

“Age” effects on Korean adults’ learning of English morphosyntax^a

James Emil Flege

School of Health Professions
University of Alabama at Birmingham

University of Illinois, Champaign IL May 6, 2010

Second annual Second Language Acquisition and
Teacher Education (SLATE)

In honor of Professor Molly Ann Mack

^aIn this revised SLATE talk I have replaced the presentation of phonetic data (available in Flege, 2007) with details of the Flege et al. (1999) study that could not be included in the original *JML* publication due to page limits. This research presented here was supported by NIH grant DC02892..

Aim of this talk

Research examining second language acquisition (SLA) has repeatedly shown that earlier is better with respect to naturalistic L2 learning.

The research I will present today examined 240 Korean living in the Washington, D.C. area. Flege, Yeni-Komshian & Liu (1999) administered a grammaticality judgment test (GJT) to these immigrants that was derived from the classic study by Johnson & Newport (1989).

The results of Flege et al. (1999) confirmed J&N (1989) in showing that the Koreans' GJT scores decreased as age of arrival (AOA) in the US increased. The results were difficult to interpret, however, because several variables that might be expected to influence the GJT scores were confounded with AOA.

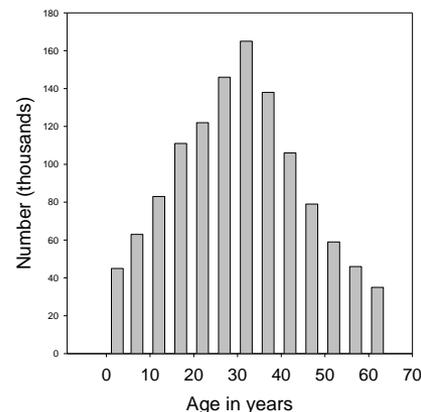
As will be described in detail later, an aim of Flege et al. (1999) was to untangle these confounds using novel analysis techniques.

AOA

AOA is used widely in second language acquisition (SLA) research to index immigrants' age at the time of first exposure to an L2 (e.g., Flege et al. 1995). AOA is a useful selection variable because it can be reported accurately and because it is possible—although not necessarily easy—to recruit immigrants whose arrival ages range from early childhood to 60 years and beyond.

It is widely assumed in the SLA literature that immigrants' age of first extensive exposure to an L2 is related to their state of neurological maturation when L2 learning began.

At right: *The chronological ages of immigrants who became legal residents of the US in 2006.*



AOA

Regarding “age” effects on GJT scores, DeKeyser (2000) concluded that:

Somewhere between the ages of 6-7 and 16-17, everybody loses the mental equipment required for the abstract patterns underlying a human language ... The severe decline of the ability to induce abstract patterns implicitly is an inevitable consequence of fairly general aspects of neurological maturation (2000, pp. 518-519)."

Scovel (2000) proposed that a critical period for L2 learning results from a reduction in cerebral “plasticity”.

Most SLA researchers accept these general premises. However, debate continues regarding what are the neurological underpinnings of a “critical period” for L2 learning (Lenneberg, 1967), when the hypothesized critical period closes, whether the closure is permanent, and whether the closure occurs at different stages of neurological (or cognitive) maturation for different linguistic domains.

Maturational state

Johnson & Newport (1989) was one of the first to examine the effect of AOA on the learning of English morphosyntax.

These authors (henceforth J&N 1989) recruited 46 Chinese and Koreans here at the University of Illinois. These immigrants had arrived in the US between the ages of 3-39 years and had been living here from 3-26 years.

The immigrants' knowledge of English morphosyntax was evaluated using a grammaticality judgment test (GJT) that included sentences testing 12 types of grammatical patterns in English. These sentence types were thought to represent the "*most basic aspects of English sentence structure*" (1989, p. 72)

Maturational state

Half of the 276 sentences in the GJT were grammatical (G) and the other half were ungrammatical (U) versions of the same sentence. For example:

Last night the old lady died in her sleep

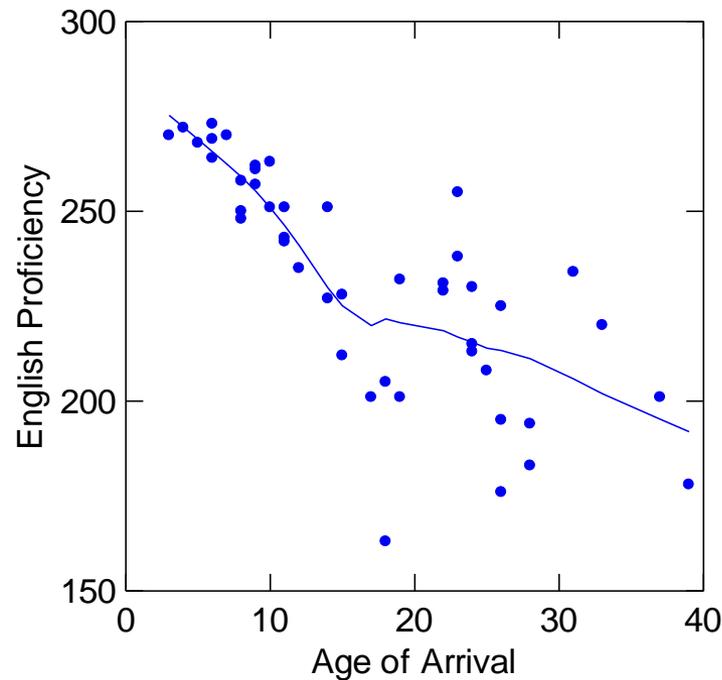
**Last night the old lady die in her sleep*

.The immigrants' task was to decide if each sentence was grammatical (G) or ungrammatical (U) in English.

The dependent variable was the percentage of correct responses. The performance of AOA-defined subgroups of nonnative speakers was compared to that of native English speakers by t-tests.

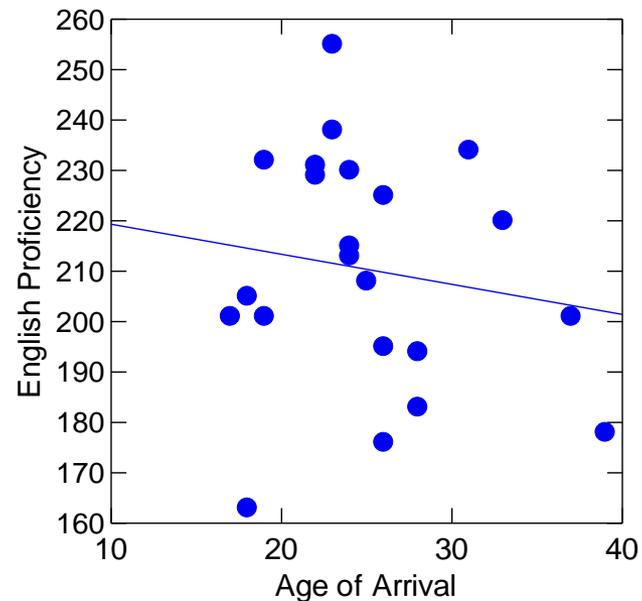
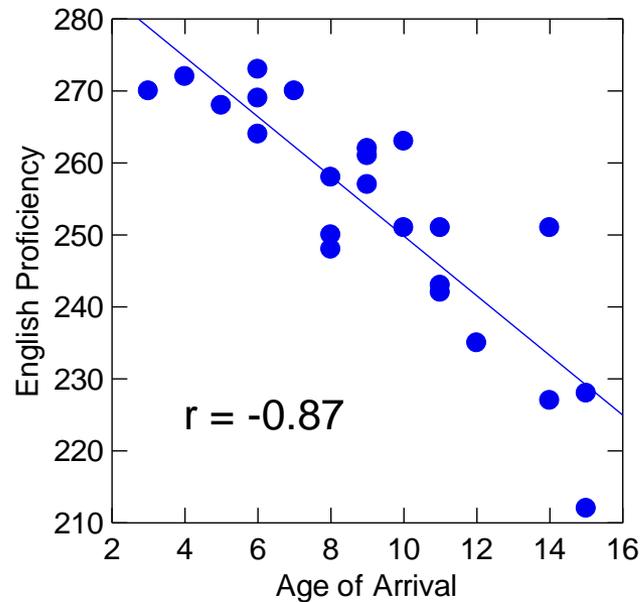
Maturational state

The GJT scores obtained by J&N (1989) generally decreased as a function of AOA: the later the arrival in the US, the lower the scores tended to be.



