Evidence for plasticity in studies examining second language speech acquisition

James E. Flege
Univ. of Alabama at Birmingham

ISCA Workshop on Plasticity in Speech Perception
15-17 June 2005, London, UK

Outline
1. The “Doom” (no plasticity) hypothesis
2. The “Full Access” hypothesis
3. Factors affecting L2 speech perception & production
4. Access to the properties of L2 speech sounds
5. Plasticity

The Doom hypothesis
• Most phonetic segments in an L2 differ phonetically from those in the native language (L1)
• L1-L2 phonetic differences varies
  --substantial differences (“new” L2 sounds)
  --smaller differences (L2 sounds have L1 “counterpart”)

The Doom hypothesis
Hypothesis
“Individuals who have become perceptually attuned to the L1 sound system can not gain perceptual access to L1-L2 phonetic differences”
The Doom hypothesis

Consequences
- Lack of sensory/perceptual access to L1-L2 differences prevents development of accurate long-term memory representations for L2 sounds
- In turn: prevents accurate production of L2 phonetic segments

1. Phonological grid
- The phonemes of the L2 are heard as phonemes of the L1 (“new wine in old bottles”)
- It difficult or impossible to learn to produce/perceive phonemic distinctions based on features not used to contrast phonemes in L1

2. Categorical perception
- From the perspective of the L1 phonetic system, many L1-L2 phonetic differences represent within-category phonetic variation
- Example: Initial /t/ may be realized in Spanish and English with VOT values of 20 and 65 ms, respectively

2. Categorical perception
- Easier to discriminate between- than within-category differences
- Example: VOT of word-initial English /t/ varies widely as a function of degree of stress, vowel context, speaking rate
- Native English speakers tend to “filter out” VOT differences in /t/ tokens (e.g., 20 vs 65 ms)

Consequences
- L1-L2 phonetic differences get “filtered out”
- Even if detected at sensory level, not used to develop accurate perceptual representations
- Inaccurate L2 segmental production
The Doom hypothesis

also related to

3. The Critical Period hypothesis (Lenneberg)

3. CP hypothesis

- Lenneberg (1969) proposed a critical period for L1 acquisition
- extended to L2 because of the observation that most people learning an L2 after the age of 12 speak the L2 with foreign accent
- Adults unsuccessful in L2 speech learning because of lost neural “plasticity”

Consequences

The “filtering out” of L1-L2 phonetic differences may be more severe for adult than child L2 learners

The Speech Learning Model (SLM)

Hypothesis

The processes and mechanisms that guide successful L1 speech acquisition—including the ability to establish new phonetic categories—remain intact across the life span

Outline

1. The “Doom” (no plasticity) hypothesis
2. The “Full Access” hypothesis
3. Factors affecting L2 speech perception & production
4. Access to the properties of L2 speech sounds
5. Plasticity

The Speech Learning Model (SLM)

Hypothesis

Adults & children can eventually gain access to the properties needed to differentiate L2 speech sounds (meaning) even if such properties are not used to differentiate L1 speech sounds

the “Full Access” hypothesis
The Speech Learning Model (SLM)

Limitation

the SLM does not specify

• How long it takes L2 learners to “gain access” to phonetic properties of L2 sounds
• If more time/input is needed for certain properties than others

The Full Access Hypothesis

Motivation from two sources

• Research with monolinguals hearing L1 sounds
• Research with monolinguals hearing foreign speech or foreign-accented L1 speech

The Full Access Hypothesis

• For monolingual adults, within-category phonetic information is accessible
  “... although the retrieval of this information will depend on the level of processing ...”  
  (Pisoni & Tash, 1974)
• Example: Werker & Logan (1985)

Flege (1984)

Stimuli: English speech samples produced by native English and French adults in Chicago:

• /tu/ syllables (word “two”)
• Hybrid /tu/ syllables (cross-splice /t/ or /u/ segments from original /tu/ to single, native-produced segment (/u/ or /t/)
• First 30-ms of /t/ (release bursts)

Flege (1984)

Aim: determine if native English adults can detect within-category differences in French-accented English

Flege (1984)

• Stimuli presented in pairs
  – One produced by native English speaker
  – One produced by native French speaker
• Listeners’ task: decide which of 2 stimuli was “foreign”
• No training or feedback
Flege (1984)

English monolingual could detect
• “backness” (non-fronting) of French-accented /u/
• Too-short VOT values in French-accented /t/
• Dental place in French-accented /t/

Interpretation

English monolingual adults can detect within-category differences distinguishing English phonetic segments produced by
• native English speakers
• French speakers of English

Flege (1984)

Interpretation

Therefore …
Shouldn’t a native English adult who learns French as an L2 be able to detect even larger English-French phonetic differences? … even if they are “within-category”?

