“Age” effects on second language acquisition*

James Emil Flege

Professor Emeritus
Speech and Hearing Sciences
University of Alabama at Birmingham

New Sounds 2010  1-3 May, 2010
Poznań, Poland

*The research reported here was supported by grants from the National Institute for Deafness & Other Communicative Disorders (DC00257, DC02892) to the University of Alabama at Birmingham
First, a personal note

At the age of 56 I retired from UAB and moved to the small Italian town pictured above. Virtually no one spoke English, and I needed to interact with scores of people in order to reconstruct our house. So now I had the two magic ingredients necessary for successful L2 learning: opportunity and need. The experience of becoming bilingual at a comparatively late age has naturally caused me to reflect on the research findings we obtained over the course of 25 years in Birmingham.
Earlier is better

- There is no doubt that “earlier is better” with respect to learning a second language (L2)
- Seems especially true for speech perception and speech production

(assessed by evaluating articulation of specific vowels & consonants, overall degree of foreign accent, etc.)
Earlier is *surely* better!

It is widely believed by ordinary people, and even many “experts” that …

- Everyone who learns a second language (L2) in adulthood will speak it with a foreign accent…
Earlier is **surely** better!

It is widely believed by ordinary people, and even many “experts” that …

- Everyone who learns a second language (L2) in adulthood will speak it with a foreign accent…

- Everyone who started learning their L2 in early childhood will, sooner or later, speak it **without** a detectable foreign accent
Earlier is better!? 

Well, not always

- Careful research has yielded important exceptions
- Such evidence must be considered by those who want to understand the underlying cause(s) for the advantage of starting young
This talk has 3 aims

1. Provide *examples* of “age” effects in L2 acquisition research
2. *Outline* four hypotheses regarding “age” effects on ultimate proficiency in an L2
3. *Consider* strengths and weaknesses of each hypothesis
To begin, let’s clarify the term “age” as it is used in L2 research.

In our research in Birmingham the term “age” has been used to refer not to the chronological age of participants, but to their chronological age, in the past, when they first began to learn an L2.
Aim #1

My research typically examined:

• naturalistic acquisition arising from immersion (not study of L2 in classroom setting)

• learning that occurred after the L1 system was established (at least partially)

• groups of immigrants differing in chronological age of arrival (AOA) in a country where an L2 had to be learned
Aim #1

My own research typically examined:

- naturalistic acquisition arising from immersion (not study of L2 in classroom setting)
- learning that occurred after the L1 system was established (at least partially)
- groups of immigrants differing in chronological age of arrival (AOA) in a country where an L2 had to be learned
Aim #1

My own research typically examined:

• naturalistic acquisition arising from immersion (not study of L2 in classroom setting)

• learning that occurred after the L1 system was established (at least partially)

• groups of immigrants differing in chronological age of arrival (AOA) in a country where an L2 had to be learned
AOA

For example

- Native speakers of Italian who learned English upon immigrating to Canada
- Native speakers of Korean who learned English after immigrating to the U.S.
For samples of participants like these: the age of arrival (AOA) in the host country marks the first substantial contact with the L2 as spoken by native speakers

"AOA" = chronological age at the time L2 learning began

Note: most of our participants began to learn their L2 within several months after arriving in the host country. In research where this is not usually the case, the AOA value needs to be adjusted appropriately
Age/AOA is of not of interest in itself.

The variable is used in research because it is assumed, explicitly or implicitly, to be related to some other variable(s) that has been hypothesized to affect learning directly.
Potential “causative” variables include

- extent of neural “plasticity”
- state of development of L1 structures (e.g., phonetic categories)
- type/amount of L2 input received
AOA

age/AOA has proven useful in L2 research because

• It can be reported accurately by participants

• It is possible—although not necessarily easy—to recruit participants differing in AOA
The range of possible AOA values in an immigrant population represents a continuum.

