neufeld, 1977, 1979; Seliger, Krashen, & Ladeforged, 1975; Snow & Hoefnagel-Hohle 1978; Tahta, Wood, & Loewenthal, 1981). This primary attention given to phonology has been attributed to two major presumptions: (a) phonology is the linguistic domain most susceptible to maturational effects, since pronunciation entails neuromotor involvement on a neuromuscular basis (Scovel, 1969, 1988); (b) phonology is the most examinable finite subsystem of a language (Seliger, 1978). In recent years, nonetheless, attention has shifted to some extent toward morphosyntax. The major contribution to this shift was made by Johnson and Newport's 1989 study.

The Johnson and Newport 1989 Study

In an attempt to fill the void in the previous research with respect to convincing evidence of an age effect in the ultimate attainment of L2 syntax, Johnson and Newport examined the morphosyntactic proficiency of 46 native Korean and Chinese speakers who had begun their

[2] Given that the original grammaticality judgment test used in JN89 was inaccessible, the present study adopted DeKeyser's (2000) modified version. Due to this methodological gap, the current study cannot be considered as a direct replication of JN89, but rather as its indirect partial replication.

acquisition of English at differing ages. To this end, they recruited participants who had arrived in the United States between the ages of 3 and 39. All of them had a minimum five-year length of residence (LoR), and at least three years of uninterrupted stay in the U.S. Among these participants, 23 were early arrivals who came prior to the age of 15, while the others were late arrivals who came after the age of 17. In order to prevent too much variation between participants' diverse experiential backgrounds, Johnson and Newport also ensured that all the participants had received some schooling in the U.S., and were either university faculty or students.

A grammaticality judgment test[3] was used, in order to assess the participants' morphosyntactic proficiency. The test encompassed 12 rule types which are regarded as the most fundamental aspects of English sentence structure: past, plural, third person singular, present progressive, determiner, pronominalization, particle movement, subcategorization, auxiliary, yes/no question, wh-question, and word order. The test consisted of 276 sentences, among which were 136 grammatical sentences and 140 ungrammatical counterparts. All of these sentences were presented in the oral mode on a tape recorded by a female native speaker of American English. After listening to each sentence, the participants were instructed to make a judgment of its grammaticality. Along with this test, they were also asked to fill out a language questionnaire requesting information on experimental variables (length of exposure, amount of initial exposure, and years of English classes) and attitudinal variables (identification, self-consciousness, and motivation).

The study demonstrated a strong age effect on the acquisition of L2 syntax by showing that the participants' performances were negatively correlated with the age of arrival (AoA) for the group as a whole (r=-.77, $p < 0.01$). More specifically, Johnson and Newport observed a different pattern of age effect for the early and for the late arrivals. Among the early arrivals, the participants whose AoAs were between the ages of three and seven showed full attainment of native-like competence, which is often referred to as the *ceiling effect*. Those who began their acquisition between ages 8 and 16 showed a linear decline in performance with increasing age. The late arrivals' performance, however, was not only low but also highly variable. Their performance was marked by a random distribution of scores without any significant correlation between scores and AoA. Johnson and Newport pointed to this particular pattern as providing the distinct evidence in support of the CP:

> If the explanation for late learners' poorer performance relates to maturation, performance should not continue to decline over age. . . . Instead, there should be a consistent decline in performance over age for those exposed to the language before puberty, but no systematic relationship to age of exposure, and a leveling off of ultimate performance, among those exposed to the language after puberty. This is precisely what was found (p. 70).

Replications of JN89

Given that it demonstrated a specific pattern of age effects featuring maturational constraints, JN89 came to be generally regarded as a landmark CP study in the domain of L2

[3] The test used in JN89 was modeled on the syntax test of Linebarger, Schwartz, and Saffran's (1983) study.

syntax. Due to the authors' distinctive findings, JN89 has served as a strong impetus for subsequent research.

A number of studies have attempted to replicate JN89. Among these, some have produced findings compatible with those of JN89. Johnson and Newport (1991) themselves conducted another study, similar to the previous one. Their second study was conducted in an attempt to see whether the finding of JN89 could be replicated with respect to the properties of language relating to Universal Grammar (UG).[4] The participants were 21 native Chinese speakers of English who had arrived in the United States between the ages of 4 and 38. A negative correlation was found between AoA and the performance for the whole group. More specifically, there was a steady decline in performance according to the AoA, and a sudden drop of capacity at ages 14 to 16. Their performance then leveled off to barely above the chance level. Hence, this pattern of age function was regarded as consistent with that of JN89. In the following year, Johnson (1992) conducted a more strict replication of JN89 using the same participants. All other procedures and the materials were identical to those used in the original study. This time, however, a written version of the grammaticality judgment test was used, in order to compare its effects with those of the oral version of the test. In other words, modality was taken into consideration as a potential factor affecting participants' performance. Johnson still found a correlation between AoA and proficiency (r= -.54), although the correlation was weaker than that discovered by the original study (r= -.77). In addition, an analogous pattern of age effect was identified in this replication.

Shim (1993) assessed Korean non-native speakers of English who varied in AoA. She measured the participants' syntactic ability, focusing especially on Subjacency and Anaphoric Binding. The grammaticality judgment test given for assessing proficiency revealed that early learners who were exposed to English prior to the age of five performed within the range of the native controls, and that those whose English acquisition began between the ages of 6 and 11 showed some age effects. Late learners' L2 acquisition was found to occur in an inconsistent manner.

More recently, another replication was performed by DeKeyser (2000). The study was carried out with 57 Hungarian-speaking immigrants in the U.S. using a modified version of the grammaticality judgment test of JN89. In consideration of the possibility that an extensively long test would cause fatigue effect and thereby skew the test result, the test was shortened from 278 to 200 sentences. The findings of the study were notably similar to those of JN89 with a negative correlation again being found between AoA and performance. Further, the language acquisition of the early arrivals was marked by uniform attainment of native-like competence, while that of the late arrivals was featured by a random distribution of scores. A minor divergence was also detected, however, in that DeKeyser's study did not find the declining performance as the subjects' ages approached 15, as had been shown in JN89. Based on this finding, DeKeyser suggested that a prematurational decline might not necessarily be a feature of the CP, as claimed by Johnson and Newport.

In spite of some divergent outcomes, the four studies presented above did generally confirm Johnson and Newport's (1989) findings. All of them noted a clear discrepancy between the performance of early and late arrivals by demonstrating the ceiling effect in the acquisition of

[4] In their 1991 study, Johnson and Newport investigated the universal syntactic principle, *subjacency*, which places restrictions on the types of structures from which wh-movement can take place (see Johnson & Newport, 1991, for more details). By testing the *subjacency* principle, the authors attempted to determine whether or not there is critical period effect for the availability of UG in second language acquisition.

the early arrivals and showing that the proficiency of the later arrivals was marked by performance without a consistent pattern. While the aforementioned studies have supported JN89, there have been other replications that have challenged it.

Flege, Yeni-Komshian, and Liu (1999) investigated 240 Korean native speakers of English who arrived in the U.S. at differing ages ranging from 1 to 24. All of the participants were ensured as being well exposed to English, with more than 15 years of residence in the U.S.. Their knowledge of morphosyntax was measured via a 144-item test, a subset of the JN89 stimuli. The results showed that the ceiling effect was observed among early arrivals (AoA $\leq$ 12), and that a constant decline in performance was found among late arrivals (AoA$\geq$13) . This linear decline can be contrasted with the randomness shown in JN89. Another adverse result was found in the relation between experiential variables (the amount of education and the use of L2) and performance. Although the effect of AoA on the test was statistically significant, it became nonsignificant when other experiential variables, confounded with AoA, were controlled for. Further analyses revealed that the acquisition of rule-based generalizable aspects of morphosyntax (regular past tense on plural formation, third-person singular morphology on present tense verbs, or case assignments on personal pronouns) was mainly affected by the amount of education in the U.S., and that the acquisition of lexically-based irregular and ungeneralizable aspects of morphosyntax depended on participants' use of English. This finding led the authors to conclude that age effect is not maturationally based, but rather results from other variables, mainly amount of education in the U.S. and use of L2.