The Full Access Hypothesis

But …
• Flege (1984) used sensitive paired comparison task
• Perhaps adults cannot detect small L1-L2 phonetic differences in the “real world” when attending to meaning

The Full Access Hypothesis

Flege & Hammond (1982)

Aim: determine if monolingual English adults can detect and store within-category phonetic differences due to foreign accent through everyday exposure to foreign-accented English
Flege & Hammond (1982)

- Tested 50 native English college students
- Enrolled in 1st year Spanish classes at University of Florida
- Classes taught in English by Spanish-accented teachers

Task: Try to produce English words inserted into a carrier phrase (The__is on the__) with a “Spanish accent”

<table>
<thead>
<tr>
<th>Lexical Items</th>
<th>Substitute</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>nose, cheese, bone</td>
<td>s/z</td>
<td>141 (47%)</td>
</tr>
<tr>
<td>vice, red, raw</td>
<td>s/z</td>
<td>129 (43%)</td>
</tr>
<tr>
<td>fig, pig, wig</td>
<td>s/z</td>
<td>127 (42%)</td>
</tr>
<tr>
<td>hook, hook, crook</td>
<td>s/z</td>
<td>61 (20%)</td>
</tr>
<tr>
<td>shell, sheet, sheep</td>
<td>s/z</td>
<td>49 (16%)</td>
</tr>
<tr>
<td>bean, phone, bone</td>
<td>s/z</td>
<td>1 (0.3%)</td>
</tr>
<tr>
<td>rape, tube, road</td>
<td>s/z</td>
<td>0 (0.0%)</td>
</tr>
</tbody>
</table>

Flege & Hammond (1982)

- Spanish-accented English /t/ often has VOT midway between
  - Spanish short-lag
  - English long-lag
- This effect of cross-language phonetic interference occasionally causes Spanish-accented /t/ to be heard as /d/

Flege & Hammond (1982)

- The students never realized /t/ as /d/ when imitating a Spanish accent
- Did they fail to detect VOT shortening in Spanish-accented /t/?

  Acoustic analysis suggest otherwise

Flege & Hammond (1982) conclusion

- As knowledge of Spanish-accented English developed, within-category phonetic variation was detected and stored in memory
- VOT finding not an artifact of combining Spanish short-lag + English long-lag values (frequency histogram)
L2 speech perception & production

- Individuals first exposed to L2 as children (early learners) more native-like than individuals first exposed as adults (late learners)
- However, some late learners in native range
- Example: Flege (1991)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Age</th>
<th>AOA</th>
<th>Length of Residence</th>
<th>English Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish monolingual</td>
<td>10</td>
<td>30</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>English monolingual</td>
<td>10</td>
<td>26</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Early S-E bilinguals</td>
<td>10</td>
<td>23</td>
<td>2</td>
<td>21</td>
<td>82%</td>
</tr>
<tr>
<td>Late S-E bilinguals</td>
<td>10</td>
<td>34</td>
<td>20</td>
<td>14</td>
<td>66%</td>
</tr>
</tbody>
</table>
Flege (1991) participants

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Age</th>
<th>AOA</th>
<th>Length of Residence</th>
<th>English Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish monolingual</td>
<td>10</td>
<td>30</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>English monolinguals</td>
<td>10</td>
<td>26</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Early S-E bilinguals</td>
<td>10</td>
<td>23</td>
<td>2</td>
<td>21</td>
<td>82%</td>
</tr>
<tr>
<td>Late S-E bilinguals</td>
<td>10</td>
<td>34</td>
<td>20</td>
<td>14</td>
<td>66%</td>
</tr>
</tbody>
</table>

Flege (1991)

word-initial /t/ (utterance initial)

Mean VOT (ms)

Participant (ascending order, by group)

Flege (1991)

word-initial /t/ (utterance medial)

Mean VOT (ms)

Participant (ascending order, by group)

L2 speech perception & production

- Language use affects L2 performance of both early and late bilinguals
- Example: studies examining 72 Italian immigrants to Canada
  - Piske, MacKay & Flege (2001)
  - MacKay, Meador & Flege (2001)
### Participants

<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>

### Average identification of stops (/b d g p t k/) in Final Position

<table>
<thead>
<tr>
<th></th>
<th>% Correct Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td>Late</td>
</tr>
<tr>
<td>Native English mean</td>
<td>80</td>
</tr>
</tbody>
</table>

### Categorial Discrimination of /i/–/e/ and /a/–/u/ (Mean Discrimination A')

<table>
<thead>
<tr>
<th>Type</th>
<th>Early</th>
<th>Late</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native English mean</td>
<td>0.50</td>
<td>0.75</td>
</tr>
</tbody>
</table>