Ages of immigrants who became legal residents of the US in 2006.
However, many published studies have examined groups of participants whose AOA falls within particular ranges.
A possible research design

Select adult immigrants having AOA values that fall within these two ranges:

- AOA = 2 to 10 years
- AOA = 15 to 23 years
Groups selected in this way might be called

- “early” L2 learners
- “late” L2 learners

**Expectation:**
L2 proficiency greater for early > late
A more complex design with $n = 24$ participants in ten AOA-defined groups

Flege et al. (1995)

<table>
<thead>
<tr>
<th>Group name</th>
<th>Mean AOA</th>
<th>AOA range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native English</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Italian-3</td>
<td>3.1</td>
<td>1.9-4.1</td>
</tr>
<tr>
<td>Italian-5</td>
<td>5.2</td>
<td>4.2-6.4</td>
</tr>
<tr>
<td>Italian-7</td>
<td>7.5</td>
<td>6.5-8.6</td>
</tr>
<tr>
<td>Italian-9</td>
<td>9.6</td>
<td>8.7-10.6</td>
</tr>
<tr>
<td>Italian-11</td>
<td>11.6</td>
<td>10.6-12.6</td>
</tr>
<tr>
<td>Italian-13</td>
<td>13.6</td>
<td>12.7-14.8</td>
</tr>
<tr>
<td>Italian-15</td>
<td>15.8</td>
<td>15.0-16.8</td>
</tr>
<tr>
<td>Italian-17</td>
<td>17.5</td>
<td>16.8-18.5</td>
</tr>
<tr>
<td>Italian-19</td>
<td>19.3</td>
<td>18.5-20.2</td>
</tr>
<tr>
<td>Italian-21</td>
<td>21.5</td>
<td>20.2-23.2</td>
</tr>
</tbody>
</table>
One aim of Flege et al. (1995) identify the AOA of the first group that differed significantly from the native English comparison group.

<table>
<thead>
<tr>
<th>Group name</th>
<th>Mean AOA</th>
<th>AOA range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native English</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Italian-3</td>
<td>3.1</td>
<td>1.9-4.1</td>
</tr>
<tr>
<td>Italian-5</td>
<td>5.2</td>
<td>4.2-6.4</td>
</tr>
<tr>
<td>Italian-7</td>
<td>7.5</td>
<td>6.5-8.6</td>
</tr>
<tr>
<td>Italian-9</td>
<td>9.6</td>
<td>8.7-10.6</td>
</tr>
<tr>
<td>Italian-11</td>
<td>11.6</td>
<td>10.6-12.6</td>
</tr>
<tr>
<td>Italian-13</td>
<td>13.6</td>
<td>12.7-14.8</td>
</tr>
<tr>
<td>Italian-15</td>
<td>15.8</td>
<td>15.0-16.8</td>
</tr>
<tr>
<td>Italian-17</td>
<td>17.5</td>
<td>16.8-18.5</td>
</tr>
<tr>
<td>Italian-19</td>
<td>19.3</td>
<td>18.5-20.2</td>
</tr>
<tr>
<td>Italian-21</td>
<td>21.5</td>
<td>20.2-23.2</td>
</tr>
</tbody>
</table>
• All (?) research has shown more native-like performance in L2 by “early” than “late” AOA learners
• True for all (?) linguistic domains, including:
  – morphosyntax
  – speech production & perception
“Age” effect: examples

Example of “age” effects in two domains:

- morphosyntax
- speech production
“age” effect: morphosyntax

– recruited 46 Chinese and Korean adults with variable AOA
– all students or faculty at Univ. of Illinois
– minimum length of residence in the US: 3 years
Johnson & Newport (1989) evaluated knowledge of 12 types of grammatical patterns representing “the most basic aspects of English sentence structure” (1989, p. 72)
Stimuli: grammatical and ungrammatical versions of 138 sentences such as:

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

*Last night the old lady **die** in her sleep
“age” effect: morphosyntax

- Participants’ task: decide if each sentence was
  
  *Grammatical* (G)
  
  *Ungrammatical* (U)

- Dependent variable: % correct
“age” effect: morphosyntax

Scores generally higher for early than late learners.
In the domain of L2 speech:

- Flege (1991) measured a well-known acoustic phonetic dimension in speech production: VOT (voice onset time)
- Measurement focused on production of word-initial /t/ tokens
“age” effect: speech

Groups of native Spanish Ss differed in AOA: “early” vs. “late”

<table>
<thead>
<tr>
<th></th>
<th>Chron. Age</th>
<th>Mean AOA</th>
<th>Length of Residence</th>
<th>English Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish monolingual</td>
<td>30</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>English monolinguals</td>
<td>26</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Early S-E bilinguals</td>
<td>23</td>
<td>2</td>
<td>21</td>
<td>82%</td>
</tr>
<tr>
<td>Late S-E bilinguals</td>
<td>34</td>
<td>20</td>
<td>14</td>
<td>66%</td>
</tr>
</tbody>
</table>
“age” effect: speech

early learners produced English /t/ like English monolinguals
Late learners produced English /t/ with VOT values that were too short (midway between Spanish, English)
Aim 1: Summary