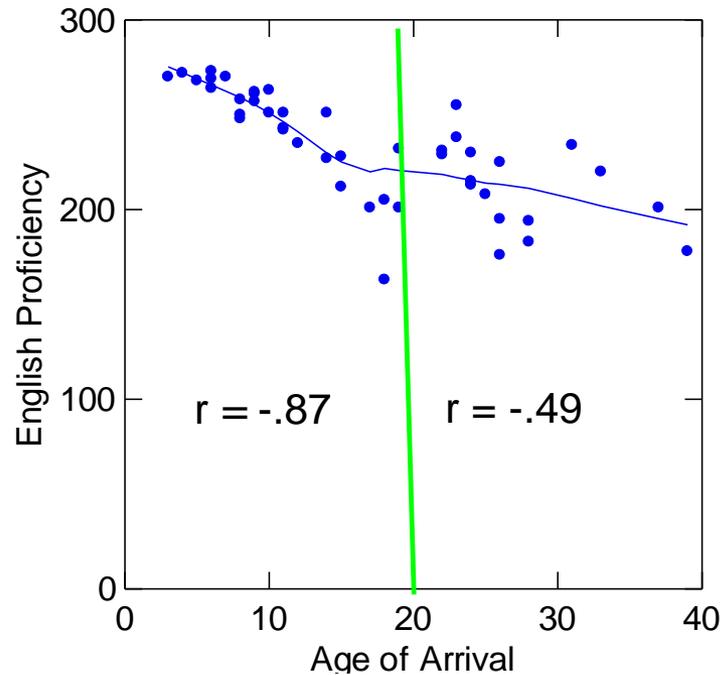
Maturational state

A significant AOA-GJT correlation was observed for immigrants having AOAs less than 15 years but not for those who arrived in the US later in life. This was interpreted to mean that the immigrants' "maturational states" upon arrival in the US was an important determinant of their GJT scores.



Maturational state

Hakuta & Bialystok (1994) later noted that simply shifting the AOA “cut point” from 15 to 20 years resulted in significant AOA-GJT correlations for both subsets of participants. This observation undermined the “maturational state” interpretation.



Maturational state

Although it was of considerable theoretical importance, the J&N (1989) study was limited in several ways:

- both Korean and Chinese learners of English were tested, rather than a single L1 group;
- the sample size was small;
- the lower limit on years of residence in the US was set at just 3 years, which means that some of the immigrants were probably still in fairly early stages of L2 learning;
- the number of GJT items presented was large, creating the possibility of errors due to inattention or boredom;
- no information regarding frequency of L2 use was provided.

Maturational state

The aim of Flege, Yeni-Komshian & Liu (1999) was to provide additional insight into the learning of English morphosyntax by Korean immigrants to the US.

We administered a GJT to 240 Koreans living in the Washington, D.C. area. The Koreans had arrived in the US between the ages of 1-23 years and had been living there for a minimum of 8 years when tested.

Most of our participants were young adults. The chronological age of our Koreans participants differed little (*mean* = 26 years, *range* = 17-46) from members of the native English comparison group (*mean* = 27 years, *range* = 20-45).

Maturational state

Koreans who had lived in the US for at least 8 years were selected on the basis of AOA and gender (half male, half female).

The AOA of individual participants ranged from 1-23 years. The mean AOA values of the ten AOA-defined subgroups ranged from 3-22 years (Flege et al., 1999, Table 1; see also Yeni-Komshian et al., 2000).

Recruiting and testing the 240 Koreans took nearly 2 years. In the first year Korean students helped us recruit other Koreans who could be located on or near the University of Maryland-College Park campus.

In the second year we recruited Koreans with the help of the pastors of nine Korean churches located in the District of Columbia, Maryland, and northern Virginia. These latter participants were tested in a quiet room located at their church rather than on the University of Maryland campus.

Maturational state

Most (127) sentences in our 144-item grammaticality judgment test were taken from the J&N (1989) study.

Our GJT was shorter, by design, than that of J&N (1989). We eliminated items from the J&N (1989) test that did not serve to distinguish Early from Late learners. The 16 new items that were added tested lexically specified subject/object raising.

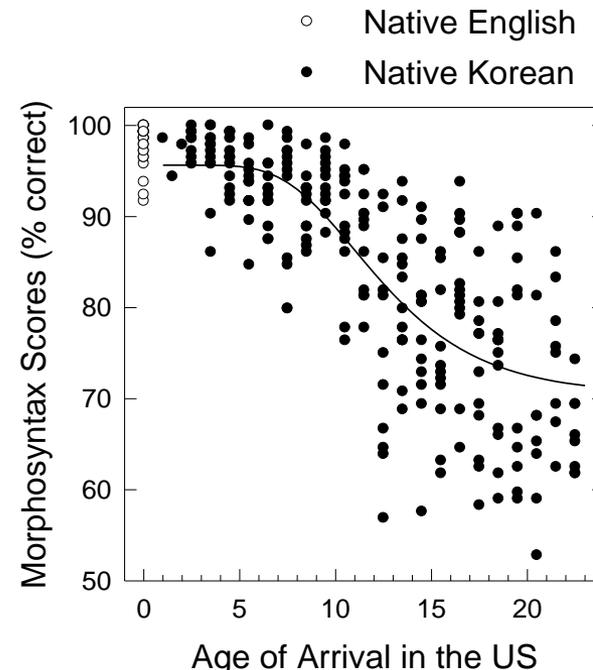
As in the original J&N (1989) study, half of the sentences were grammatical and half were ungrammatical in standard English.

Maturational state

Here are the percent correct GJT scores obtained for the 240 Koreans and 24 native English (NE) controls.

We used *t*-tests to compare each of the ten AOA-defined Korean groups ($n = 24$ each) to the scores obtained from the 24 NE controls.

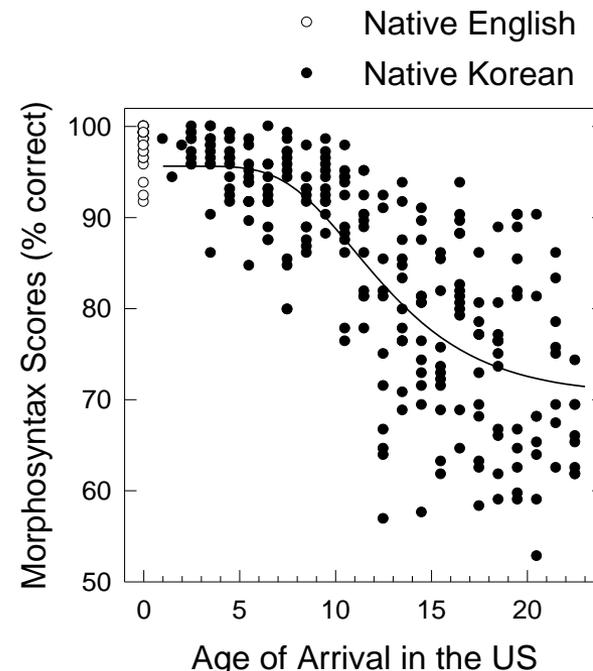
Korean groups having mean AOAs of 7-22 years obtained significantly lower scores than the NE controls did whereas Koreans with mean AOAs of 3 and 5 years did not differ significantly from the NE group ($p < .01$).



Maturational state

A Gumpertz-Makeham growth function was fit to the percent correct scores to visually organize the data. Inspection of this curve suggests that

- For Koreans having an AOA of less than about 8 years, GJT scores resembled those of the NE speakers;
- Koreans with AOAs ranging from about 9-15 years obtained increasingly lower scores as AOA increased;
- For Koreans who arrived after about 15 years of age the GJT scores seemed to decrease at a slower rate as a function of AOA than for Koreans who arrived earlier.



Maturational state

What explains the relation between the GJT scores and AOA?

As already mentioned, J&N (1989) suggested that the “maturational state” of their Korean and Chinese participants was an important determinant of the GJT scores obtained.

The “maturational state” in question might have depended on some aspect of neurological development, overall neural “plasticity”, or cognitive development at the time of immigration.

Another (non-exclusive) possibility to consider is that the GJT scores varied as a function of input. On this view, the seeming potency of AOA as predictor of the GJT scores arose from its correlation with variables other than maturational state.