Another partial replication by Bialystok and Miller (1999) yielded somewhat different findings. Bialystok and Miller examined syntax acquisition by two different non-native groups: Spanish speakers and Chinese speakers, both groups acquiring English. Each non-native participant group was divided into two subgroups of younger learners beginning their acquisition before age 15 and older ones after age 15. To measure their participants' morphosyntactic proficiency, Bialystok and Miller used a grammaticality judgment task which encompassed five grammatical categories: plural, determiner, future tense, present progressive tense, and collocation restriction. These categories were carefully selected on the basis of structural similarities and dissimilarities between the L1s (Chinese and Spanish) and the L2 (English). The results showed that the proficiency patterns were different for the two learner groups. For Spanish speakers, participants whose AoAs were up until the age of eight performed within the range of native speakers, while those whose AoAs were after the age of eight exhibited declining performance throughout the entire sample. In the case of the Chinese speakers, none of them reached native-like achievement, and AoA was found to influence proficiency across all AoAs, rather than being limited to the putative CP. This apparent discrepancy between the two groups led the authors to dispute any role played by maturational constraints in SLA.

A similar pattern to the age function found in Bialystok and Miller's study, was replicated in Birdsong and Molis's study (2001). These researchers examined 61 native speakers of Spanish, with 29 early arrivals (AoA $\leq$ 16) and 32 late arrivals (AoA$\geq$17). After the participants were given the same grammaticality judgment test and language questionnaire (originally used in JN89), their syntactic knoweldge was measured. The findings revealed three identifiable differences from JN89: (a) the existence of a postmaturational effect after the putative offset of the CP; (b) evidence of the native-like attainment of the late arrivals; and (c) the possibility that the outcome of L2 acquisition depends on L1-L2 pairings and L2 use.

The three aforementioned studies found some consistent patterns which can be used to challenge JN89. The first of these, postmaturational effects, is at odds with Johnson and

Newport's argument that if there is a CP, acquisition that has begun after its closure must be marked by a random distribution of scores. In addition, all three studies point to the possiblity that the apparent impact of an age effect on the test score may in fact result from other confounding factors, not from the age of acquisition. Variables such as amount of education in the U.S., use of L1 and L2, and L1-L2 pairings have been pointed to as potential confounding factors causing age differences.

Studies that replicated JN89 have yielded inconsistent findings and thus there is a clear need to conduct further replications of JN89. The present study is conducted as another attempt to replicate JN89. Birdsong and Molis (2001) have attributed the diverse outcomes gained by JN89 replications to procedural differences and/or to variations from the original. Guided by this insight, the current replication attempts to employ procedural and methodological approaches which are as similar as possible to those used in JN89.

THE STUDY

Research Questions

The research questions addressed in this replication are the same as those of JN89:

1. Is there an age-related effect on learning the grammar of a second language? If so, what is the shape of the function relating age to learning and to ultimate performance? In other words, where does the relationship plateau or decline?

2. What areas of grammar are the most, and least, problematic for learners of different age groups?

The first research question is one of the most fundamental issues that need to be answered in order to ascertain the existence of the CP. If the findings of JN89 are replicated, two major outcomes are to be expected: (a) AoA negatively correlates with performance; (b) the particular pattern of age effect reported in JN89 is duplicated with the acquisition starting prior to the AoA of 15 showing that the ceiling effect is followed by a linear decline in performance, whereas the acquisition after the AoA of 15 is marked by a random distribution of performance.

The second question is addresses the relationship between AoA and performance on different grammatical rule categories. As shown in JN89, if there is a CP, it can be predicted that the various rule categories will not affect younger learners' acquisition, but yet cause some differential effects on older learners' acquisition.

Methodology

Participants

The participants of this study consisted of 34 native speakers of Korean. This study's participant composition is different from that of JN89. In their study, Johnson and Newport recruited participants from a mixed group of native Chinese and Korean speakers. Their participant selection was based on the presumption that the speakers of Korean and Chinese can

be categorized into a single linguistic group because both languages are typologically distant from English. In fact, given that Korean and Chinese are marked by their distinctive linguistic features, categorizing the L1 speakers of both languages as a single linguistic group renders an examination of potential L1 effects problematic. Thus, the present study selected participants from one homogenous language background, Korean. In this regard, it must be noted that previous studies have already acknowledged the need to replicate JN89 using participants of one L1 background (Birdsong & Molis, 2001; DeKeyser, 2000).

The two major criteria used in recruiting participants were age of arrival in the U.S., and length of residence. The participants varied from ages three to 34 in their AoAs. Among the 34 participants, 18 were early arrivals who came to the U.S. prior to the age of 15, while 16 were late arrivals who came after the age of 16. Consistent with the minimum criteria of LoR used in JN89, it was ensured that all the participants had resided in the U.S. for at least five years in an uninterrupted stay. The average LoR was 17.6 years for the early-arrival group and 8.9 years for the late-arrival group. In addition, to prevent too much variation resulting from diverse educational backgrounds, all the participants had to have some exposure to formal education in the U.S. In spite of this criterion related to exposure to education in the U.S., however, the participants' levels of education still varied: from an 11th-grader in high school, to professionals holding doctoral degrees.[5]

Instruments

Data in this study were gathered using two instruments: a grammaticality judgment test (GJT) and a language background questionnaire. The tests used to assess the participants' syntactic proficiency was DeKeyser's (2000) modified version of the original GJT used in JN89, hence a partial replication of JN89.

Grammaticality judgment test. The test used to assess the participants' syntactic proficiency was DeKeyser's (2000) modified version of the original GJT used in JN89. Guided by Johnson and Newport's comments and some methodological criticisms of JN89, DeKeyser made a few changes on the original prompts. First, the 12 grammatical rule categories have been modified into 11 rule categories. One category, auxiliary, which was used as an independent rule category in JN89, has been subsumed under the other categories such as present progressive and word order. The 11 rule categories used in the GJT are past, plural, third-person singular, present progressive, determiner, pronominalization, particle movement, subcategorization, yes-no question, WH question, and word order. Second, the 276 sentences of the original test have been cut down to 200. The length of the original test had been criticized for the potential fatigue effect. DeKeyser attempted to avoid this methodological flaw by shortening the number of test items. In addition, DeKeyser deleted or reorganized several subcategories in order to include structures which tended to be problematic for his Hungarian native participants. Despite these modifications, however, it should be noted that most grammatical rule types and test items used in JN89 remained the same (cf. DeKeyser, 2000).

Out of the 200 items from DeKeyser's GJT, 100 items were grammatical and the other 100 were ungrammatical. In the present study, four items, consisting of two grammatical and two ungrammatical sentences, were given as practice items and 196 items were given as the test. The

[4] In JN89, the participants' levels of education varied in a range similar to those seen in the current replication: from college students to professors holding doctoral degrees.

196 items were divided into two parts. The ungrammatical and grammatical sentences of each pair were randomly placed into the opposite halves and then scrambled. The prompts of 200 sentences were recorded by a female native speaker of English. Each item was repeated twice with a 3-second interval. There was a 6-second interval between the items. The complete list of items is given in Appendix A.

Language background questionnaire. A language background questionnaire was designed and administered in order to collect information on each participant's date of birth (age at time of testing), country of origin, native language (L1), age of arrival (AoA) to the U.S., length of residence (LoR) in the U.S., language environment (e.g., years of formal instruction before immigration, years of education in the U.S., and percentage language use of in respect to Korean and English). The questionnaire is presented in Appendix B.

Procedure

The test and the language background questionnaire were administered either in groups, or individually in places such as at church, on campus, or in participants' homes. The participants were asked to first fill out a consent form for both the experimental study and the language questionnaire. The grammaticality judgment test followed the questionnaire. On the GJT, the participants had to respond to the test by marking either *correct* or *incorrect* on each item. Four practice items were given as a trial, and then the actual 196 test items followed. There was a five-minute break after the first 98 items, but the participants could ask to stop at any time. The test lasted for approximately one and a half hours.

Coding/Scoring

Participants' responses on the test were coded and scored using SPSS Version 11. Each response was scored dichotomously as correct/incorrect. Correct responses were coded as '1' and incorrect or missing responses as '0.' All responses for each participant were summed up into a total score. Each participant also received a percentage score which was calculated by dividing the total correct answers by the total number of applicable items.

Statistical Analysis

The data set drawn from the GJT was analyzed in two rounds of analyses. First, the correlation between AoA and overall performance was calculated to look for the possible existence of a significant age effect. Then, a pairwise series of t-tests were run to see if there is any significant difference between adjacent sub-age groups. For the purpose of this analysis, the participants were divided into six groups: the native control group, and five age groups, 3-7, 8-10, 11-15, 16-24, and 25-34. Correlation analyses were also conducted on the dependent variable of the overall performance to see if there were any significant differences in the performance of early arrivals (AoA $\leq$ 15) and late arrivals (AoA$\geq$16). In the second round of analyses, correlations were run to investigate the effects of AoA on the acquisition of 11 grammatical rule types.