- **Age/AOA** effects on L2 performance are widespread and robust
- However, the best explanation for these effects remains uncertain—even controversial
Next aim

Briefly evaluate each of four hypothesized causes of “age” effects on ultimate L2 proficiency
Aim #2: hypotheses

“Age” effects might arise from…

1. Normal neural maturation
2. Cognitive changes across the life span
3. Changes in how the L1 and L2 systems interact
4. Differences in the amount/kind of L2 input

OR: some combination of 2 or more factors
H1: As humans mature, some aspect(s) of development reduces the effectiveness of one or more mechanism/process that—earlier in life—assured rapid and successful L1 acquisition

Best known example of the maturational constraint hypothesis:

Critical Period (CP) hypothesis
Regarding age effects on grammaticality judgment test scores, DeKeyser concluded:

“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.”

(R. DeKeyser, 2000, 518-519)
According to Scovel (2000)

- The critical period for L2 speech learning has a “neuromotor etiology”
- An L2 will be learned less well after the critical period because of reduced cerebral “plasticity”
H1: Maturational constraints

No convergence of views as to …

• What is the neurological basis of the critical period (CP)
• When the CP ends
  - 5-6 years?
  - 12 years? 15 years?
  - different ages for different linguistic domains?
H1: Maturational constraints

However, CP advocates agree that …

• CP ends at the about the same age (= state of development) for everyone because the end of the CP is triggered by normal maturation

• Less L2 proficiency attained by those who began learning L2 after than before CP
H1: Maturational constraints

It is also agreed that …

• There will be no systematic decline in L2 proficiency as a function of age/AOA after the Critical Period has ended

• Why? Once past the CP, everyone suffers equally the ill effects of passing the CP
Now let us consider...

- Evidence for/against CP hypothesis
- Discussion will necessarily be selective and brief (time limits)
H1: Maturational constraints

The CP hypothesis correctly predicts

Early > Late

more native-like performance by early than late learners
H1: “for”

- Consider evidence from 2 studies showing Early > Late
- Both studies examined degree of foreign accent in samples of immigrants differing in AOA

Flege et al. (1995)
Flege et al. (1999)
In both studies:

- Participants began learning English as an L2 upon immigration (to Canada or US)
- 10 AOA-defined subgroups of $n = 24$ each
In both studies:

- Participants repeated English sentences after hearing a native-speaker model.
- Sentences later rated for degree of perceived foreign accent by native-English listeners.
Flege et al. (1995)

240 native Italian immigrants to Canada (black circles)

H1: “for”
H1: “for”

Flege et al. (1999)

240 native Korean immigrants to the US
H1: Maturational constraints

The CP hypothesis also correctly predicts that

No--or very few--adults who began learning L2 after the end of the critical period (12 years? 15 years?) will speak L2 without a detectable foreign accent, even after years of L2 use.
H1: “for”

% of Ss in AOA-defined subgroups of n = 24 with “accent free” pronunciation

criterion: did rating for each individual fall within 2 SDs of mean rating for native English comparison group
A number of problems exist for the CP hypothesis

Problem 1: does not predict foreign accent in early learner’s production of L1

- In the Korean study, foreign accent rated in both English and Korean sentences (Yeni-Komshian et al. 2001)
H3: “against”

Yeni-Komshian et al. (2001) ratings of

English (○)
Korean (■)

Early learners of English L2 have an foreign accent in L1 (Korean)
Problem 2: foreign accent in L2 speech of many early learners

- Foreign accent evident in early learners already considered (Flege et al. 1995, 1999)
- These early learners were adults who learned English as children and had used English daily for decades
Foreign accent also evident in early learners who are still children when tested

- Flege et al. (2006) examined Korean children who had lived in North America for 4 or 6 yrs
- All immersed in English, all attended English-medium schools
H1: “against”

Significantly lower ratings for native Korean than English children
H1: “against”

Indicated a detectable foreign accent despite an average 5 years of immersion
H1: “against”

Foreign accents still evident 1.1 years later (Time 2)
H1: “against”

For a closer look:
assigned Korean children to subgroups \((n=6\) each) based on AOA

black bars: foreign accents of AOA-defined subgroups
H1: “against”

