Origins and development of the Speech Learning Model*

James E. Flege
Univ. of Alabama at Birmingham

Keynote lecture presented at the
1st ASA Workshop on L2 Speech Learning
Simon Fraser Univ., Vancouver, BC
April 14-15, 2005

**corrected version*

Outline

1. Purpose of the Speech Learning Model (SLM)
2. Historical background
3. Core aspects of the SLM
4. Some testable predictions generated by the SLM
5. How to falsify the SLM
6. What is most needed now?

Outline

1. **Purpose of the Speech Learning Model (SLM)**
2. Historical background
3. Core aspects of the SLM
4. Some predictions generated by the SLM
5. How to falsify the SLM
6. What is most needed now?

Primary aim of the SLM

- The aim of the SLM is to account for variation in the extent to which individuals learn – or fail to learn – to accurately produce and perceive phonetic segments (i.e., vowels & consonants) in a second language (L2)

SLM Purpose

Our research has focused on questions such as:

- Is it impossible for learners of an L2 to produce certain L2 speech sounds accurately? Are there “un-learnable” L2 sounds?
- Do un-learnable L2 sounds exist, but only for adults?
- How does the perception of speech sounds encountered on the phonetic surface of an L2 influence their eventual production?

Outline

1. Purpose of the Speech Learning Model (SLM)
2. **Historical background**
3. Core aspects of the SLM
4. Some predictions generated by the SLM
5. How to falsify the SLM
6. What is most needed now?

Zeitgeist: 1979-1984

To understand the SLM, it is important to consider the state of research in the first 5 years of development of what later became known as the SLM. Some topics to be considered today are:

- Critical Period (CP) hypothesis
- Contrastive Analysis (CA) hypothesis
- Focus on abstract linguistic units, not phonetic substance
- Categorical Perception/filtering via the phonological "grid"
- Unidirectional L1→ L2 interference

Critical Period Hypothesis (CPH)

- Eric Lenneberg (1969) proposed that as humans mature neurologically, their ability to learn the L1 diminishes. Among the evidence he cited was the effect of severe deprivation on L1 development (e.g., “wolf” children)
- Lenneberg casually observed that it is difficult to learn to pronounce an L2 without a foreign accent after the age of 12 years (“puberty”)
- This casual remark triggered extension of CPH from L1, Lenneberg’s focus, to L2 learning