RESULTS

Age of Acquisition and Performance on the Test

The first research question concerned the possible existence of a significant age effect on participants' performance. A strong negative correlation between AoA and performance on the grammaticality judgment test was found across the entire AoA span (r = -.848, $p < 0.01$). This strong negative correlation was detected in JN89 (r = -.77, $p < 0.01$) and also in subsequent studies (e.g., Bialystok & Miller, 1999; Birdsong & Molis, 2001; DeKeyser, 2000; Flege et al, 1999; Johnson, 1992; Johnson & Newport, 1991; Shim, 1993). This general finding affirms that younger learners are better than older learners on the GJT measuring the acquisition of major syntactic rules of English. In other words, the acquisition of an L2 gets harder with increasing age.

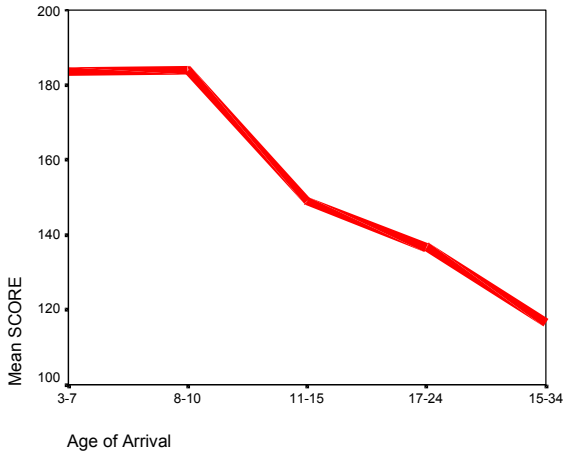
In order to obtain a more detailed understanding of the relationship between AoA and performance, a sub-age-group analysis was conducted. The participants were divided into five age groups: 3-7, 8-10, 11-15, 16-24, and 25-34. Table 1 presents the mean, the standard deviation (SD), and the range of correct items for each group.

TABLE 1
Mean, Standard Deviation, and Range of Correct Items on Grammaticality Judgment Test for Five Age Groups (N=34)

	Age of Arrival (AoA)				
	3-7 (N=7)	8-10 (N=4)	11-15 (N= 6)	16-24 (N=10)	25-34 (N=6)
Mean	183.6	184	149	136.6	116.5
SD	4.79	3.36	19.27	15.95	27.33
Range	190-177	186-179	178-117	151-108	179-87

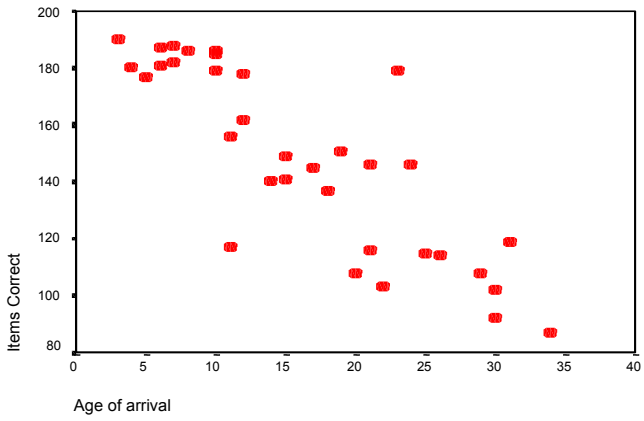
An independent sample t-test was used to see if there were significant differences between adjacent groups. No significant discrepancy was detected between the age groups of 3-7 and 8-10. As can be seen in Table 1, the 3-7 and 8-10 groups were found to be almost identical in both the means and SD ($p = .879$, $p < 0.05$ level). They also overlap in the ranges of correct items. A significant difference was detected, however, between the 8-10 and 11-15 groups ($p = .012$, $p < 0.05$ level). The 11-15 group performed significantly worse than the first two age groups. The range of scores achieved by the 11-15 group did not overlap with that of the 8-10 group. The average mean of performance of the 11-15 group was higher than that of the 16-24 group, and the 16-24 group was higher than that of the 25-34 group. Likewise, the performance of the age groups whose AoA is past age 10 is marked by a linear decline with increasing AoA. Thus, it can be said that it is around AoA 10 that the decline in performance is sharply marked. The relationship between the five different AoA groups and their performances is shown in Figure 1.

FIGURE 1
Different AoA Groups and their Mean Scores on the GJT (N=34)



Another issue of interest concerns the differences between early arrivals (AoA ≤ 15) and late arrivals (AoA≥16). For early arrivals, the correlation between AoA and performance was found to be significant (r= - *0.686*, p < 0.01). The performance of early arrivals was marked by a plateau ceiling effect. By contrast, there is a linear decline of performance in participants whose AoA ranged from 11 to 15. Figure 2 displays the scatterplots of the participants' performance.

FIGURE 2
Scatterplots of Participants' Scores on the GJT (N=34)



The results for early arrivals are seemingly analogous to that of JN89, given that both studies demonstrated a statistically significant correlation between AoA and performance (JN89, r = - *.87,* p < 0.01, vs. the current study, r= - *0.686*, p < 0.01). Note, however, that the results obtained in the present study can be differentiated from that of JN89 in several respects. First, the current study showed that the ceiling effect appears from age three to ten. This does not echo the findings of JN89 in which the ceiling effect ceased at the age of seven. More importantly, in JN89, the linear decline in performance started at the age of 7 and ended at the age of 15. This

declining performance was also marked by statistical significance. According to JN89, the declining performance that marks the gradual offset of the CP should be statistically significant. And yet, while the current study did observe some decline in performance between ages 10 and 15, no statistical significance was found for this decline. Consequently, it is hard to conclude that the declining performance observed in this study indicated any gradual offset of the CP.

As for late arrivals (AoA$\geq$16), the present study found a statistically significant correlation between AoA and performance (r = -.667, $p < 0.01$). This indicates the existence of a postmaturational effect after the putative closure of the CP. This finding can be contrasted with JN89's random distribution of late arrivals' performance.

Another notable pattern for late arrivals in the current study is that their performance was marked by increasing variation. There was virtually no variation among the early arrivals, whereas the variance of scores among the late arrivals became larger as their AoAs increased. As seen in Table 1 and Figure 2, scores were widely distributed among the late arrival groups suggesting more variation among late learners, especially among those whose AoA surpassed the age of ten.

To summarize, both similarities and dissimilarities have been found between the current study and JN89. The two studies are similar in the following three aspects: (a) a strong negative correlation between AoA and performance was found for both groups of participants; (b) the acquisition of early arrivals was marked by the ceiling effect at the level of native-like competence; and (c) more variation among individuals was observed for late arrivals. Nevertheless, the two studies show some discrepancy: a prematurational decline was found only in JN89, and a postmaturational effect only in the current study.

Age of Acquisition and Grammatical Rule Types

The second question raised by this study is that of how AoA affects the acquisition of different grammatical rules. In order to analyze the relationship between AoA and performance with respect to 11 grammatical rules, it was necessary to decide how to quantify participants' performance on each grammatical rule. Johnson and Newport (1989) noted that when a participant marks a grammatical sentence as ungrammatical, one can only guess which grammatical point is problematic to him/her. However, when a participant fails to respond to a targeted grammatical structure by marking an ungrammatical sentence as grammatical, it seems to be more apparent that the participant is missing the targeted grammatical point. Thus, in this study, the ungrammatical sentences were deemed to be a more reliable index of the participants' knowledge. Therefore, following JN89, the current study excluded participants' performance on the grammatical items focusing instead on their performance on the ungrammatical items.

A correlation analysis was conducted so as to find the relationship between AoA and scores on each of the 11 rule types. As may be seen in Table 3, most of the rule types showed significant negative correlations with AoA. These results were interpreted as evidence for the age of exposure to L2 as having an effect on the acquisition of grammatical structures. In other words, participants experience more difficulty in acquiring grammatical rules with increasing AoA.