Chronological age of AOA-defined subgroups shown in red

Chronological Age
- native Korean children ($n = 6$)
H1: “against”

next, assigned the **native English** children to subgroups (n=6 each) based on **chronological age**

ratings for the **NE children** are shown next to ratings of Korean children having similar chronological ages
H1: “against”

All groups of Korean children (black bars) received signif. lower ratings than all groups of native English children (striped bars)
H1: “against”

even Korean children who arrived in US at age 6 and had lived there for 5 years
Problem 3: post-CP declines

- CP hypothesis predicts that L2 proficiency will continue to decline as the CP is neared, but not after.
- Why? All “post-critical period” learners suffer equally the ill effects attributed to having passed the CP.
However, L2 proficiency does decline systematically after the end of the supposed CP (see work by D. Birdsong).

To see this, consider correlations between AOA and degree of foreign accent in Italian and Korean immigrants.
H1: “against”

AOA–foreign accent correlations for 240 native Italian immigrants to Canada

<table>
<thead>
<tr>
<th>AOA range</th>
<th>n</th>
<th>correlation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 to 23</td>
<td>240</td>
<td>-0.85</td>
<td>.0001</td>
</tr>
<tr>
<td>12 to 23</td>
<td>125</td>
<td>-0.53</td>
<td>.001</td>
</tr>
<tr>
<td>15 to 23</td>
<td>95</td>
<td>-0.33</td>
<td>.01</td>
</tr>
</tbody>
</table>

Flege et al. (1995)
H1: “against”

AOA – Foreign accent correlations for 240 native Korean immigrants to the US

<table>
<thead>
<tr>
<th>AOA range</th>
<th>n</th>
<th>correlation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5 to 23</td>
<td>240</td>
<td>-0.85</td>
<td>.0001</td>
</tr>
<tr>
<td>12 to 23</td>
<td>120</td>
<td>-0.49</td>
<td>.001</td>
</tr>
<tr>
<td>15 to 23</td>
<td>83</td>
<td>-0.29</td>
<td>.01</td>
</tr>
</tbody>
</table>

Flege et al. (1999)
H1: “against”

Problem 3

- Consider another example of diminished L2 performance after putative end of CP
- Flege & MacKay (unpubl.) three groups of Italian immigrants to Canada
- Had mean AOA values of 10, 18, 26 years
H1: “against”

CP correctly predicts:

higher ratings (milder foreign accent) for AOA-10

than AOA-18
H1: “against”

CP does **not** predict:

significantly higher ratings for AOA-18 than AOA-26

Why not? no one thinks the CP ends after age 18
Problem 3

Re-examination of Johnson & Newport (1989) morphosyntax data calls into question the evidence most frequently cited in the literature in support of the “maturational constraint” (aka CP) hypothesis.
J & N assumed a CP ending at 15 years 

Accordingly, divided their small sample (n = 45) into even smaller subgroups with a cutpoint at AOA = 15 giving ...
H1: “against”

Strong correlation between GJT scores and AOA for pre-CP Ss (AOA < 15)

\[ r = -0.87 \]
Strong correlation between GJT scores and AOA for pre-CP Ss (AOA < 15) but no significant correlation for post-CP Ss (AOA > 15)
Importantly, though:

simply shifting the AOA “cut point” from AOA = 15 to AOA = 20 gives a different outcome

analysis by Hakuta & Bialystok (1994)
H1: “against”

Problem 4: confounds

The “age” effect obtained by Johnson & Newport (1989) may disappear when factors confounded with AOA are controlled.
H1: “against”

- Already mentioned: Flege et al. (1999) obtained foreign accent ratings for 240 Korean immigrants
- Same participants responded to grammaticality judgment test (GJT) derived from the Johnson & Newport (1989) study
Korean participants in Flege et al. (1999)

- AOA range: 3-22 years
- minimum US residence: 8 years

Finding: the morphosyntax (GJT) scores showed a steady decrease as AOA increased
H1: “against”
H1: “against”

The AOA-morphosyntax correlation was only slightly weaker than the AOA-foreign accent correlation.

<table>
<thead>
<tr>
<th></th>
<th>Age of arrival (AOA)</th>
<th>Length of residence (LOR)</th>
<th>Use of English</th>
<th>Years of educ. in USA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>foreign accent ratings</strong></td>
<td>-0.85*</td>
<td>0.37*</td>
<td>0.61*</td>
<td>0.83*</td>
</tr>
<tr>
<td><strong>GJT scores</strong></td>
<td>-0.75*</td>
<td>0.39*</td>
<td>0.54*</td>
<td>0.78*</td>
</tr>
</tbody>
</table>
Surprisingly, the AOA-morphosyntax correlation was very similar to the correlation between morphosyntax and Education.