Unconfounding AOA

Flege et al. (1999) developed two analysis techniques to help better understand the role of AOA.

1. Two GJT subscores were computed for each participant. The subscores were established on the basis of principal components analyses which identified test items that patterned together. I will refer to the two subscores, each based on 44 sentences (half grammatical, half ungrammatical), as the “Rule-based” and “Lexically-based” GJT items.
2. The two subscores and the overall GJT scores were then examined in matched subgroup analyses, the intent of which was to reduce or eliminate factors that were confounded with our Korean participants' AOA in the US.

Unconfounding AOA

The Rule-based and Lexically-based GJT subscores were functional rather than syntactically motivated groupings. The items making up the two subscores were drawn from multiple sentence “types” (see Flege et al., 1999, p. 90 ff and Appendix 1) .

The Rule-based GJT items tested knowledge of regular, productive and generalizable rules of the surface morphology of English. All involved case or number assignment on nouns or else person or tense markers on verbs (e.g., regular past tense on plural formation, third-person singular morphology on present tense verbs, or case assignment on personal pronouns). For example:

(The) boys are going to the zoo this Saturday.

The man paints (painted) his house yesterday.

Them (They) worked on the project all night.

Unconfounding AOA

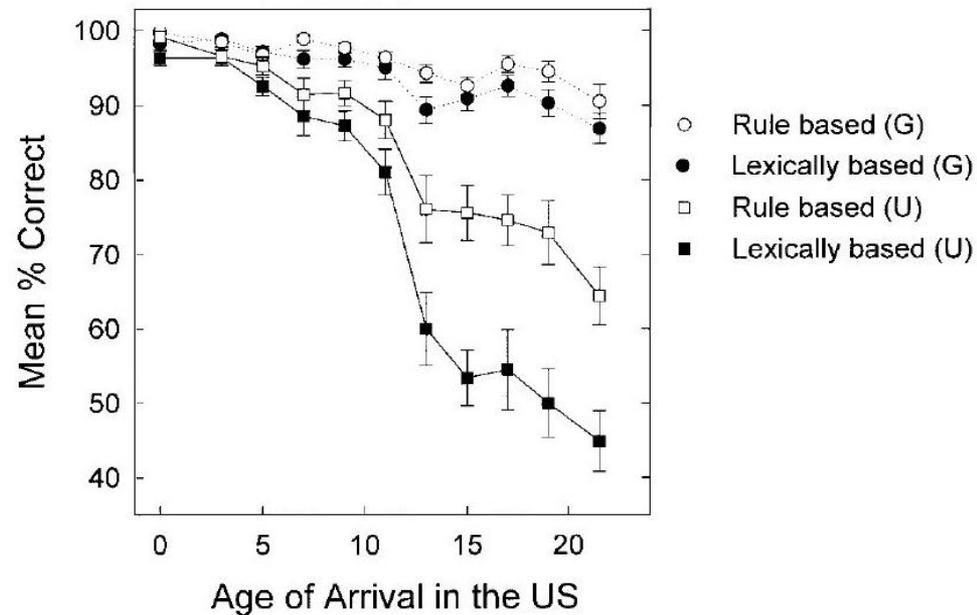
The Lexically-based subscores, on the other hand, tested irregular and ungeneralizable aspects of English morphosyntax involving the proper assignment of particles or prepositions with verbs or knowledge of idiosyncratic features of particular English verbs.

Some sentences tested that participants' knowledge of which preposition should precede a nominal complement (e.g., **The farmers were hoping rain*), the use of a particle in phrasal verbs (e.g., **The little boys laughed the clown*), or the placement of particles in phrasal verbs (e.g., **The man climbed the ladder up carefully*).

All ungrammatical Lexically-based sentences could be made grammatical by replacing the verb (for example, changing “lets” to “permits” in **The man lets his son to watch TV*). The ungrammatical Rule-based sentences could not be corrected in this way.

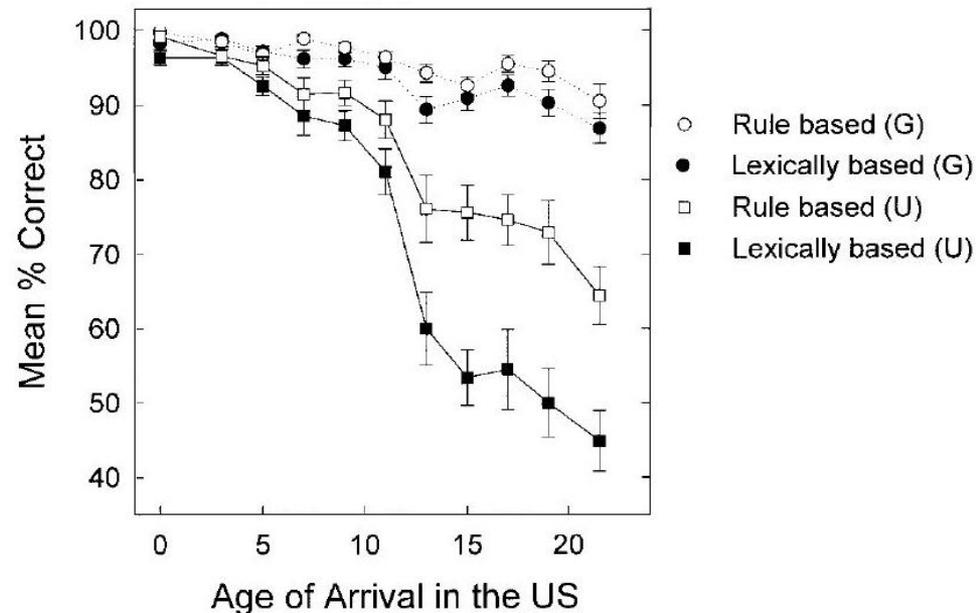
Flege et al. (1999)

Here we see the mean Percent Correct scores obtained for both the grammatical and ungrammatical versions of the Rule-based and Lexically-based subscores. (The group with an AOA of "0" is the native English control group.)



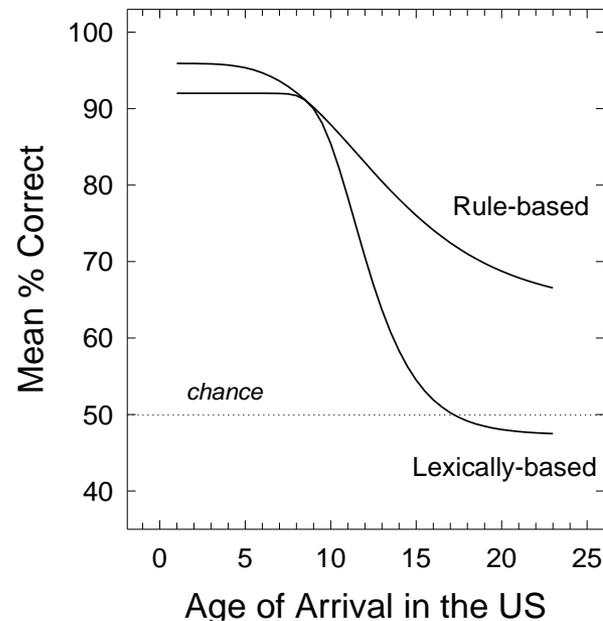
Flege et al. (1999)

For the NE speakers and Korean groups having mean AOAs of 3-11 years, the difference between the Rule-based and Lexically-based subscores was non-significant. However, for the Korean groups having mean AOAs of 13-21 years, significantly higher scores were obtained for Rule-based than Lexical-based subscores.



Flege et al. (1999)

Gumpertz-Makeham growth functions have been fit here to the Percent Correct scores obtained for ungrammatical Rule-based and Lexically-based GJT items. A far more rapid decrease in Percent Correct scores seems to have existed in the AOA range of about 9-15 years for the Lexically-based than Rule-based subscores.



Flege et al. (1999)

For this talk I selected three groups of Koreans for further analysis: “Early” (AOA = 1 - 8.5 years), “Mid” (AOA = 9.5 - 14.5 years) and “Late” (AOA = 15.5 - 22.5 years).