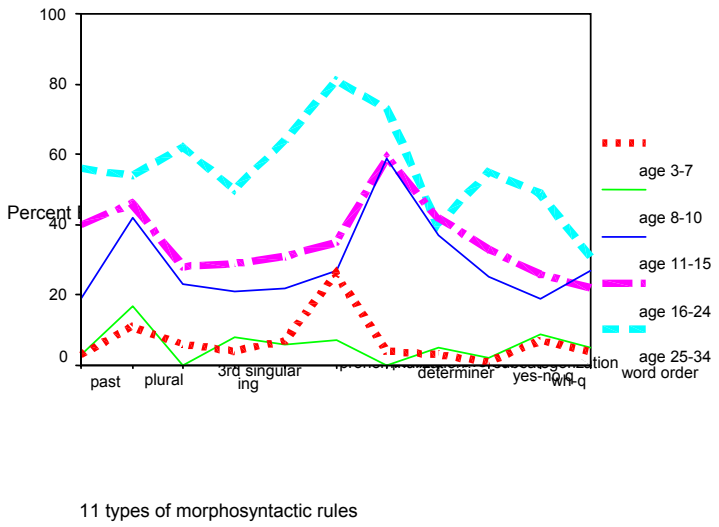
In order to obtain a more detailed error pattern for the11 grammatical rule types, the participants were divided into five age groups: 3-7, 8-10, 11-15, 16-24, and 25-34. Next, the mean percentage of errors for each age group on each rule pattern was calculated. Figure 3 shows the error patterns for the five age groups across the 11 grammatical rules.

TABLE 3
Correlation Coefficients between AoA and 11 Grammatical Rule Types (N=34)

Rule Types	Correlations
Past	.967**
Plural	.963**
Third singular	.911*
Present Progressive	.917**
Pronominalization	.944*
Particle movement	.791
Subcategorization	.957*
Determiner	.894*
Yes-No question	.970**
Wh question	.943*
Word Order	.894*

** Significant at the 0.01 level (p<0.01).
* Significant at the 0.05 level (p<0.05).

FIGURE 3
Mean Percentage of Errors on 11 Grammatical Rule Types (N=34)



What can be inferred from Figure 3 is that the two early age groups (AoA ≤ 10) performed better than the other age groups did. The early age groups' performance showed little difference across the 11 grammatical rules. On the other hand, the three late arrival groups (ages 11-15, 16-24, and 25-34) seem to have made more errors in certain rule types. Subsequent statistical analyses revealed significant correlations with AoA for five rule types: past (r= .811, p < 0.01), determiner (r= .745, p < 0.01), yes-no question (r= .718, p < 0.01), third person singular (r= .674, p < 0.01), and plural (r= .61, p < 0.01). On the other hand, low correlations were found for the other structural types: word order (r= .503, p < 0.01) and particle movement (r=.407, p <

0.01). The five structures with high correlations suggest that these structures are more vulnerable to an age effect. In other words, learners experience more difficulty in acquiring these structures as their age with increases. By contrast, the two structures with low correlations show that these structures are less susceptible to an age effect; consequently, learners can easily acquire them irrespective of their AoA.

Another statistical analysis was conducted to test whether the sub-rule divisions used in the test are a proper way of measuring participants' knowledge on grammatical rule types. Each grammatical rule used in the GJT consists of several sub-rule divisions, as shown in Table 4.

TABLE 4
Grammatical Rule Types and their Subdivisions in the Grammaticality Judgment Test

Past Tense	PTO (Past tense marking omitted) PTIR (Irregular verb regularized) PTII (Regular ending on irregular stem)
Plural	PLO (Plural marking omitted) PLIR (Irregular plural regularized) PLM (Mass noun used with plural markers)
Third-Person Singular	TPSO (Omission in obligatory context) TPSM (marked on main verb after modals)
Present-Progressive	PPMO (Omission of –ing in obligatory context) PPAO (Omission of auxiliary [be verb])
Determiner	DO (Determiners omitted) DAN (Determiners with abstract nouns)
Pronominalization	PROM (Omission in an obligatory context) PRGE (Gender error)
Particle Movement	PMS (Phrasal verb separation not allowed) PMTF (Particle moved too far)
Subcategorization	Sub (Subcategorization)
Yes-No Question	YNA (Misuse of auxiliary) YNDS (Do support) YNDT (Double tense marking)
WH Question	WHNI (No auxiliary inversion) WHNA (No auxiliary (do) supplied)
Word Order	WOSVO (SVO violated) WOAP (Adverb-misplacement)

DeKeyser (2000) pointed to a possible error emerging from the assumption that "all instances of a given structure are equally good tests of a particular structure" (p. 512). To avoid this error, he did not sum up the participants' scores on all the items belonging to the same grammatical rule, but rather analyzed participants' performance on each individual item. Guided by a similar concern, the current study has summed up the scores on items categorized under each sub-rule division. Next, the correlations of the composite scores with the AoAs were computed. It was hypothesized that if the sub-rule divisions under each grammatical rule are

equally good in measuring the knowledge of the given rule, participants' performances will not vary greatly across these divisions.

The correlations between participants' average test scores on each subdivided rule and AoA are listed in Table 5 according to whether they are high ($p < 0.01$), or marginal ($.01 < p < 0.05$), or low ($p \geq 0.05$). The analysis has revealed that participant's performances on sub-rule divisions belonging to the same grammatical rule rarely varied. For instance, sub-rules under past, i.e., PTII, r =-.717, PTIR, r= -.696, PTO, r= -.585; those under determiner, i.e., DAN, r= -.707, DO, r= -.662; those under plural, i.e., PLO, r = -.613, PLIR, r = -.470, PLM, r= -.441. Thus, it was confirmed that the 11 rule categories and their sub-rules are a suitable test for measuring participants' grammatical knowledge.

In sum, the current findings are compatible with that of JN89. In both JN89 and the current study, early arrivals ($AoA \leq 10$) performed within the range of native speakers with respect to virtually all aspects of grammatical rules, whereas late arrivals ($AoA \geq 11$) demonstrated different patterns of difficulty in acquiring different rule types. In addition, in both studies, determiners and plurals were found to be the most difficult structures, while basic word order was found to be relatively easy.

TABLE 5
The Correlations between AoAs and Test Scores on Sub-rule Divisions (N=34)

$p < .01$		$0.01 < p < .05$	$p \geq 0.5$
-.717 PTII	-.589 WOSVO	-.413 WOAP	-. 207 PMS
-.707 DAN	-.585 PTO	-.407 PMTF	
-.696 YNA	-.570 PLIR		
-.689 PTIR	-.567 PROM		
-.662 DO	-.534 YNDS		
-.648 PRGE	-.533 WHNA		
-.646 SUB	-.509 WHNI		
-.624 PPAO	-.490 YNDT		
-.613 PLO	-.470 PPMO		
-.612 TPSO	-.441 PLM		
-.600 TPSM			

*PTO (Past tense marking omitted), PTIR (Irregular verb regularized), PTII (Regular ending on irregular stem), PLO (Plural marking omitted), PLIR (Irregular plural regularized), PLM (Mass noun used with plural markers), TPSO (Omission in obligatory context), TPSM (Marked on main verb after modals), PPMO (Omission of –ing in obligatory context), PPAO (Omission of auxiliary (be verb)), DO (Determiners omitted) , DAN (Determiners with abstract nouns), PROM (Omission in an obligatory context), PRGE (Gender error), PMS (Phrasal verb separation not allowed), PMTF (Particle moved too far), Sub (Subcategorization),YNA (Misuse of auxiliary) ,YNDS (Do support), YNDT (Double tense marking), WHNI (No auxiliary inversion), WHNA (No auxiliary (do) supplied), WOSVO (SVO violated),WOAP (Adverb-misplacement)

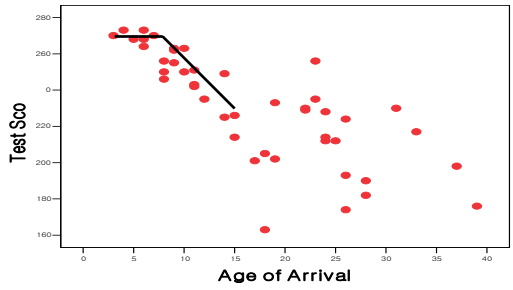
DISCUSSION

Age of Acquisition and Performance

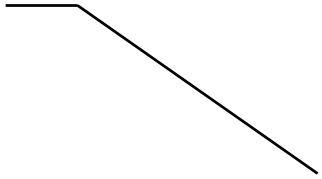
The present study confirms that AoA plays a crucial role in determining L2 learners' proficiency by showing a statistically significant negative correlation between AoA and performance (r = -*.848*, $p < 0.01$). This finding, consistent with a number of previous studies, confirmed the general belief that younger learners do better than older learners in the acquisition of an L2.