### Overall FA

<table>
<thead>
<tr>
<th>Type</th>
<th>Early</th>
<th>Late</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native English mean</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
L2 speech perception & production

Whatever the explanation …
No general conclusions regarding L2 plasticity – or lack thereof – can be reached without considering quantity and quality of L2 input

Access to L2 sounds

McAllister, Flege & Piske (2002)

Aim: determine if adults learn to perceive and produce phonetic distinctions based on a feature not used to contrast meaning in L1

Outline

1. The “Doom” (no plasticity) hypothesis
2. The “Full Access” hypothesis
3. Factors affecting L2 speech perception & production
4. Access to the properties of L2 speech sounds
5. Plasticity

L2 speech perception & production

Basis for “L1 use” effect?
1. Not enough L2 native-speaker input for low-L1-use group?
2. Too much Italian-accented English input?
3. When L1 system is stronger, exerts stronger effect on the L2?

McAllister et al. (2002)

• Examined production & perception of Swedish vowels
• Participants
  – 20 native English adults
  – 20 native Spanish adults
• All experienced speakers of Swedish (> 10 years in Stockholm)
McAllister et al. (2002)

Focused on distinctions between long-short Swedish vowels (4 pairs)
• mid vowels – almost exclusively duration
• high & mid vowels – duration differences accompanied by spectral quality differences

McAllister et al. (2002)

Stimuli
• naturally produced Swedish words containing long or short vowels
• versions of same stimuli with modified quantity (long → short, short → long)
• Modifications created non-words

McAllister et al. (2002)

[Length] not used to distinguish vowel phonemes (meaning) in English or Spanish
• 20 Estonians included as control group
• [length] not used to distinguish Estonian vowel phonemes

McAllister et al. (2002)

Participants produced real-word stimuli after hearing a definition (e.g., say “castle” after being cued by “place where a king and queen live”)
• Later, original & modified stimuli presented in random order following the same definitions
• Perceptual task: decide if each stimulus was “correct”

McAllister et al. (2002)

Separate analyses for
• mid vowel pairs (contrasts based almost entirely on duration)
• high and low vowel pairs (contrasts based on duration & spectral quality)
McAllister et al. (2002) conclusions
- Estonians benefited from the presence of a [length] feature in L1
- English & Spanish Ss did better when Swedish long-short contrasts accompanied by feature (spectrum) used in L1
- They demonstrated difficulty using new (non-L1) feature

McAllister et al. (2002) conclusions
• However, the above-chance performance of most native English & Spanish Ss – and excellent performance of some English Ss – indicated that adult L2 learners can acquire sensitivity to a new feature ([length])

McAllister et al. (2002) • Also measured duration of vowels
• Production results for /øː/-/ø/ much like perceptual results for same contrast

Proceedings of ISCA Workshop on Plasticity in Speech Perception (PSP2005); London, UK; 15-17 June 2005
Access to L2 sounds

- English /r/ very dissimilar from any Italian vowel
- Distinguished from other English vowels by a feature ([rhotic]) not used to distinguish Italian vowels

Aim of 3 studies: determine if Italian adults can learn to use the rhotic feature
- Munro, Flege & MacKay (1996)


Perceptual assimilation experiment
- Classifications English /r/ in terms of Italian vowel categories less consistent than for other English vowels
- Lower goodness of fit ratings than for any other English vowel

Munro et al. (1996)

- 240 Italian immigrants
  - Differed in age of arrival
  - Many years of residence in Canada
- Delayed repetition of English words
- Vowels rated by listeners, 5-point scale

Munro et al. (1996)

- Mean Age of Arrival in Canada (years)
  - /r/ ratings
  - Correctly produced vowel, no FA
  - Very strong FA
  - Wrong vowel
Munro et al. (1996)

N per group (max = 24) whose /æ/ received a rating within 2 SD of NE mean

Mean Age of Arrival in Canada (years)

Flege et al. (2003)

Examined /æ/ produced by 5 groups (n=18 each)
- Native English
- 2 groups of Italian early learners matched for AOA in Canada (7 years) but differing in amount of Italian use (7% vs. 43%)
- 2 groups of late learners (AOA = 20) differing in Italian use (10% vs. 53%)

Flege et al. (2003)

Mean ratings of /æ/ production

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>correct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>acceptable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>distorted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wrong vowel</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Early bilinguals

Late bilinguals

Flege et al. (2003)

- Dependent variables
  - mean ratings (parametric)
  - N of listeners who gave “good/acceptable” judgments (non-parametric)
- In both analyses
  NE > late but not early learners
- Perhaps “no new feature” ([rhotic]) applies just to late learners

Flege et al. (2003)