Doubts re: CPH

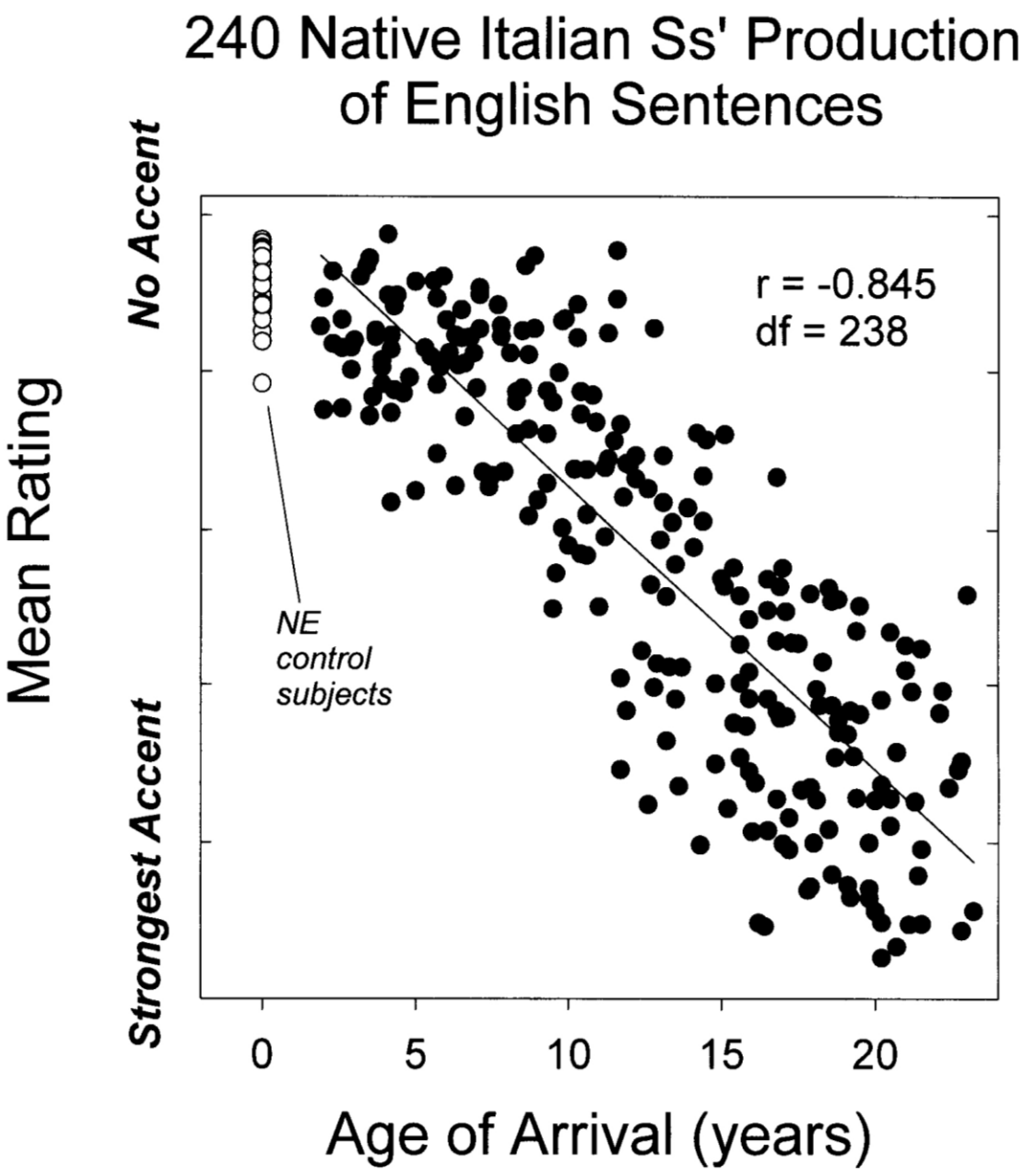
- It is certainly true that “earlier is better” as far as learning an L2 is concerned, especially the phonetic & phonological characteristics of an L2
- However, in the early 1980s no one had produced objective evidence for:
  1. A sharp drop in L2 learning success at a particular chronological age (say, 12 years), neuro-endocrinological status (e.g., “puberty”), or state of neurological development
  2. Success in L2 learning for <u>all</u> children
  3. Failure for <u>all</u> adults