	1 - Early	2 - Mid	3 - Late	F	Tukey .05
N of males/females	42/42	35/38	42/41		
Age of arrival in U.S.	5.5 (1.0-8.5)	12.0 (9.5-14.5)	18.7 (15.5-22.5)	880.9	1 < 2 < 3
Chronological age	22.8 (17-30)	24.7 (19-47)	31.0 (22-47)	75.9	1 < 2 < 3
Exp. to English in Korea	0.5 (0.5-1.5)	0.9 (0.5-4.5)	5.7 (0.5-10.5)	235.5	1, 2 < 3
Years to speak E "comfortably"	1.7 (0-5.5)	1.6 (0.5-7.5)	2.8 (0.5-9.5)	12.4	1, 2 < 3
Length of residence in U.S.	17.7 (11-25)	13 (8-30)	12.8 (7-25)	37.2	1 > 2, 3
Ratio of English/Korean use	1.87 (0.63-5.00)	1.23 (0.55-2.25)	0.97 (0.38-2.81)	73.2	1 > 2 > 3

Flege et al. (1999)

Members of the Early, Mid and Late subgroups differed significantly in chronological age as well as language use.

	1 - Early	2 - Mid	3 - Late	F	Tukey .05
N of males/females	42/42	35/38	42/41		
Age of arrival in U.S.	5.5 (1.0-8.5)	12.0 (9.5-14.5)	18.7 (15.5-22.5)	880.9	1 < 2 < 3
Chronological age	22.8 (17-30)	24.7 (19-47)	31.0 (22-47)	75.9	1 < 2 < 3
Exp. to English in Korea	0.5 (0.5-1.5)	0.9 (0.5-4.5)	5.7 (0.5-10.5)	235.5	1, 2 < 3
Years to speak E "comfortably"	1.7 (0-5.5)	1.6 (0.5-7.5)	2.8 (0.5-9.5)	12.4	1, 2 < 3
Length of residence in U.S.	17.7 (11-25)	13 (8-30)	12.8 (7-25)	37.2	1 > 2, 3
Ratio of English/Korean use	1.87 (0.63-5.00)	1.23 (0.55-2.25)	0.97 (0.38-2.81)	73.2	1 > 2 > 3

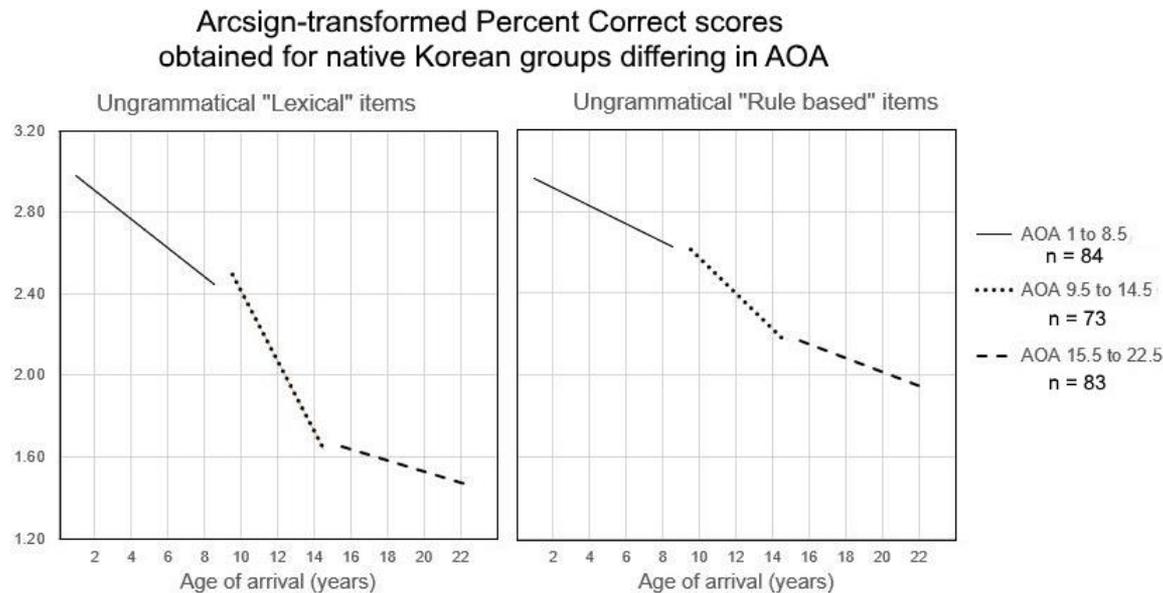
Flege et al. (1999)

Members of the Early group had lived significantly longer in the US than those in Mid and Late groups. Members of the Late group had studied English longer in Korea before immigrating than had members of the other two groups, and reported needing significantly longer to learn English after arriving in the US.

	1 - Early	2 - Mid	3 - Late	F	Tukey .05
N of males/females	42/42	35/38	42/41		
Age of arrival in U.S.	5.5 (1.0-8.5)	12.0 (9.5-14.5)	18.7 (15.5-22.5)	880.9	1 < 2 < 3
Chronological age	22.8 (17-30)	24.7 (19-47)	31.0 (22-47)	75.9	1 < 2 < 3
Exp. to English in Korea	0.5 (0.5-1.5)	0.9 (0.5-4.5)	5.7 (0.5-10.5)	235.5	1, 2 < 3
Years to speak E "comfortably"	1.7 (0-5.5)	1.6 (0.5-7.5)	2.8 (0.5-9.5)	12.4	1, 2 < 3
Length of residence in U.S.	17.7 (11-25)	13 (8-30)	12.8 (7-25)	37.2	1 > 2, 3
Ratio of English/Korean use	1.87 (0.63-5.00)	1.23 (0.55-2.25)	0.97 (0.38-2.81)	73.2	1 > 2 > 3

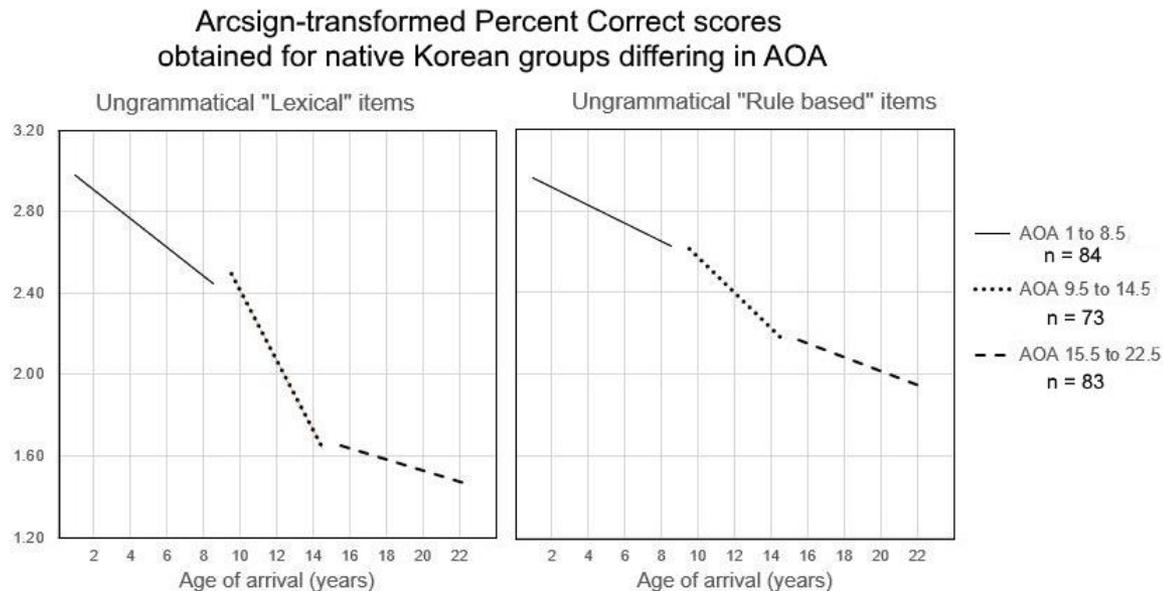
Flege et al. (1999)

Here I have fit linear functions to the three groups' ungrammatical Lexically-based and Rule-based GJT subscores. The slopes for the Early and Mid groups for both subscores differed significantly from zero (Bonferroni $p < .05$) indicating that the scores decreased as a function of AOA. However, the slopes obtained for the Late group did not differ significantly from zero.



Flege et al. (1999)

Scores obtained by members of the Late group for the Lexically-based items were at or near a chance level of responding. This might mean that the Late learners lacked the capacity for certain aspects of L2 learning. Their scores for Rule-based items were above chance, yet did not show a significant decline as AOA increased. This might mean that members of the Late group had reached an asymptotic (“end state”) level of performance.