A primary concern of the study has been to explore the relationship between AoA and performance. In JN89, early arrivals' acquisition showed the ceiling effect at the level of native-like competence (3≤ AoA ≤7), and a linear decline in performance (8 ≤ AoA ≤ 15). As for the late arrivals, their acquisition showed no consistent pattern due to the age effect. According to Johnson and Newport (1989), the pattern of the age function in their study indicated that "age effect is present during a time of ongoing biological and cognitive maturation and absent after maturation is complete" (p. 90). Hence, the researchers' conclusion was that this shape of the relationship between AoA and performance ought to show up as evidence of maturational effects. In the present data, however, the ceiling effect shown for the participants whose AoA was between ages three and ten was followed by a linear decline in performance for participants whose AoA exceeded age ten. Thus, the shape of the age function in the present data was different from that seen in JN89. Figure 4 presents the age functions of JN89 and of the present study, respectively.

FIGURE 4
The Age Functions of JN89 and of the Present Study



JN89



The Present Study

The point that arises at this juncture is whether this diverging feature, observed in the current study, constitutes counterevidence against the CP. Certainly, some studies that observed a similar pattern have interpreted it as counterevidence that supported their position against the CP (e.g., Bialystok & Miller, 1999; Birdsong & Molis, 2001; Flege et al., 1999). In this section, in order to answer this question, major structural features of the age function that have been detected in the current data are scrutinized, and further mined for their potential meanings.

The Ceiling Effect among Early Arrivals

One key structural feature detected in both JN89 and the present study is the ceiling effect among early arrivals. As is shown in Figure 4, JN89 reported that it extends from age three to eight, while the current study sees it as stretching from age three to ten. The ceiling effect in both studies emerged due to the fact that, during a particular AoA span, virtually all participants achieved uniform mastery of L2 knowledge. Most CP studies, whether they are in favor of or against the CP, have reported this ceiling effect (e.g., between ages 3 and 15, Birdsong & Molis [2001]; between ages 3 and 15, DeKeyser [2000]; between ages 2 and 12, Flege et al. [1999]; between ages 5 and 15, Patkowski [1980]). In spite of its apparent presence, however, its causal

factors have yet to be adequately explored. One of the most plausible among a number of proposals as to what causes ceiling effect has been the CP account. According to the CP account, due to the heightened sensitivity available within the CP, the successful acquisition of a target behavior is warranted regardless of how variable the exterior factors are (Bornstein, 1989; Colombo, 1982). Advocates of the CP firmly believe that early arrivals, as evidenced by their uniform achievement of native-like competence, benefit from this peak sensitivity. Thus, the ceiling effect shown in the current study could be used to support the existence of the CP.

Absence of a Prematurational Decline among Early Arrivals

The first discrepancy between JN89 and the present study was found in the absence of a prematurational decline. Johnson and Newport (1989) asserted that if there is a CP, "there should be a consistent decline in performance over age for those exposed to the language before puberty" (p. 70). Accordingly, for Johnson and Newport, a prematurational decline appearing before the end of the CP is an inevitable feature of the CP. While the finding of JN89 confirmed this assertion, the data from the present study did not do so in that there was no statistically significant decline in performance prior to puberty.

It is still an issue of controversy, however, whether a prematurational decline is indeed an inevitable structural feature of the CP. In the original Critical Period Hypothesis (CPH), Lenneberg (1967) posited that the heightened sensitivity drops off abruptly around puberty. This indicates that there is no transitional phase during which sensitivity gradually declines. In other words, the original CPH did not stipulate a prematurational decline. This abruptness and rigidity of the original CPH, however, has become an unpopular notion among many researchers. Instead, the gradual decline of maturational sensitivity has become a more appealing idea. Oyama (1979) suggested that the terms "biological" and "maturational" have tended to be too narrowly defined in their early usage. According to her, a CP could be more certainly thought of as a product of a nature-nurture interaction; accordingly, its gradual nature would have to be taken into consideration. More recently, Newport (1990) posited that the offset of the CP is not abrupt. Rather she sees it as a curvilinear function which clearly marks the peak sensitivity to learning and the decrease of this ability to some plateau. Similarly, Birdsong (in press) has defined the CP as including an abrupt onset or increase of sensitivity, a plateau of peak sensitivity, and a gradual offset or decline. By stressing the fact that the span of a CP begins at the moment when sensitivity starts to increase and ends at the point where sensitivity is absent, Birdsong claims that a critical period includes "all heightened sensitivity, the transition as well as the uppermost level" (p. 3).

Likewise, while the recent literature seems to favor the idea of a prematurational decline, this appealing notion remains speculative. No researcher has provided any convincing theoretical foundation for this proposal (DeKeyser, 2000). In addition, few studies, even those penned by the researchers in favor of the CP, have provided clear evidence of a prematurational decline. This issue, therefore, remains open to further theoretical and empirical inquiry.

Discontinuity Point of the Age Function

Another discrepancy between JN89 and the present study is found in the different cutoff points of the age function: age 15 in JN89, and age 10 in the present study. It has been assumed that if there is a CP, its endpoint can be defined as the discontinuity point of the age function

(Bialystok & Miller, 1999). In other words, if the age-related decline is caused by maturational changes, an identifiable change, one that is reflected in a certain type of discontinuity of the age function, should be observable around the closure of the CP. In the present data, a clear discontinuity of the age function was found around age ten, yet the finding conflicts with most studies supporting the CP.

In the original CPH, Lenneberg suggested the cutoff point of the CP as puberty. This assumption has been generally accepted and also empirically supported. The clearest demonstrations were made by Johnson and Newport (1989) and by Patkowski (1980). Both studies showed that when the samples were divided into two AoA groups, with age 15 as the dividing line, there was a different pattern for the pre-15 and for post-15 group. In addition, DeKeyser (2000) reported a clear discontinuity of the age function at age 15.

Nevertheless, it has been suggested that the cutoff point of the CP may vary to some extent. As noted above, there has been a growing acceptance of the idea that language acquisition is open to external stimuli and varies considerably among individuals. Such a notion has been called for given that the variations detected across individuals and across groups. As a matter of fact, some studies have reported varying cutoff points (e.g., age of 11 in Shim [1993] and age of 12 in Flege et al. [1999]). In light of these diverse research findings, the deviation in the present data ─ the discontinuity point at age ten ─ does not appear to be problematic, yet it is left for further scrutiny. More importantly, the possibility cannot be ruled out that this early discontinuity is artifactual given the limited number of participants. If so, there is a need to verify the current finding with a larger population of participants.

Postmaturational Effect among Late Arrivals

The last structural feature worthy of attention is the postmaturational effect observed in the current study. In general, it has been argued that if there is a CP, no systematic age-effect should be detected after the putative closure of the CP. Instead, a random distribution of late arrivals' performance is generally regarded as an evidence of the CP. Researchers who have written in favor of the CP thus have adopted this pattern to support the CP (e.g., DeKeyser, 2000; Johnson, 1991; Johnson & Newport, 1989, 1991; Shim, 1993). On the other hand, any type of consistent age effect, frequently referred to as a "postmaturational effect," was deemed a strong piece of counterevidence against the CP (e.g., Bialystok & Miller, 1999; Birdsong & Molis, 2001). In this regard, the postmaturational effect shown in the current study appears to disfavor the CP.

It must be noted, however, that the presence of a postmaturational effects does not on its own prove that there is no CP. It is generally believed that if there is a CP, sensitivity should be available only within the CP. A postmaturational effect which reflects sensitivity still available outside of the CP, therefore, has been used to refute the CP. It should be carefully noted, though, that in order to be used as counterevidence against the CP, a postmaturational effect must mirror the decline of maturational or biological sensitivity, and not any other type of decline. If a postmaturational effect is a true reflection of diminishing maturational sensitivity, then AoA must be the only variable which can determine a linear decline of performance among late learners. If other factors beyond AoA are causing that decline, a postmaturational effect may not be deemed as an indicator of decreasing maturational sensitivity, and as a result not be available to counter the existence of the CP.