Doubts re: CPH

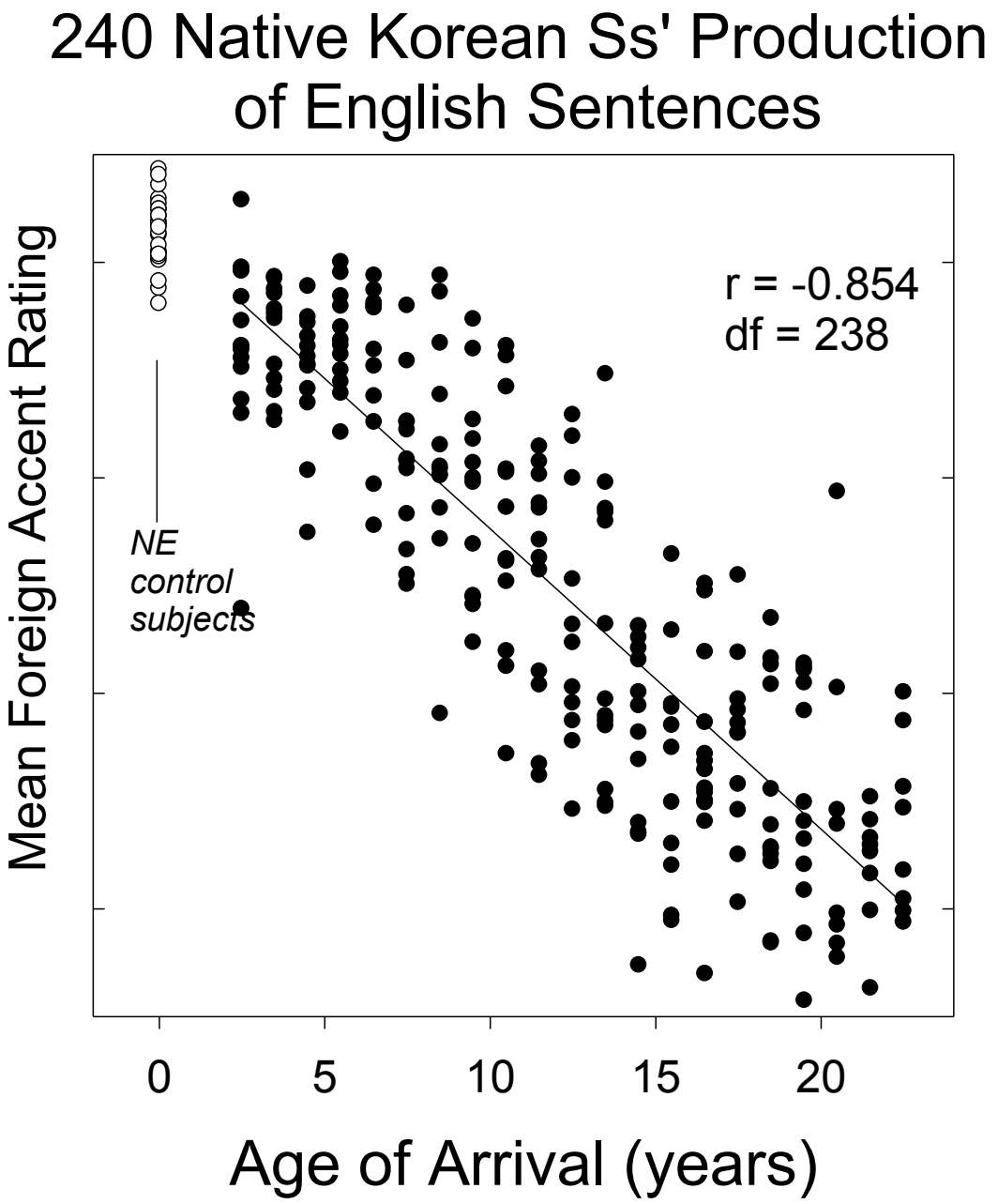
Subsequent empirical research failed to sustain predictions generated by the CPH

- We will now consider 2 studies that examined relatively large (n = 240) groups of immigrants to North America
- All participants (Ss) were adults at the time of testing
- All were selected of the basis of their age of arrival (AOA) in a country where an L2 had to be learned for everyday use
- The Ss recorded English sentences that were later rated for overall degree of perceived foreign accent (FA) by native English-speaking listeners

Results of a study (Flege et al., 1995) examining degree of foreign accent (FA) in English sentences spoken by native Italian immigrants to Canada, selected on the basis of AOA. A continuous scale was by native English-speaking listeners to estimate FA. Scores obtained for the native English control group are shown with unfilled circles



Results of a similar study by Flege et al., 1999. This study examined FA in English sentences spoken by native Korean immigrants to the U.S. As in the earlier study, the Ss were selected on the basis of AOA. A 9-point rating scale was used by native English-speaking listeners to estimate FA.