Flege et al. (1999)

Both of these interpretations are likely to be wrong.

Members of the NE group had been exposed to English full time for an average of 27 years. The Lexically-based scores obtained for 12 (50%) of them were not at ceiling. The Late learners used English somewhat less than half of the time and had lived in the United States for an average of 13 years. We can infer that they had received roughly 1/4 of the English input that the NE speakers had received.

	NE	1 - Early	2 - Mid	3 - Late
PC all items	4/24 (17%)	4/84 (5%)	0/73 (0%)	0/83 (0%)
PC lexical-based	12/24 (50%)	20/84 (24%)	2/73 (3%)	1/83 (1%)
PC rule-based	19/24 (79%)	25/84 (30%)	5/73 (7%)	1/83 (1%)

The values shown here indicate the number of participants in four groups who obtained percent correct scores of 100%, and so were at ceiling.

Flege et al. (1999)

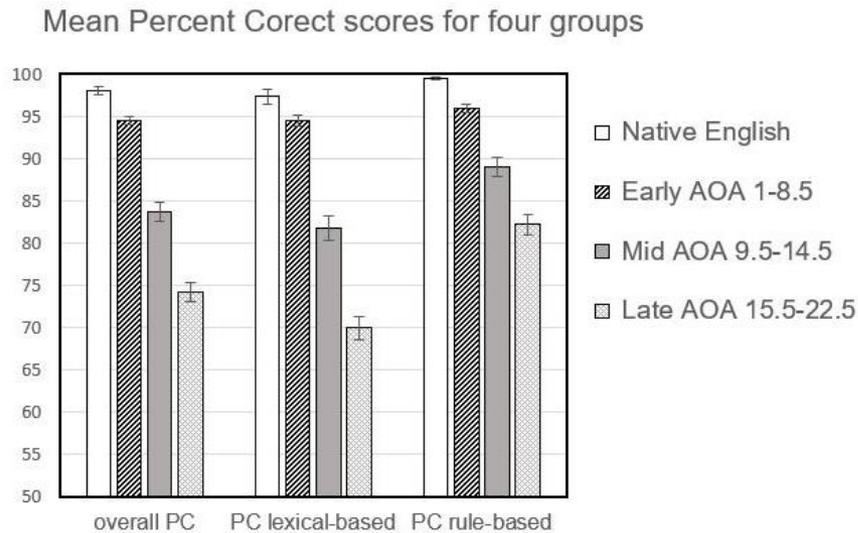
The Rule-based scores obtained for five (21%) of the NE speakers were not at ceiling.

Let's assume that the NE speakers were usually exposed to grammatically correct English sentences spoken by other native English speakers. The same might hold true to a lesser extent for the Korean Late learners. If so, their Rule-based GJT scores might have reflected the grammatical errors they heard when interacting with other Koreans, not a reduced capacity for language learning.

	NE	1 - Early	2 - Mid	3 - Late
PC all items	4/24 (17%)	4/84 (5%)	0/73 (0%)	0/83 (0%)
PC lexical-based	12/24 (50%)	20/84 (24%)	2/73 (3%)	1/83 (1%)
PC rule-based	19/24 (79%)	25/84 (30%)	5/73 (7%)	1/83 (1%)

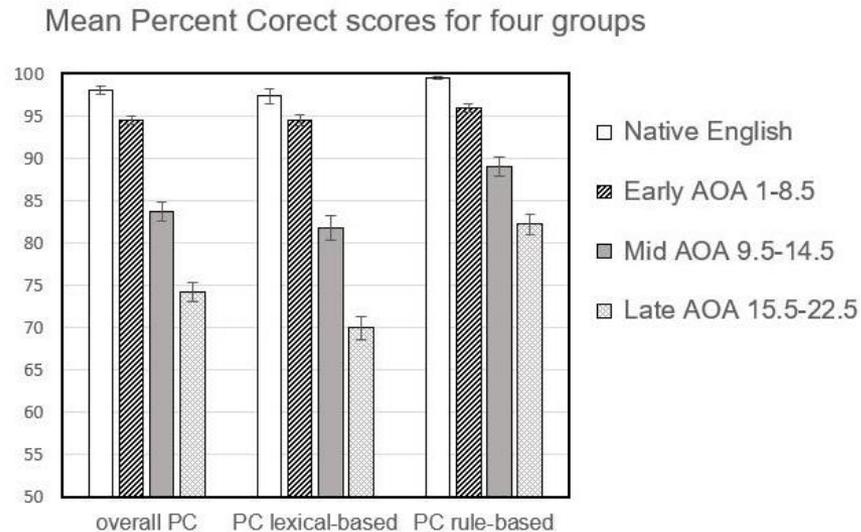
Flege et al. (1999)

Here we see the mean Percent Correct scores obtained by the Early, Mid and Late Korean learners of English for all 144 GJT items as well as the Lexically-based and Rule-based subscores.



Flege et al. (1999)

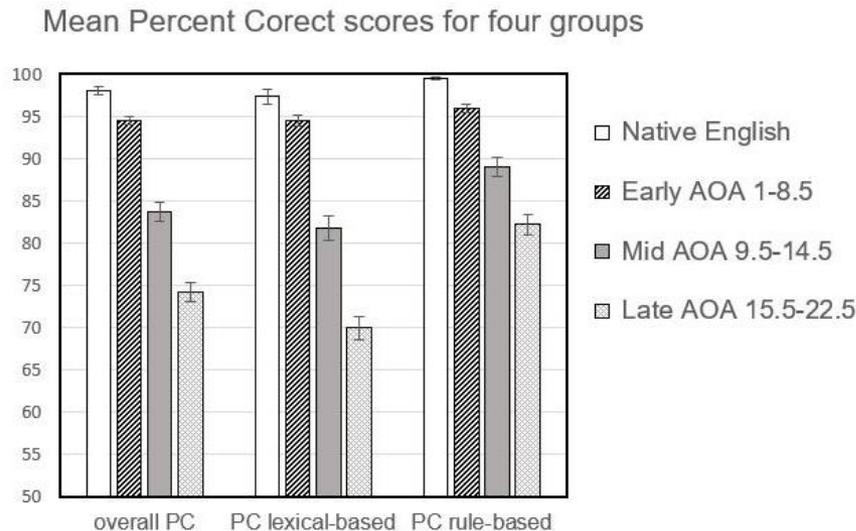
All three Korean groups obtained significantly lower scores than the NE speakers did; this held true for all three outcome measures ($p < .05$).



Given unequal cell sizes and heterogeneity of variance, parametric tests were inappropriate for the 9 native vs nonnative comparisons. I performed Mann-Whitney U tests instead. Eight tests yielded U values ranging from -4.2 to -7.4 and were significant at a Bonferroni-corrected level of $p < .001$. The test for the Early group's Lexically-based subscores yielded a U value of -2.9; it was significant at a Bonferroni corrected .05 level.

Flege et al. (1999)

For all three outcome variables: Significantly higher scores were obtained for the Early group than for the Mid group who, in turn, obtained significantly higher scores than the Late group did ($p < .01$).



For consistency, I again used Mann-Whiney tests to compare scores for the Mid group to those obtained for Early and Late groups (six tests in all). All six comparisons were significant at a Bonferroni-corrected 0.01 level.

Unconfounding AOA

It is of theoretical interest to understand the basis of the native-nonnative differences as well as the differences between Korean groups differing in AOA.

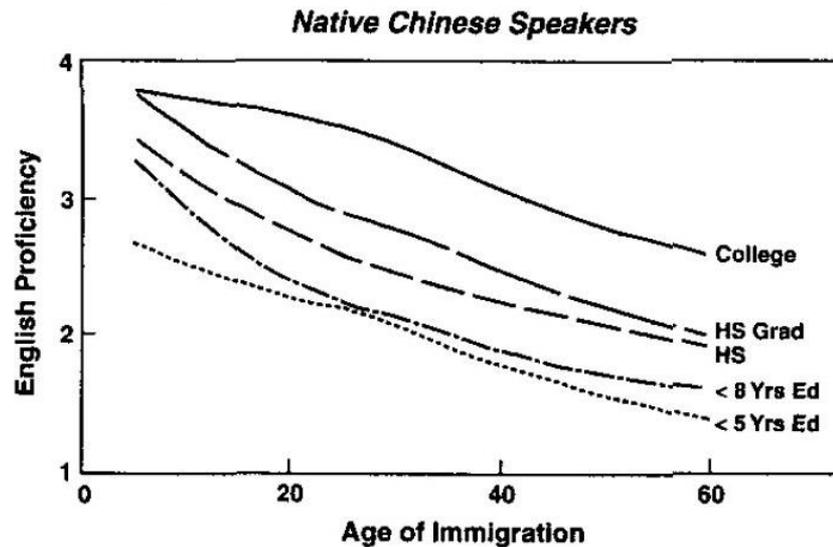
As a starting point, consider that four variables obtained on a language background questionnaire correlated significantly ($p < .001$) with all three outcome measures.