In this regard, the empirical findings of the present study were revealing. They demonstrated that the postmaturational decline found in the study resulted not from AoA, but rather from the use of L1 and L2. First, the correlations between the experiential variables and the score (use of Korean, $r = .787$, $p < 0.05$; use of English, $r = -.787$, $p < 0.05$) were higher than those between the AoA and the score ($r = -.618$, $p < 0.05$). Further, the partial correlation analysis also confirmed this finding. When AoA was partialed out, the correlation between performance and each variable remained robust (i.e., use of Korean, $r = -.578$; use of English, $r = .407$). However, when each experiential variable was partialed out, the correlation of the AoA and the score became nonsignificant. Given that the present study has identified two experiential variables, use of Korean and use of English as being more plausible factors in causing a decline in performance after the putative endpoint of the CP, it could be concluded that the postmaturational effect found in the present study does not reflect a maturational decline, and hence cannot constitute counterevidence against the CP. There has been a prevalent but probably unsupported presumption that the presence of a postmaturational effect reflects its declining maturational sensitivity. No such assumption was supported by this study, however. The study's findings highlight the need for the postmaturational effects to be carefully examined in light of the causal factors, focusing especially on whether they indeed result from declining maturational sensitivity, or rather from other factors.

To summarize, the observation of the ceiling effect among early arrivals appears to work in favor of the existence of the CP. Still, other structural features such as a prematurational decline, the discontinuity point of the age function, and the postmaturational effect, could be used either for or against the CP, due to its insufficient theoretical and empirical foundations. Given this study's mixed findings on the relationship between AoA and performance, it seems to be too early to arrive at any firm conclusion as to the existence of the CP.

Age of Acquisition and Grammatical Rule Types

Different age effects were detected between early and late arrivals with respect to their acquisition of 11 grammatical rule types. Younger arrivals' performance was not susceptible to any particular type of linguistic structure; that is, they experienced few problems in mastering virtually all grammatical rules. By contrast, in the case of late arrivals notably different patterns of difficulty were found with respect to grammatical rule type. More specifically, the five grammatical structures ─ past, plural, determiner, yes-no questions, and third singular ─ were found to be most problematic, whereas the two grammatical structures ─ basic word order and particle movement ─ were found least problematic.

A possible causal factor for the different patterns of difficulty may be the interference from the participants' first language. The four English (L2) structures that were found difficult in this study do not have any corresponding structures in Korean (L1). Thus, it can be reasonably assumed that it is the absence of similar target structures in the L1 which brings about this difficulty. That the difficulty will emerge from the dissimilarity between L1 and L2 has already been predicted by several researchers. Kellerman (1995) explicitly formulated a hypothesis as to an interaction among L1 features, L2 features, and AoA. According to him, learners attempting to acquire certain features of an L2 which do not have L1 equivalents must have acquired those features before the CP ends; otherwise, they will never acquire them. By contrast, features of L2 having L1 analogues can in principle be mastered at any stage, regardless of the age of onset of learning. Similarly, yet more specifically, Bialystok and Miller (1999) posited that if there is a

CP, there should be different L1 effects for younger and for older learners. Younger L2 learners within the CP can have access to both the language-specific mechanism and L1 knowledge, and hence they can acquire any L2 linguistic feature and minimize the impact of the L1 on their ultimate attainment. By contrast, older L2 learners outside of the CP are left open to the influence of L1 linguistic structures, since they have access only to L1 knowledge and the general learning mechanism. Therefore, older learners' L2 construction will depend on the structural similarity, or dissimilarity, of the L1 and L2. With respect to L2 features having L1 equivalents, older learners may be able to acquire them at any age; when it comes to L2 features having no L1 equivalents, however, they can achieve only a limited success at best. What has been generally predicted by both Kellerman (1995) and Bialystok and Miller (1999), was illustrated by the current study: certain dissimilarity effect between L1 and L2 in the acquisition of late arrivals.

Another plausible explanation for the different patterns of difficulty was put forth by DeKeyser (2000). According to him, the saliency of grammatical items plays a role in determining ease or difficulty of items in the acquisition of older learners. Since older learners begin their acquisition after the putative closure of the CP, they count on their cognitive mechanism for the acquisition of syntax; this makes these learners more dependent on their explicit, analytic, problem-solving capacities. As a result, older learner first acquire the more salient patterns (e.g., basic word order), then the less salient ones (e.g., determiners, plurals). DeKeyser's view can be summed up as follows: younger learners, who benefit from their biological superiority, can easily master any linguistic structure, whereas older learners, who rely on their cognitive ability, can only achieve a limited mastery of L2 linguistic rules.

What should be noted with respect to the two accounts cited above is that both suggest a possible role for a cognitive mechanism in late starters' acquisition. Bialystok and Miller (1999) proposed that the manifest influence of L1 transfer on late arrivals' acquisition, which does not appear in early arrivals' acquisition, can be ascribed to a general cognitive mechanism. Similarly, DeKeyser's idea of a saliency effect is rooted in the Fundamental Difference Hypothesis (FDH), which proposes that an implicit language-learning mechanism used by children is converted into an explicit cognitive problem-solving ability in adults' acquisition.

In sum, it can be said that no consensus has been reached as to what causes different patterns of difficulty prevalent among late arrivals. The two most plausible accounts have been reviewed in this study: L1 transfer and perceptual saliency. Both accounts suggest a possible involvement of cognitive mechanism in older learners' acquisition. Even this suggested involvement remains speculative, however, given the lack of any substantial proof. What is apparent at this point is, first, that two distinct patterns of performance ─ uniform success among early arrivals, and diverse outcomes among late arrivals ─ speak to the involvement of different mechanisms in the acquisition of early and late arrivals; and second, that early arrivals' uniform success can be best explained via the involvement of an innate mechanism linked to biological superiority. The findings of the present study on the relationship between AoA and grammatical rule types can, therefore, be considered to constitute another supporting argument for the CP.

CONCLUSION

The present study was conducted as an attempt to partially replicate JN89 which has been regarded as the landmark investigation of the CP in the acquisition of L2 syntax. Although not all findings of this study were in complete agreement with those of JN89, the present study has nonetheless provided some modest support for the existence of the CP. The evidence for the CP that this study has gleaned can be summed up as follows:

1. A clear ceiling effect was found in the acquisition of early arrivals ($3 \leq$ AoA ≤ 10); participants' uniform success was observed within this particular period.

2. There were found to be different patterns of grammar acquisition for the early arrivals and for the late arrivals. The early arrivals were uniformly successful in acquiring virtually all grammatical rules; by contrast, the late arrivals showed a different range of mastery across the 11 grammatical rule types.

In addition to the above findings in favor of the CP, the study was also able to identify some of the unresolved issues which deserve the attention of future CP researchers. First of all, the structural characteristics of the CP have yet to be clarified. In particular, the idea of a prematurational decline as a structural feature constituting the CP needs to be substantiated both theoretically and empirically. Second, the postmaturational effect, which has been cited as one of the most convincing pieces of counterevidence against the CP, needs to be further scrutinized in terms of its causal factors. The literature to date appears to have given most of its attention to the presence of a postmaturational effect, and yet what brings about this effect has never been made clear. Third, the study has addressed the possibility that different mechanisms may be involved in the acquisitions of the early arrivals and the late arrivals ─ possibly a biologically innate mechanism for the early arrivals and a cognitive mechanism for the late arrivals. Identifying the role of a different mechanism governing the late arrivals' acquisition might bring us to a better understanding of the role played by the CP in the acquisition of early starters. Particularly promising for future CP research would be further investigations of the FDH addressed in the recent replication by DeKeyser (2000) and of the language transfer effect addressed by Kellerman (1995) and by Bialystok and Miller (1999).

The present study has also revealed some methodological limitations. The first is the use of length of residence (LoR) as a major criterion for recruiting participants. Following JN89, the current study used five-year LoR in order to select participants who had reached their final stage of acquisition. The underlying assumption here is that a certain length of residence in a target L2 country will guarantee enough exposure to L2 input whereby L2 learners fully develop their L2 and reach their final stage of L2 acquisition. This assumption, however, has been regarded as highly arbitrary and problematic (Birdsong, in press). In the absence of a more detailed scrutiny of the learners' acquisition processes and linguistic environments, it is hard to determine whether learners have reached their final stage of L2 acquisition. Thus, a methodological refinement is unequivocally necessary for any type of CP investigation that aims to examine L2 learners' ultimate attainment. Second, a small number of the participants (n=34) proved to be problematic to the point of rendering the results of this study less generalizable. This also made the statistical analyses less reliable, due to the artifactual effects commonly associated with a smaller sample

size. The third and final limitation stems from the interpretation of L1 transfer. It has been suggested that L1 transfer may be a possible way of accounting for the different patterns of performance of early and late arrivals. This interpretation of L1 transfer needs to be treated with caution, however, given the study only examined participants from one L1 background, Korean. In order to gain a more reliable measurement of L1 transfer effects, it would be best to compare more than two participant groups coming from different L1 backgrounds.