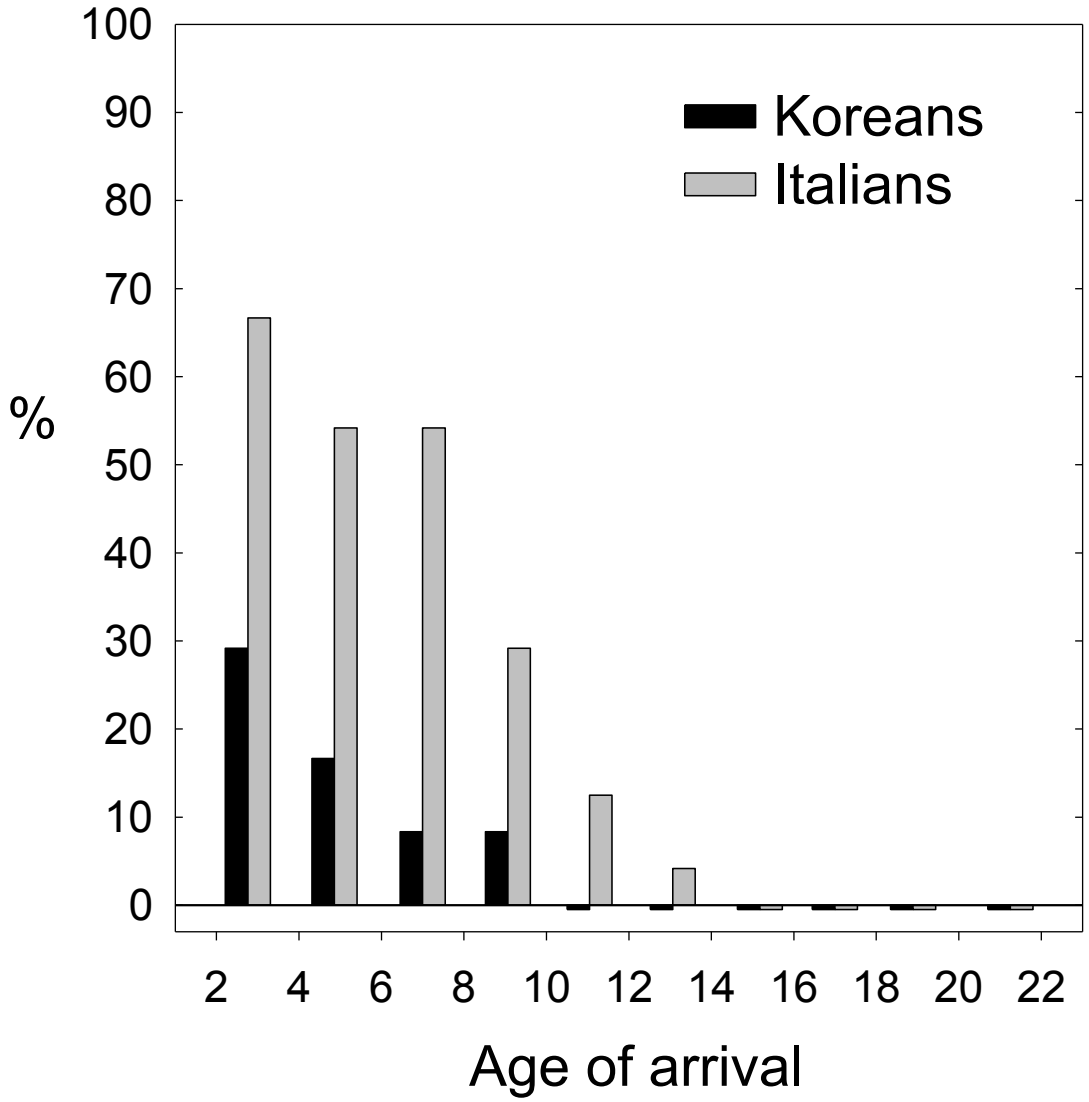


Index of "accent free"

- Lenneberg (1969) had cited the presence of foreign accent (FA) as evidence that successful L2 learning might be limited by a critical period.
- We naturally wondered if any of our Italian or Korean Ss had managed to produce our English test sentences without an obvious FA.
- We used a statistical method as a first approximation to answering this question.
- Specifically, we determined if the sentences spoken by each non-native S had obtained a mean rating that fell within 2 SDs of the mean rating obtained for the 24 Ss in the native English control group