The correlation between the outcome measures and AOA is familiar to SLA researchers, but the equally strong correlations between AOA and Years of Education in US schools may come as a surprise to some.

% Correct	AOA	Years of Educ.	E/K use ratio	LOR
PC all items	-0.753	0.742	0.563	0.385
PC lexically-based	-0.739	0.723	0.579	0.389
PC rule-based	-0.598	0.618	0.451	0.389

Unconfounding AOA

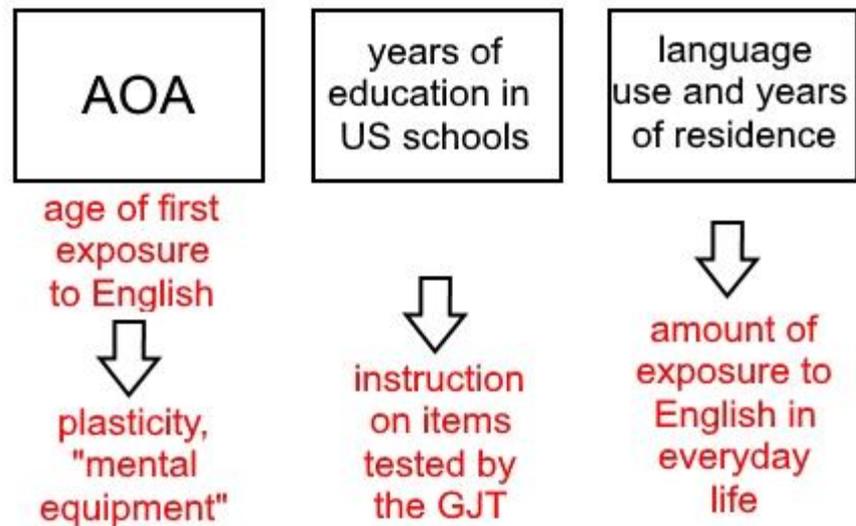
The education-AOA we obtained for the Koreans immigrants is reminiscent of the census data for Chinese immigrants reported by Hakuta et al. (2003). These authors found that Chinese immigrants' self-rated ability in English differed systematically as a function of education level across a very wide range of AOAs.



Unconfounding AOA

It's not a simple matter to know what is the best predictor(s) of GJT scores. Many are convinced that "earlier is better" because a relatively late exposure to an L2 reduces the capacity for L2 learning.

But other possible explanations exist and must be evaluated.



Unconfounding AOA

As in previous research examining L2 learning by immigrants, Flege et al. (1999) observed multi-collinearity among variables associated with the Koreans' ages of arrival (AOA) in the United States. The Koreans' AOA was correlated with their:

- chronological age at the time of test, $r = .68$;
- self-estimated use of English and Korean, $r = -0.56$ and 0.66 ;
- years residence in the US, $r = -0.42$,
- years of education in the US, $r = -0.92$.

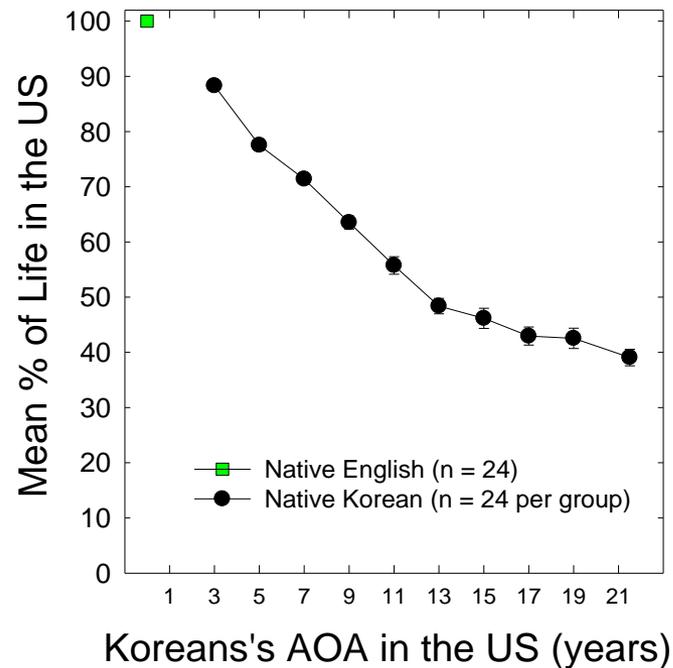
Moreover, these variables were all correlated significantly with one another ($p < .01$).

What a mess!

The problem with AOA

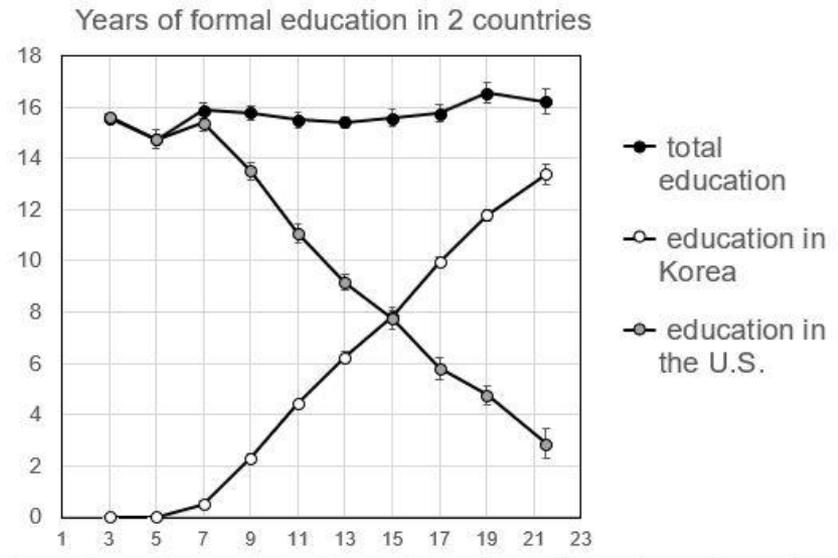
Let's consider some of the variables confounded with AOA in greater detail.

Members of the NE control group (green box) had lived 100% of their lives in the US. The Koreans, on the other hand, had spent only portions of their lives in the US. The mean percentages of the Koreans' lives spent in the US decreased systematically as AOA increased.



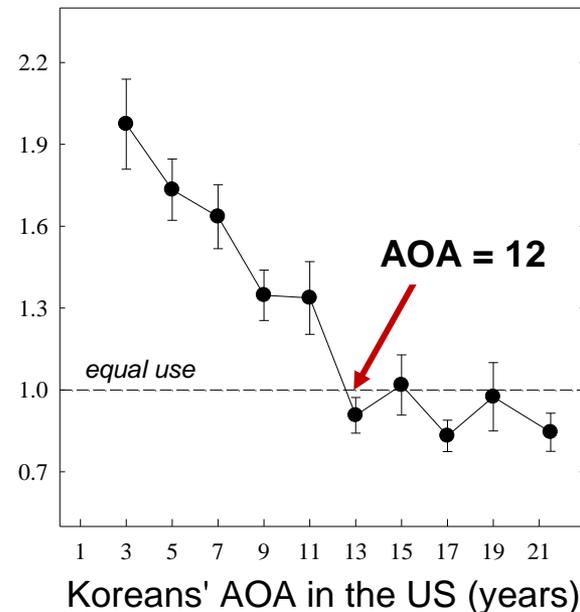
Unconfounding AOA

Most Koreans had obtained a BA or BS degree (or were about to do so). Koreans who arrived before the age of 15 years had received more education in the US than in Korea whereas the reverse held true for those who arrived after the age of 15 years.