- Acoustic analysis, however, suggested that late learners do acquire sensitivity to [rhotic] feature
- Examined Bark-transformed F3-F2 differences: index of [rhotic]

F3-F2, Bark transforms

(index of rhotic dimension)

English monolinguals
Early low
Early high
Late low
Late high
Italians in Italy
Flege et al. (2003)

- Late bilinguals produced significantly larger F3-F2 differences (less “rhotic” vowels) than early bilinguals

However, late bilinguals produced significantly smaller F3-F2 differences (more “rhotic” vowels) than Italian monolinguals

- Indicated that late bilinguals acquired sensitivity to [rhotic] feature
- A few late learners were native-like

/r/-/l/

- Word-initial English liquids /r/ and /l/ notoriously difficult for Japanese adults
- both English liquids differ phonetically from any Japanese consonant, including the single Japanese liquid /R/
- However, both /r/ and /l/ tend to be heard as instances of Japanese /R/

Difficulty of English /r/ and /l/ may be due to
- Failure to use a feature/property needed to distinguish English /r/-/l/
- Continued use of inappropriate L1 feature/property
- Both

Iverson et al. (2003)

Figure 1: Formant frequencies for the English /r/ and /l/ sounds used in this study (from Flege & Todd, 1994). The sounds varied in terms of the second (F2) and third (F3) formants during the initial consonant. The formant frequencies were spread equally using the MLS scale (representing frequency and pitch).

Word-initial English liquids /r/ and /l/ notoriously difficult for Japanese adults
both English liquids differ phonetically from any Japanese consonant, including the single Japanese liquid /R/
However, both /r/ and /l/ tend to be heard as instances of Japanese /R/
Japanese adults made insufficient use of F3 to distinguish /r/-/l/
• Too much use of F2

Will perceptual difficulties identified by Iverson et al. (2003) extend to all Japanese adults with little/no English-language experience?
Perhaps not

Yamada (1991)
• Tested 152 undergraduates in Japan
• As for Ss in Iverson et al. (2003), had never lived outside of Japan
• Stimuli: 4 naturally produced non-words pairs (e.g., rosti-losti)
• 1 hour of identification (2AFC) training

Substantial individual variation
• 2/152 could differentially identify /r/ and /l/ before training
• 9 others (6%) did so in all 4 contrasts by end of training
• 53 others (35%) did so for at least 1 of 4 pairs
• However, 77 (51%) remained at chance

MacKain et al. (1981)
• Tested groups of Japanese adults who differed in
  – use of English (29% vs. 55%)
  – Years of English use in the US (0.7 vs. 2.3 years)
• Synthetic /r/-/l/ continuum
• Identification & discrimination tests
MacKain et al. (1981)

- 6/7 inexperienced Japanese Ss showed “near-chance performance” on both tasks
- All 5 experienced Ss resembled NE in showing “categorical perception” of English word-initial /r/-/l/

Flege et al. (1996)

Tested groups of Japanese adults (n = 12 each) who differed in years of residence in the US
- Inexperienced  1 - 4 years
- Experienced    12 - 29 years

Flege et al. JASA

Identification of equally familiar minimally paired English words beginning in “r” and “l”

Why was /r/ identified better than /l/?

- According to the Speech Learning Model (Flege 1995, 1999, 2002, 2003), the likelihood of a new category formation increases with perceived L1-L2 phonetic distance
Japanese speakers identify initial English /r/ and /l/ tokens as /R/.
- But /r/ is rated as less /R/-like than /l/.
- The /r/-/l/ rating difference increases as a function of length of residence in the US (Aoyama & Flege, under review).

Perhaps, with English-language experience, some Japanese adults begin to perceive the phonetic properties that distinguish /r/ from /R/, and so establish a new category for it.

Outline
1. The “Doom” (no plasticity) hypothesis
2. The “Full Access” hypothesis
3. Factors affecting L2 speech perception & production
4. Access to the properties of L2 speech sounds
5. Plasticity

Plasticity
- Much inter-subject variability L2 learners, especially late learners (e.g., Hazan et al. 2002)
- Many different explanations have been offered
- Poorly understood

My hunch
Most inter-subject variability is due to variation in the quantity and/or quality of L2 input received
1. If L2 learners are strongly motivated to speak L2 well, they will receive much native-speaker input

maybe what are thought of as “motivational” differences are really input differences

2. If bilinguals use the L2 often, they necessarily use the L1 seldom

This may reduce the degree of L1 activation

and so reduce L1 \rightarrow L2 interference

Plasticity

If some some early learners differ from native speakers, but other early learners show native-like performance, not possible to make general statements about plasticity regarding early learners

True also for late learners

End

• Thanks for your attention