The great contributions made by JN89 to CP research in the field of SLA cannot be underestimated. It has advanced the CP investigation both theoretically and empirically. Nevertheless, the findings of the current partial replication underscore the need to reanalyze JN89 in terms of both its theoretical assumptions and its methodological procedures, and to address a number of lingering or newly-surfaced controversial issues that have been pinpointed by both JN89 and the followers. Hence, further replications of JN89 need to be conducted to bring a better understanding of the issues relating to the CP in SLA. It is hoped that both the theoretical issues and the methodological limitations brought to light by this partial replication will serve as a good foundation for future investigations.

ACKNOWLEDGEMENTS

I would like to thank Dr. Robert DeKeyser for letting me use his grammatical judgment test in this paper. Thanks also go to Dr. ZhaoHong Han, Dr. James Purpura, and the anonymous reviewers for providing helpful comments on earlier versions of this paper.

REFERENCES

Asher, J., & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal, 53,* 334-341.

Bialystok, E., & Hakuta, K. (1994). *In other words: In science and psychology of second language acquisition.* New York: Basic Books.

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age difference for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 161-181). Mahwah, NJ: Erlbaum.

Bialystok, E., & Miller, B. (1999). The problem of age in second language acquisition: Influences from language, task, and structure. *Bilingualism: Language and Cognition, 2,* 127-145.

Birdsong, D. (in press). Second language acquisition and ultimate attainment. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics.* London: Blackwell.

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., Summeren C. V., Planken, B., & Schils, E. (1995). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19*, 447-465.

Bornstein, M. H. (1989). Sensitive periods in development: Structural characteristics and causal interpretations. *Psychology Bulletin, 105,* 179-197.

Colombo, J. (1982). The critical period concept: Research, methodology, and theoretical issues.

Psychological Bulletin, 91, 260- 275.

Curtiss, S. R. (1977). *Genie: A psycho-linguistic study of a modern day "wild child."* New York: Academic Press.

Curtiss, S. R. (1988). The special talent of grammar acquisition. In L. K. Obler & D. Fein (Eds.), *The exceptional brain: Neuropsychology of talent and special abilities* (pp. 362-386). New York: Guilford.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, 499-534.

Flege, J. E. (1987). A critical period for learning to pronounce foreign languages? *Applied Linguistics, 8,* 162-177.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second language acquisition. *Journal of Memory and Language, 41,* 78-104.

Harley, B., & Wang, W. (1997). The Critical Period Hypothesis: Where are we now? In A. De Groot, & J. Kroll (Eds.), *Tutorials in Bilingualism* (pp. 19-51). Mahwah, NJ: Lawrence Earlbaum.

Ioup, G. (1984). Is there a structural foreign accent?: A compassion of syntactic and phonological errors in second language acquisition. *Language Learning, 34,* 1-17.

Johnson, J. S. (1992). Critical period effects in second language acquisition: The effect of written versus auditory materials on the assessment of grammatical competence. *Language Learning, 42,* 217-248.

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

Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: the status of Subjacency in the acquisition of a second language. *Cognition, 39,* 215- 258.

Kellerman, E. (1995). Age before beauty: Johnson and Newport revisited. In L. Eubank, L. Selinker & M. Sharwood Smith (Eds.), *The current state of Interlanguage: Studies in honor of William Rutherford (*pp. 244-264). Cambridge, UK: Cambridge University Press.

Lenneberg, E. H. (1967). *Biological foundations of language.* New York: Wiley.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called a grammatic aphasics. *Cognition, 13*, 361-392.

Neufeld, G. (1977). Language learning ability in adults: A study on the acquisition of prosodic and articulatory features. *Working Papers on Bilingualism, 12,* 45-60.

Neufeld, G. (1979). Towards a theory of language learning ability. *Language Learning*, *29,* 227-241.

Newport, E. L. (1984). Constraints on learning: Studies in the acquisition of American sign language. *Papers on Reports on Child Language Development, 23,* 1-22.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14,* 11-28.

Newport, E. L., & Supalla, T. (1987). *A critical period effect in the acquisition of a primary language.* Unpublished manuscript, University of Illinois.

Oyama, S. C. (1979). The concept of the sensitive period in developmental studies. *Merrill-Palmer Quarterly, 25,* 83-103.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language acquisition. *Language Learning, 30,* 449-72.

Schumann, J. H. (1978). Social and psycholinguistic factors in second language acquisition. In J. Richards (Ed.), *Understanding second and foreign language acquisition* (pp. 163-178). Rowley, MA: Newbury House.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning, 19,* 245-253.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human research.* Cambridge, MA: Newbury House.

Seliger, H. (1978). Implications of a multiple critical periods hypothesis for second language learning. In W. Ritchie (Ed.), *Second language acquisition research: Issues and implications* (pp. 11-19). New York: Academic Press.

Seliger, H. W., Krashen, S.D., & Ladeforged, P. (1975). Maturational constraints in the acquisition of second language acquisition. *Language Science, 36,* 20-22.

Shim, R. (1993). Sensitive periods for second language acquisition: A reaction-time study of Korean-English bilinguals. *Ideal, 6,* 43-64

Snow, C. M., & Hoefnagel-Hohle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development, 49*, 1114-1128.

Tahta, S., Wood, M., & Loewenthal, K. (1981). Foreign accents: Factors relating to transfer of accent from the first language to a second language. *Language and Speech, 24,* 265-272.

White, L., & Genesee, F. (1996). How native is near-native?: The issue of ultimate attainment in adult second language acquisition. *Second Language Research, 12,* 238-265.

APPENDIX A

Grammaticality Judgment Test

Practice Items

1. A snake bit she on the leg.*
2. Susan is making some cookies for the oak tree.
3. The baby bird has fall from the oak tree.
4. The little boy was counting all his pennies last night. *

Real Test Items

First set of 98 Questions

1. Last night the old lady die in her sleep.*
2. Last night the books falled off the shelves.*
3. Mr. Murphy hidded his money under his mattress.*
4. Yesterday the teacher sent Allison to the principal.
5. Sandy filled a jar with cookies last night.
6. John sing for the church choir yesterday.*
7. Three boy played on the swings in the park.*
8. Janie slept with her teddy bear last night.
9. A policeman gave Alan a ticket for speeding yesterday.

10. A bat flewed into our attic last night.*

11. Many houses were destroyed by the flood last week.
12. The farmer bought two pig at the market.*
13. A shoe salesman sees many feet throughout the day.
14. The boy lost two teeth in the fight.
15. I need to get some information about the train schedule.
16. Our neighbor bought new furnitures last week.*
17. Two mouses ran into the house this morning.*
18. Teachers often give useful advice to their students.
19. John's dog always waits for him at the corner.
20. Mrs. Sampson clean her house every Wednesday.*

21. John can plays the piano very well.*
22. Mary will goes to Europe next year.*
23. Every Friday our neighbor washes her car.
24. Our new neighbor should turn his radio down a bit.
25. The boy has been lie to his father.*
26. The little boy is speak to a policeman.*
27. Janet is wearing the dress I gave her.
28. Tom working in his office right now.*
29. The Johnsons may be moving to Chicago this fall.
30. Mrs. Newport will is leaving the party early.*

31. Red is a beautiful color.
32. Peter made out the check but didn't sign.*
33. Tom is reading a book in the bathtub.
34. Bob trying to fix Jim's car with his new tools.*
35. Mrs. Johnson went to library yesterday.*
36. The children are playing in the garden till dark these days.*
37. The boy is helping the man build a house.
38. The men played the basketball in the backyard.*
39. Beauty is something that lasts forever.
40. After a life like that he will go straight to the hell.*

41. John took a sweater along but didn't put on.*
42. Mike wrote the letter but didn't send it.
43. The girl cut herself on a piece of glass.
44. The drunk slept his hangover off in the guest room.*
45. Peter did not have any money on her.*
46. Mary fell but she did not break any bones.
47. John knew but she did not tell.*
48. Mary looked at the flowers but didn't buy them.
49. The man climbed up the ladder carefully.
50. The new neighbors carried on a long conversation.