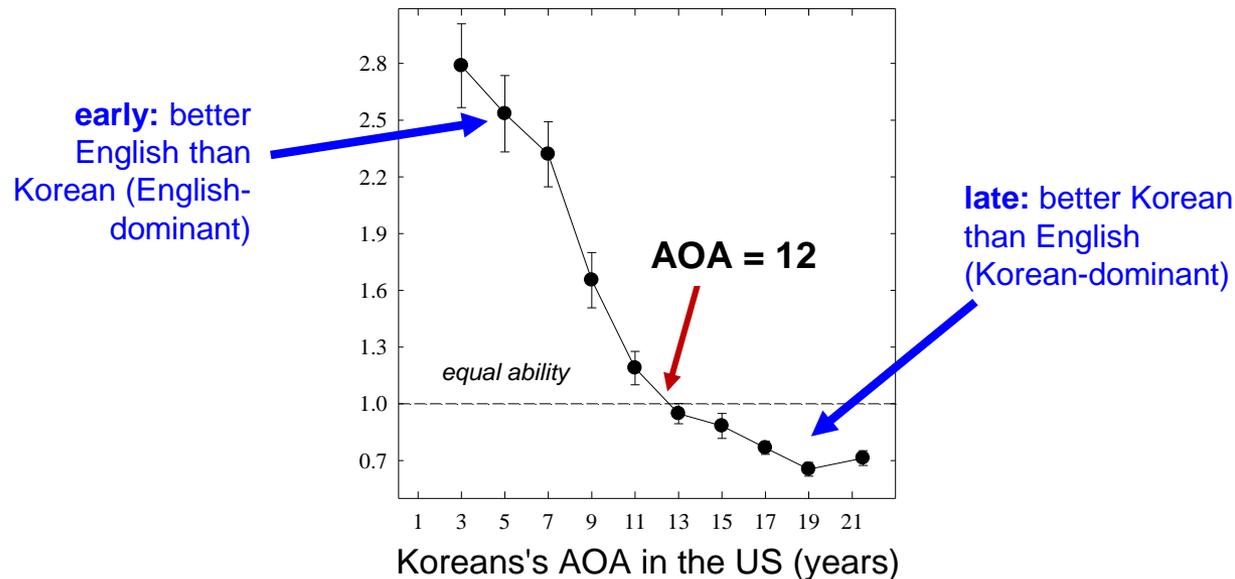
Unconfounding AOA

Here we see the mean ratio of self-rated frequency of English/Korean use. Koreans who arrived in the US before the “magic age” of 12 years used English more than Korean whereas Koreans who arrived after the magic age tended to use English and Korean with roughly equal frequency.



Unconfounding AOA

Here are the mean ratios of self-rated proficiency in English/Korean. Most Koreans who arrived in the US before the “magic age” of 12 judged themselves to be more proficient in English than Korean whereas the reverse held true for those who arrived after the magic age.



Unconfounding AOA

Responses to 39 items on a questionnaire were submitted to a principal components analysis. The resulting factor scores were examined in a step-wise multiple regression analysis.

The factor we chose to call “Age of L2 learning” accounted for substantially more variance in the GJT scores than the factor called “Length of residence” in the US.

Results of a multiple regression analysis examining the GJT scores obtained for 240 Korean immigrants

Step	Variable	R^2	Change	F value	Probability
1	F1: Age of L2 learning	0.494	0.494	313.7	.001
2	F10: Length of residence	0.547	0.053	33.6	.001
3	F2: English media input	0.570	0.023	14.5	.001
4	F8: Sound processing ability	0.593	0.023	14.5	.001
5	F9: Integrative motivation	0.616	0.023	14.4	.001
6	F11: Instrumental motivation	0.625	0.010	6.0	.015

Unconfounding AOA

This finding might be interpreted to mean that the Koreans' age of first exposure to English, as indexed by age of arrival (AOA), was by far the most important determinant of the GJT scores and that the quantity of English language input, as indexed by LOR, was of little importance.

In my view, such an interpretation is incorrect for two reasons.

Results of a multiple regression analysis examining the GJT scores obtained for 240 Korean immigrants

Step	Variable	R^2	Change	F value	Probability
1	F1: Age of L2 learning	0.494	0.494	313.7	.001
2	F10: Length of residence	0.547	0.053	33.6	.001
3	F2: English media input	0.570	0.023	14.5	.001
4	F8: Sound processing ability	0.593	0.023	14.5	.001
5	F9: Integrative motivation	0.616	0.023	14.4	.001
6	F11: Instrumental motivation	0.625	0.010	6.0	.015

Unconfounding AOA

First, AOA had a very high loading (.912) on the “Age of L2 learning” factor but five other variables that might be relevant for the GJT scores also had high loadings:

- years of education in the U.S. (-.856)
- use of Korean with spouse (.786)
- use of Korean with close friends (.729)
- use of Korean at social gatherings (.737)
- use of English at social gatherings (-.712)

The multi-collinearity lurking under the surface of the “Age of L2 learning” factor makes interpretation of this factor difficult if not impossible.

Unconfounding AOA

Second, a novel form of analysis revealed that when variables that were confounded with AOA were held constant the effect of AOA disappeared in some cases.

We used a subgroup matching technique in an attempt to untangle the confound between AOA, on the one hand, and language use and amount of formal education on the other hand.

This technique, although simple and straightforward, does not seem to have been used previously because it requires a larger sample than is typical for SLA research.

Unconfounding AOA

One subgroup analysis compared groups of 20 Koreans each who were matched for AOA but differed in years of formal education in the US.

We selected pairs of Koreans having the same AOA who differed by at least four years of US education. Members of the “Much Education” subgroup had received an average of 6.2 more years of US education than members of the “Little Education” subgroup.

The matching process eliminated AOA differences between the two subgroups (*mean* = 12.3 years for both subgroups, *range* for the 40 Koreans = 3.5- 21.5 years, *SE* = 0.893).

Matching Korean participants for AOA also had the effect of eliminating significant between-group differences in self-reported frequency of English and Korean use. However, the LOR difference between the matched subgroups (Much Education = 17.0, Little Education = 13.6 years) remained significant ($p < .05$).

atched subgroups

The Much Education subgroup obtained significantly higher Rule-based GJT subscores than the Little Education subgroup did. However, the two subgroups did not differ significantly for the overall percent correct GJT scores or the Lexically-based subscores.

This suggests that amount of formal education in US schools affected the Koreans' knowledge of rule-based regularities of English grammar but not aspects of morphosyntax reflecting idiosyncratic properties of particular lexical items.

	US Educ. 15.1 AOA 12.3	US Educ. 8.9 AOA 12.3	<i>F</i> (1,38)	<i>p</i> =
PC all items	86.6%	80.0%	3.290	0.0776
PC lexically-based	84.3%	78.5%	1.394	0.2452
PC Rule-based	93.4%	84.7%	9.150	0.0045

Matched subgroups

We evaluated this interpretation in a control analysis which compared subgroups of 20 Koreans each who had the same amounts of US education as the subgroups just consider but whose AOA, on a pairwise basis, differed by at least 7 years.

The resulting AOA difference between the two subgroups averaged 8 years (means = 7.2 vs 16.2). Years of education for the 40 Koreans in the control analysis averaged 12.0 years ($SE = 0.557$) and ranged from 5-18 years).

In addition to a difference in Years of US education, the “Early” and “Late” learners in the control analysis differed significantly ($p < .001$) in LOR (means = 14.8 vs 11.3 years), English use (4.2 vs 3.4), Korean use (2.7 vs 3.7) and the ratio of English to Korean use (1.76 vs 0.97).

Matched subgroups

All three between-group differences reached significance in the control analysis ($p < .001$). This was expected because the two unmatched subgroups differed along multiple dimensions (Years of US education, English use and Korean use as well as AOA), thereby reflecting characteristics evident in the entire sample of 240 Koreans.

The control analysis showed that the lack of significant between-group differences for two outcome variables in the preceding matched subgroup analysis was not due to insufficient statistical power. The absence of effects could therefore be attributed to characteristics of the participants making up the matched subgroups.

	US Educ. 15.1	US Educ. 8.9		
	AOA 7.2	AOA 16.2	F(1,38)	p =
PC all items	93.4%	76.5%	45.066	0.0000
PC lexically-based	94.0%	73.4%	57.769	0.0000
PC Rule-based	95.5%	84.5%	16.607	0.0003

Matched subgroups

Another matched subgroup analysis compared 20 Koreans each who differed in language use but were matched for AOA.