<table>
<thead>
<tr>
<th></th>
<th>Age of arrival (AOA)</th>
<th>Length of residence (LOR)</th>
<th>Use of English</th>
<th>Years of educ. in USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>foreign accent ratings</td>
<td>-.85*</td>
<td>.37*</td>
<td>.61*</td>
<td>.83*</td>
</tr>
<tr>
<td>GJT scores</td>
<td>-.75*</td>
<td>.39*</td>
<td>.54*</td>
<td>.78*</td>
</tr>
</tbody>
</table>
The morphosyntax (GJT) scores were also correlated with two other **potentially relevant variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Age of arrival (AOA)</th>
<th>Length of residence (LOR)</th>
<th>Use of English</th>
<th>Years of educ. in USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>foreign accent ratings</td>
<td>-.85*</td>
<td>.37*</td>
<td>.61*</td>
<td>.83*</td>
</tr>
<tr>
<td>GJT scores</td>
<td>-.75*</td>
<td>.39*</td>
<td>.54*</td>
<td>.78*</td>
</tr>
</tbody>
</table>
H1: “against”

Possible that

• The variables *correlated* with AOA were responsible for drop in the morphosyntax scores, *not* AOA …

  … and if not *age/AOA* then not state of neurological development when L2 learning began, etc, etc
H1: “against”

Let’s now consider variables that were confounded with the Koreans’ AOA.
H1: “against”

% of their lives the Koreans had lived in the US

(means for subgroups of n = 24)
H1: “against”

Years of education in English-medium US schools
H1: “against”

ratio of self-rated frequency of daily English/Korean use
H1: “against”

ratio of self-rated English/Korean use
crossover between “mostly L2” and “roughly equal” L1-L2 at AOA = 12

AOA = 12

equal use

Koreans' AOA in the US (years)
H1: “against”

ratio of self-rated proficiency in English/Korean

early better English than Korean (English-dominant)

late better Korean than English (Korean-dominant)

equal ability

Koreans's AOA in the US (years)
H1: “against”

ratio of self-rated proficiency in English/Korean

“dominance” cross-over at AOA = 12
H1: “against”

- These confounds with AOA undermine the CP hypothesis, but do not disprove it.
- Importantly, however: the “age” effect on the GJT scores disappeared when variables confounded with AOA were controlled.
Flege et al. (1999) used a “subgroup matching” technique to control for factors confounded with AOA.

- Simple and straightforward.
- However, not often used because it requires a larger sample than is typical in L2 research.
H1: “against”

3 dependent variables examined in matched subgroup analyses

• foreign accent ratings
• “rule based” subset of morphosyntax (GJT) items
• “lexically based” subset of morphosyntax items
H1: “against”

Relation between AOA and the 2 types of morphosyntax scores
H1: “against”

Step 1

• Randomly select 2 subgroups of \( n = 20 \) participants each having mean AOA values of
  – 10 years ("pre-CP")
  – 17 years ("post-CP")
H1: “against”

Step 1

• The characteristics of the two randomly selected subgroups mirrored characteristics of the entire sample of n = 240

• So necessarily: in addition to AOA, the two groups differed in …
H1: “against”

Step 1

- Years of formal education in US schools (more for “early” than “late”)
- Years of US residence (longer for “early” than “late”)
- Self-reported frequency of English use (more for “early” than “late”)
H1: “against”

Results for the **randomly selected** subgroups of n = 20 differing in AOA (10 vs. 17 years)

<table>
<thead>
<tr>
<th></th>
<th>Early</th>
<th>Late</th>
<th>F(1,38)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pronunciation</strong></td>
<td>5.9</td>
<td>3.4</td>
<td>P &lt; .01</td>
</tr>
<tr>
<td><strong>Lexically-based GJT</strong></td>
<td>92%</td>
<td>76%</td>
<td>P &lt; .01</td>
</tr>
<tr>
<td><strong>Rule-based GJT</strong></td>
<td>94%</td>
<td>85%</td>
<td>P &lt; .01</td>
</tr>
</tbody>
</table>
H1: “against”

Causality

Were the between-group differences due to “age/AOA”?