51. I want you to go to the store now.
52. The little boys laughed the clown.*
53. The man looked the new cars over yesterday.
54. This plastic gives a weird smell off.*
55. She broke her shoes very carefully in.*
56. Kevin called Nancy for a date up.*
57. John told me that his wife was ill.
58. Mary took her coat off quickly.
59. George says his prayers much too softly.
60. The student was learning in his room until late last night.*

61. The man allows his son to watch TV.
62. The man lets his son to watch TV.*
63. The girls want to watch TV.
64. Will be Harry blamed for the accident?*
65. The girls enjoy to watch TV.*
66. Is being the baby held by his mother?*
67. Has the King been served his dinner?
68. I hope you to go to the store now.*
69. Can ride the little girl a bicycle?
70. Will Harry wear his new shirt to the party?

71. Does John know the answer to that question?
72. Does Martha use her microwave oven?
73. Where Ted is working this summer?*
74. Did Bill dance at the party last night?
75. When will Sam fix his car?
76. Swam Janet in the race yesterday?*
77. Is waiting Sally in the car?*
78. Where did Arnie hunt last year?
79. Did Bobbie stayed at home last night?*
80. What is Martha bringing to the party?

81. What they sell at the corner store?*
82. The boy feeds the rabbits carrots.
83. The girl the movie likes.*
84. The dog bites.
85. When do they leave for Mexico?
86. The dinner the man burned.*
87. The boy caught the ball.
88. The woman the policeman asked a question.*
89. Who do you meet at the park every day?
90. Linda a cake baked John.*

91. The woman paints.
92. My neighbor enjoyed slowly his dessert.*

93. All our friends in the suburbs live.*
94. The student eats quickly his meals.*
95. The children play with the dog.
96. Drinks the man.*
97. Kevin usually rides his bicycle to work.
98. The students to the movies went.*

Second set of 98 Questions

1. My neighbor slowly enjoyed his dessert.
2. The student eats his meals quickly.
3. All our friends live in the suburbs.
4. Kevin rides his bicycle usually to work.*
5. The children with the dog play.*
6. The students went to the movies.
7. The man drinks.
8. Paints the woman.*
9. Bites the dogs.*
10. Linda baked John a cake.

11. The woman asked the policeman a question.
12. When Sam will fix his car?*
13. The girl likes the movie.
14. The man burned the dinner.
15. When they leave for Mexico?*
16. The boy carrots feeds the rabbits.*
17. What do they sell at the corner store?
18. Who you meet at the park every day?*
19. The ball the boy caught.*
20. Where is Ted working his summer?

21. What Martha is brining to the party*?
22. Can the little girl ride a bicycle?
23. Did Bobbie stay at home last night?
24. Danced Bill at the party last night?*
25. Did Janet swim in the race yesterday?
26. Knows John the answer to that question?*
27. Where did Arnie hunted last year?*
28. Does Martha uses her microwave oven?*
29. Is Sally waiting in the car?
30. Will wear Harry his new shirt to the party?*

31. The girls want watching TV.*
32. The man lets his son watch TV.
33. Has been the King served his dinner?*
34. The man allows his son watch TV.*

35. I hope you will go to the store now.
36. Is the baby being held by his mother?
37. I want you will go to the store now.*
38. Will Harry be blamed for the accident?
39. The girls enjoy watching TV.
40. The student was studying in his room until late last night.

41. The little boys laughed at the clown.
42. The new neighbors carried a long conversation on.*
43. George says much to softly.*
44. Mary took her coat quickly off.*
45. She broke her shoes in very carefully.
46. Kevin called Nancy up for a date.
47. This plastic gives off a weird smell.
48. The man looked the new cars yesterday over.*
49. The drunk slept of his hangover off in the guest room.
50. John said me that his wife was ill.*

51. The man climbed the ladder up carefully.*
52. John knew but he did not tell.
53. Peter made out the check but didn't sign it.
54. Peter did not have any money on him.
55. The girl cut himself on a piece of glass.*
56. Mary feel but he did not break any bones.*
57. Mike wrote the letter but didn't send.*
58. Mary looked at the flowers but didn't buy.*
59. John took a sweater along but didn't put it on.
60. The men played basketball in the backyard.

61. The boy is helping the man build house.*
62. Mrs. Johnson went to the library yesterday.
63. After a life like that he will go straight to hell.
64. The beauty is something that lasts forever.*
65. Tom is reading book in the bathtub.*
66. Bob is trying to fix Jim's car with his new tools.
67. The children playing in the garden till dark these days.*
68. Tom is working in his office right now.
69. The red is a beautiful color.*
70. The boy has been lying to his father.

71. Janet is wear the dress I gave her.*
72. Our new neighbor should turns his radio down a bit*
73. Mrs . Newport will be leaving the party early.
74. The Johnsons may are moving the Chicago this fall.*
75. The little boy is speaking to a policeman.
76. Mary will go to Europe next year.

77. John can play the piano very well.
78. Mrs. Sampson cleans her house every Wednesday.
79. John's dog always wait for him at the corner.*
80. Everyday Friday our neighbor wash her car.*

81. Our neighbor bought new furniture last week.
82. The boy lost two teeths in the fight.*
83. A shoe salesman sees many foots throughout the day.*
84. The farmer bought two pigs at the market.
85. Two mice ran into the house this morning.
86. Teachers often give useful advices to their students.*
87. Mr. Murphy hid his money under his mattress.
88. Many house were destroyed by the flood last week.*
89. I need to get some informations about the train schedule.*
90. Three boys played on the swings in the park.

91. Last might the books feel off the shelves.
92. A bat flew into our attic last night.
93. Janie sleeped with her teddy bear last night.*
94. John sang for the church choir yesterday.
95. Yesterday the teacher sented Allison to the principal.*
96. Sandy fill a jar with cookies last night.*
97. A policeman gived Alan a ticket for speeding yesterday.*
98. Last night the old lady died in her sleep.

APPENDIX B

Language Background Questionnaire

Thank you very much for participating in this study. This language background questionnaire contains questions concerning your experience with both Korean and English. I deeply appreciate your cooperation in filling out the questionnaire as accurately as possible. The information gathered through this survey will be crucial to the analysis of the data obtained from the experiment.

1. What is your date of birth? (e.g. month/date/year: 04/12/73) ____/_____/_____

2. What is your gender? (Please circle) Female/ Male

3. What is your native language? _____

4. What is your parents' native language? _____

5. How old are you now? (e.g. 19 years 2 months) _____years ____months

6. When did you come to the United States? (please write the year and the month of your arrival to the best of your knowledge)

_____

7. How old were you when you came to the United States?

Years Months

8. How long did you stay in the United States?

Years Months

9. What is your highest education? (Please circle)
High school/ Undergraduate/ Graduate

10. What language do you usually use at home? English/ Korean

11.

How much English do you use when talking to your parents?	How much Korean do you use when Talking to your parents?
_____ 100 % English	_____ 100 % Korean
_____ 75 % English	_____ 75 % Korean
_____ 50 % English	_____ 50 % Korean
_____ 25 % English	_____ 25 % Korean
_____ 0 % English	_____ 0% Korean

12.

How much English do you use when talking to your siblings?	How much Korean do you use when talking to your siblings?
_____ 100 % English	_____ 100 % Korean
_____ 75 % English	_____ 75 % Korean
_____ 50 % English	_____ 50 % Korean
_____ 25 % English	_____ 25 % Korean
_____ 0 % English	_____ 0 % Korean

13.

How much English do you use when talking to your friends?	How much Korean do you use when talking to your friends ?
_____ 100 % English	_____ 100 % Korean
_____ 75 % English	_____ 75 % Korean
_____ 50 % English	_____ 50 % Korean
_____ 25 % English	_____ 25 % Korean
_____ 0 % English	_____ 0 % Korean

14. Which language do you think you are more comfortable with?
    i. At home: Korean/ English
    ii. At work: Korean/English
    iii. With friends: Korean/ English

15. Did you learn English before coming to the United States? Yes, No

    If yes, for how long? _____ Years _____ Months

    If yes, how did you learn it?
    i. School or Institute
    ii. Private Tutoring
    iii. Through Media (e.g., TV or Radio)
    iv. Short-term language training abroad
    v. Others _____

16. Did you go to any English classes after you came to the United States? Yes/No

    If yes, for how long? _____ Years _____ Months