We selected pairs of Koreans having the same AOA (overall *range* = 4.5-19.5) but whose ratings of Korean use in a variety of contexts (an average of nine 5-point rating scales) differed by at least 1.5.

Members of the matched subgroups also differed, of course, in their mean ratings of English use, which were inversely correlated with the Korean use estimates, $r(38) = -0.84$, $p < .0001$.

Matching the two subgroups for language use eliminated significant between-groups differences in Years of US education (*means* = 12.5 for high English/low Korean use vs. 12.6 for Koreans with low English/high Korean use, $p > 0.10$), and also LOR (*means* = 15.2 vs. 14.6 years, $p > 0.10$).

Matched subgroups

For the subgroups of Koreans matched for AOA but differing in language use, only the Lexically-based subscores differed significantly.

The absence of a language use effects for the overall GJT scores and the Rule-based subscores seemed to indicate that frequency of L2 use had little effect on the learning of features of English morphosyntax that are predictable by rule and so might be learned at school. At least by those who occasionally paid attention!

	EngUse 4.5	EngUse 3.3		
	KorUse 2.1	KorUse 4.1		
	AOA 11.4	AOA 11.4	<i>F</i> (1,38)	<i>p</i> =
PC all items	88.5%	82.9%	2.449	0.1259
PC lexically-based	88.9%	80.1%	4.143	0.0489
PC Rule-based	91.7%	88.2%	1.324	0.2572

Matched subgroups

We evaluated this interpretation in a control analysis. It compared subgroups of 20 Koreans each who had the same language use values as in those in the preceding analysis but were selected without regard to AOA.

Not surprisingly, the unmatched subgroups in the control analysis showed the same pattern of significant between-group differences that were evident for Early vs Late learners considered earlier.

	Much E	Little E		
	Little K	Much K	F(1,38)	p =
Length of residence	15.9	12.6	7.5	0.0093
English use	4.5	3.2	61.4	0.0000
Korean use	2.1	4.1	299.1	0.0000
E/K use ratio	2.2	0.8	177.1	0.0000
Years of Education	15.9	7.9	160.4	0.0000

Matched subgroups

As a result, the unmatched subgroups differed significantly for all three outcome measures.

The results of this control analysis indicate that the absence of significant differences between matched subgroups for overall GJT scores and Rule-based subscores can not be attributed to a lack of statistical power.

	EngUse 4.5	EngUse 3.2		
	KorUse 2.1	KorUse 4.1		
	AOA 7.0	AOA 16.2	<i>F</i>(1,38)	<i>p</i> =
PC all items	93.5%	74.2%	69.166	0.0000
PC lexically-based	71.6%	93.5%	56.031	0.0000
PC Rule-based	80.1%	94.4%	30.488	0.0001

Summary

The Flege et al. (1999) study summarized in this talk replicated the effect of AOA on grammaticality judgment test (GJT) scores obtained by Johnson & Newport (1989) for Korean and Chinese immigrants. This was expected inasmuch as most GJT stimuli were drawn from the earlier J&N (1989) study.

Groups of Koreans have mean AOAs of 7-22, but not those having AOAs of 3 and 5 years, obtained average percent correct GJT scores that were significantly lower than those obtained for members of the native English comparison group.

This finding should not be regarded as an absolute “cut-off” for success in learning English grammar. Different results might well be obtained if a larger number of Korean immigrants were tested, if less well educated Koreans were recruited, or if an even more fine grained test of grammatical knowledge were used.

Summary

For today's talk I selected three subgroups of Koreans for further analysis:

- “Early” (AOA 1-8.5 years),
- “Mid” (9.5-14.5) and
- “Late” (15.5-22-5).

All three groups differed significantly from one another and received significantly lower scores than the NE speakers did ($p > .01$).

Although the pattern of between-group differences was straightforward, interpreting the results proved difficult – as is usually the case for studies of L2 learning through immersion – owing to confounds between AOA and other variables that might reasonably be expected to affect the GJT scores (frequency of L1 and L2 use, LOR, and years of formal education in the US).

Summary

We used two techniques to unconfound the AOA variable.

First, we derived GJT subscores of 44 items each from our 144-item GJT.

- the Rule-based items tested knowledge of regular, productive, and generalizable rules of the surface morphology of English;
- the Lexically-based items tested knowledge of irregular and ungeneralizable aspects of English morphosyntax.

Korean groups having mean AOAs of 13-21 obtained significantly lower Lexically-based than Rule-based subscores, but not those having mean AOAs ranging from 3-11 years.

Summary

Second, we carried out matched subgroup analyses in which AOA was held constant and either years of US education or language use was varied.

When AOA was held constant:

- Koreans having many years of US education (*mean* = 15.1 years) obtained significantly higher Rule-based subscores than those having less US education (*mean* = 8.9 years), but the two subgroups did not differ for the overall GJT scores or the Lexically-based subscores.
- Koreans who used English much/Korean little (means = 4.5/2.1) obtained significantly higher Lexically-based subscores than those using English little/Korean much (means = 3.3/4.1) but did not differ significantly for the overall GJT scores or the Rule-based subscores.

Summary

The pattern of results obtained in the matched subgroup analyses were coherent:

- The outcome variable most likely to be influenced by Years of US education, the Rule-based subscores, remained significant when years of education varied and AOA was held constant. The groups differing in Years of education did not differ, however, for the Lexically-based subscores.
- The outcome variable most likely to be influenced by frequency of English use, the Lexically-based subscores, remained significant when language use was varied and AOA was held constant. The groups differing in frequency of English use did not differ, however, for the Rule-based subscores.

Importantly, the results of control analyses suggested the disappearance of significant differences between the matched subgroups of 20 Koreans each was not due to insufficient statistical power.

Theoretical implications

Birdsong (2005, p. 103) noted that the observation of AOA effects in the SLA literature does not prove that Early vs Late differences arise from differences in learners' maturational state at the time of first exposure to an L2.

For Birdsong, it is “logically possible” that variables which “correlate strongly” with AOA might be responsible for the observed effects attributed to AOA. However, the confounded variables “cannot explain away the very robust effects of AOA (2005, p. 104) that have often been observed in the SLA literature.

Birdsong (2005, p. 104) suggested – and I agree – that a critical period account of L2 morphosyntax learning cannot be rejected unless and until the results of empirical research manage to explain away AOA effects.

Theoretical implications

In my opinion the findings of Flege et al. (1999), which we have considered today in detail, do indeed “explain away” AOA as a cause for Early vs Late differences, at least for the grammaticality judgment test used here.

That being the case:

it's time that SLA researchers stop treating variation in L2 input as a nuisance and start carrying out research designed to explore the role of quantity and quality of input on the learning of L2 morphosyntax.

Doing so will, in my opinion, lead to more credible explanations of “age” effects on L2 learning than are currently available in the SLA literature.

End

Thanks for your attention



References

- Bialystok, E. & Hakuta, K. (1994) *In Other Words: The Science and Psychology of Second Language Acquisition*. New York: HarperCollins.
- Birdsong, D. (2005). Interpreting age effects in second language acquisition. In J. Kolll & A. DeGroot (Eds.) *Handbook of bilingualism: Psycholinguistic approaches* (pp. 109-127) Oxford: Oxford University Press.
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *L2*, 499-534.
- Flege, J., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second language learning. *Journal of Memory and Language* 41, 78-104.
- Flege, J. (1999) Age of Learning and second-language speech., In D. Birdsong (Ed.) *Second Language Acquisition and the Critical Period Hypothesis* (pp. 101-132) Hillsdale, NJ: Lawrence Erlbaum.
- Flege, J. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. Hualde (Eds.) *Laboratory Phonology 9*. (pp. 353-380) Berlin: Mouton de Gruyter.
- Hakuta, K.; Bialystok, E. & Wiley, E. (2003). Critical Evidence: A test of the critical-period hypothesis for second-language acquisition *Psychological Science* 14, 31-38.
- Johnson, J., & Newport, E. (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.
- Lennerberg, E. (1967) *Biological Foundations of Language*. New York: Wiley..
- Scovel, T. (2000) A critical review of critical period research. *Annual Review of Applied Linguistics* 20, 213-223.
- Yeni-Komshian, G., Flege, J., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition* 3, 131-150.