If so, then maybe the age/AOA effect could be explained – for example – by a difference in neural plasticity when L2 learning began
H1: “against”

Or: were the differences between groups having a different age of L2 learning (AOA) due to variables confounded with AOA that have nothing to do with neural maturation. Such as:

- years of education in US?
- years of US residence?
- frequency of English use?
**H1: “against”**

The **randomly selected** subgroups differed in years of education in U.S.

<table>
<thead>
<tr>
<th></th>
<th>Early</th>
<th>Late</th>
<th>F(1,38)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pronunciation</strong></td>
<td>5.9</td>
<td>3.4</td>
<td>P &lt; .01</td>
</tr>
<tr>
<td><strong>Lexically-based GJT</strong></td>
<td>92%</td>
<td>76%</td>
<td>P &lt; .01</td>
</tr>
<tr>
<td><strong>Rule-based GJT</strong></td>
<td>94%</td>
<td>85%</td>
<td>P &lt; .01</td>
</tr>
</tbody>
</table>
H1: “against”

Step 2

- Select two new subgroups of $n = 20$ each having the same mean AOA values as the groups just considered (10 vs 17 years)

- However, *match* members of both subgroups for years of education in US
H1: “against”

Step 2

- New subgroups differed in AOA (10 vs. 17) but had same years of education in US (mean = 11)
- Matching for education also eliminated significant between-group differences in
  - years of US residence
  - frequency of English use
Results for the two groups differing in AOA (10 vs. 17 years) but **matched** for education

<table>
<thead>
<tr>
<th></th>
<th>AOA = 9.7 (Educ = 10.8)</th>
<th>AOA = 16.6 (Educ = 10.8)</th>
<th>F(1,38)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pronunciation</strong></td>
<td>5.2</td>
<td>3.6</td>
<td>P &lt; .01</td>
</tr>
<tr>
<td><strong>Lexically-based GJT</strong></td>
<td>81%</td>
<td>78%</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Rule-based GJT</strong></td>
<td>87%</td>
<td>89%</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
Outcome of the matched subgroup analysis:

controlling for variables confounded with AOA eliminated the significant “age” effect obtained by Johnson & Newport (1989)
H1: “against”

One wonders:

How many “age” effects reported in the literature have been due to factors confounded with AOA?
In summary

- Evidence both for/against the CP hypothesis
- What about other hypotheses?
H2: Cognitive changes

- **H2**: As humans develop cognitively, L2 learning becomes gradually less effective across the entire life span.

Assumes: the specific cognitive abilities needed for speech and language learning diminish slowly across the life span.
H2 derived from a study by Hakuta et al. (2003)

• analyzed immigrants’ self-estimates of proficiency in English L2
• Based on responses to question on US census (very large “n”!)
H2: Cognitive changes

Hakuta et al. obtained similar findings for native Korean and Spanish immigrants:

1) self-evaluated proficiency in English was related to education level

2) proficiency in English declined over the entire life span... not just in adolescence
H2: Cognitive changes

Hakuta et al. (2003) results for native Spanish immigrants
H2: “for”

H2 correctly predicts one finding of Flege & MacKay (unpubl.)

higher ratings (milder FA) at AOA = 18

than at AOA = 26 years
H2: “against”

Problem

- **H2** unlikely to account for NE vs. AOA-10 difference

- Why? Unlikely that 10-year olds have undergone cognitive changes that reduce ability to learn speech, language
Another problem

- No direct evidence (yet) linking changes in specific cognitive abilities to ultimate proficiency in L2
- One wonders: what kind of “cognitive change” between the ages of 18 and 26 years might impact speech learning?
Hypothesis 3 - changes in the nature of how L1-L2 interact

H3: as the native language (L1) system develops, the effects of cross-language interference becomes stronger because of changes in how the L1, L2 systems interact.
H3: Changes in L1-L2 interactions

Example from speech …

The sounds used to produce words in L1 and L2 differ in both

- type (number of contrastive categories)
- details of phonetic implementation
Differences between the L1 and L2 sound systems give rise to different kinds of foreign accent in L2 (e.g., German vs. Spanish accent in English)
H3: Changes in L1-L2 interactions

example of L1 vs. L2 difference in phonetic implementation:

**VOT** (voice onset time) in Spanish, English /p t k/
H3: Changes in L1-L2 interactions

- VOT is shorter in word-initial tokens of /p t k/ in Spanish than English

- Holds true for monolingual adults

Flege & Eefting (1987)
H3: Changes in L1-L2 interactions

Also for monolingual children

Flege & Eefting (1987)
H3 assumes

- that L1 categories develop slowly, at least into early adolescence
- If so: likely that L1 phonetic categories are more robust for monolingual 15-than 5-year-olds (shown, e.g., by studies examining the recognition of words in noise)
According to H3, as L1 categories develop

- learners become less likely to create new categories for L2 sounds that are similar but non-identical to corresponding L1 sounds
- Why? L1 sounds become stronger “attractors” of similar L2 sounds
H3: Changes in L1-L2 interactions

Example …

- As Spanish speakers get older, the Spanish /t/ category becomes a stronger “attractor” for English /t/ tokens
H3: Changes in L1-L2 interactions

If so, then as Spanish speakers get older

- the rated dissimilarity of English [tʰ] and Spanish [t] tokens should decrease

- the frequency of detections of English-like VOT in Spanish speech samples should decrease (or RTs increase)
H3: Changes in L1-L2 interactions

Hypotheses of the Speech Learning Model (e.g., Flege, 2002, 2003)

1. Learners of all ages remain able to auditorily detect acoustic phonetic differences between the /t/ of Spanish, English
2. As the Spanish /t/ category develops, Spanish learners of English will be less likely to create a new category for English /t/

3. If category formation is blocked, Spanish learners of English will create a **merged** (composite) /t/ category

this **merged category** will **blend** phonetic properties of Spanish and English /t/
H3: Changes in L1-L2 interactions

Predictions:

• **early learners** will establish a **new** category for English /t/, and so produce English /t/ correctly*

• **late learners** will develop a **merged** category (sharing properties of /t/ in L1, L2), leading to incorrect production of /t/ in both English and Spanish*

*more exactly: the SLM hypotheses state that the earlier/later the exposure to L2, the greater/lesser the likelihood…*
These predictions supported by production data (e.g., Flege, 1991)
H3: “for”

VOT in English /t/ as produced by Spanish early & late learners (Flege, 1991)

due to formation of new category?
due to use of merged category?
H3: “for”

Same predictions for

- Late French learners of English
- Late English learners of French

Similar VOT difference between French-English and Spanish-English

French /t/ short-lag VOT
H3: “for”

Flege (1987) tested:

• native French women who had lived in Chicago for 10 yrs

• native English (American) women who had lived in Paris for 10 yrs
H3: “for”

French women produced English /t/ with VOT between French, English
American women produced **French /t/** with VOT between English, French
H3: “for”

for both American and French late learners:

VOT in production of L1 /t/ shifted towards values appropriate for L2 /t/
Problem: little research has tested…

- If the perceived dissimilarity of L1-L2 sounds actually does decrease as L1 categories develop

- Whether such changes—if they occur—predict accuracy of L2 segmental production and/or perception (but see Baker et al., 2002)
Hypothesis 4: “Age” effects arise because early learners get more and/or better L2 input than late learners do

- H4 probably the most obvious explanation for age effects, but the least well accepted even though, for immigrants, abundant evidence that...
H4: Differences in L2 input

The later immigrants’ AOA…

• the less often L2 tends to be used
• and the more often L1 continues to be used

**Probably** also true that late learners hear L2 produced with a foreign-accented more often than early learners—but no hard evidence
H4: “for”

H4 correctly predicts

• effects of language use on L2 proficiency
• Examples: speech research with Italian immigrants to Canada

H4: “for”

Participants differed according to AOA: early vs. late

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Age</th>
<th>AOA</th>
<th>Italian Use</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NE</strong></td>
<td>18</td>
<td>50</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>early-low</strong></td>
<td>18</td>
<td>50</td>
<td>7</td>
<td>7%</td>
</tr>
<tr>
<td><strong>early-high</strong></td>
<td>18</td>
<td>49</td>
<td>8</td>
<td>43%</td>
</tr>
<tr>
<td><strong>late-low</strong></td>
<td>18</td>
<td>51</td>
<td>20</td>
<td>10%</td>
</tr>
<tr>
<td><strong>late-high</strong></td>
<td>18</td>
<td>49</td>
<td>20</td>
<td>53%</td>
</tr>
</tbody>
</table>
H4: “for”

and also % continued L1 use: low vs. high

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Age</th>
<th>AOA</th>
<th>Italian Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>NE</td>
<td>18</td>
<td>50</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>early-low</td>
<td>18</td>
<td>50</td>
<td>7</td>
<td>7%</td>
</tr>
<tr>
<td>early-high</td>
<td>18</td>
<td>49</td>
<td>8</td>
<td>43%</td>
</tr>
<tr>
<td>late-low</td>
<td>18</td>
<td>51</td>
<td>20</td>
<td>10%</td>
</tr>
<tr>
<td>late-high</td>
<td>18</td>
<td>49</td>
<td>20</td>
<td>53%</td>
</tr>
</tbody>
</table>
H4: “for”

Note:

inverse relation between L1-L2 use in these bilinguals, so the groups differing in % Italian use necessarily also differed in % English use
General finding:

- significantly less proficiency in English for participants who used English relatively seldom

this held true for …
H4: “for”

% correct identification of word-final tokens of English /b d g p t k/

**Low** use of Italian = frequent use of English

**High** use of Italian = infrequent use of English
categorial discrimination of 3 pairs of English vowels

(those in beat-bit, bet-bat, bought-but)
H4: “for”

accuracy in producing the same 6 English vowels

(those in beat, bit, bet, bat, bought, but)
H4: “for”

overall pronunciation (global foreign accent ratings)
These findings discredit claims (e.g., DeKeyser, 2000, p. 519) that input differences do not contribute to “age” effects on L2 speech learning.
Problem: unclear to what extent input differences contribute to age-related effects on L2 speech

- No study has quantified L2 input
- Research has relied on self-reports (e.g., overall % L1 or L2 use); these are estimates, not measurements
Another problem: input differences are unlikely to account for some AOA differences

• Consider foreign accent in Italian immigrants having AOA of 10, 18, 26 years (Flege & MacKay, unpubl.)
H4: “against”

H4 (input differences) might account for difference between

AOA 10

and

AOA 18
H4: “against”

but not between

AOA 18 and AOA 26
H4: “against”

Why? language use differences existed between AOA 10-18 but not AOA 18-26

<table>
<thead>
<tr>
<th>Age of Arrival</th>
<th>10</th>
<th>18</th>
<th>26</th>
<th>10&gt;18,26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of residence</td>
<td>43</td>
<td>37</td>
<td>33</td>
<td>10&gt;18,26</td>
</tr>
<tr>
<td>Use of English</td>
<td>71%</td>
<td>53%</td>
<td>47%</td>
<td>10&gt;18,26</td>
</tr>
<tr>
<td>Yrs of Educ in Canada</td>
<td>9</td>
<td>0.4</td>
<td>0.3</td>
<td>10&gt;18,26</td>
</tr>
<tr>
<td>Use of Italian (IT)</td>
<td>27</td>
<td>47</td>
<td>52</td>
<td>10&lt;18,26</td>
</tr>
<tr>
<td>Yrs of Educ in Italy</td>
<td>4</td>
<td>8</td>
<td>8</td>
<td>10&lt;18,26</td>
</tr>
<tr>
<td>Ability to pronounce IT</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>10&lt;18,26</td>
</tr>
</tbody>
</table>

Flege & MacKay, in prep
General conclusions

1. “Age” effects are robust, but...
   
   - May not be due to chronological age at the time L2 learning begins
   - Should be referred to as “age-related” effects
2. Of the four hypotheses considered, all had possible merit but none was perfect

- Possibility: all four contribute, to varying degrees, to age-related effects
- Possibility: some other factor(s) (motivation?) plays a role
3. More research is needed--current research is inadequate
   • Research should be designed so that confounded factors can be eliminated \textit{a priori} or can be controlled statistically
   • Necessary to directly evaluate factors thought to affect “ultimate” L2 learning rather than simply vary age/AOA
Conclusions

Among variables that should be examined:

- Specific aspect(s) of neurological development at time L2 learning begins
- Specific aspects of cognitive development
- State of development of L1 categories
Conclusions

Among variables that should be examined

- Age-related changes in perceived relation between sounds in L1, L2
- Amount of L2 input (measured, not estimated)
- Kind of L2 input (native-speaker? foreign-accented?)
“Take home” message

Until potential causal variables are examined in L2 research and confounded variables have been controlled, we can only speculate about the true basis (bases) for age-related differences in ultimate L2 proficiency.
End

Thanks for your attention

Supported by NIH grants DC00257 & DC02892
References


References


References


• Flege, J., MacKay, I. (unpubl.) What accounts for “age” effects on overall degree foreign accent?


References