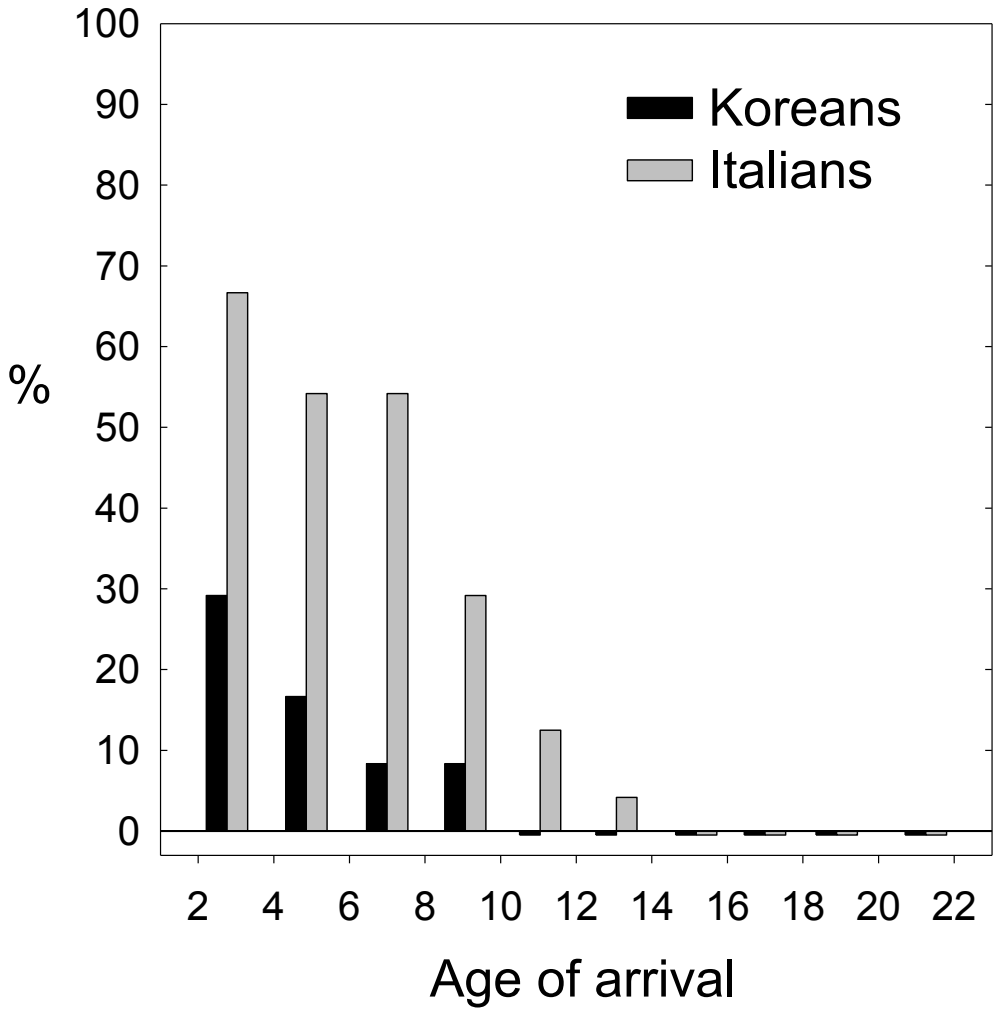
Post-hoc analysis of FA ratings obtained by Flege et al., 1995 and 1999. The percentage of Ss in each non-native group of n = 24 whose English sentences obtained a rating that fell within 2 SDs of the mean rating obtained for a control group (n = 24) of native English speakers. Non-native Ss meeting this criterion were deemed to be «accent free»

% of immigrants (n = 24 per AOA group) who spoke English w/o foreign accent



The results showed that few Ss who began learning the L2 after the age of 12 years could be said to speak it without FA, even after decades of immersion. This supported the CPH. However less than half of the s who began learning their L2 prior to the end of the «critical period» were unable to meet our fairly lax criterion of «no foreign accent». This finding clearly diverges from what the CPH predicts

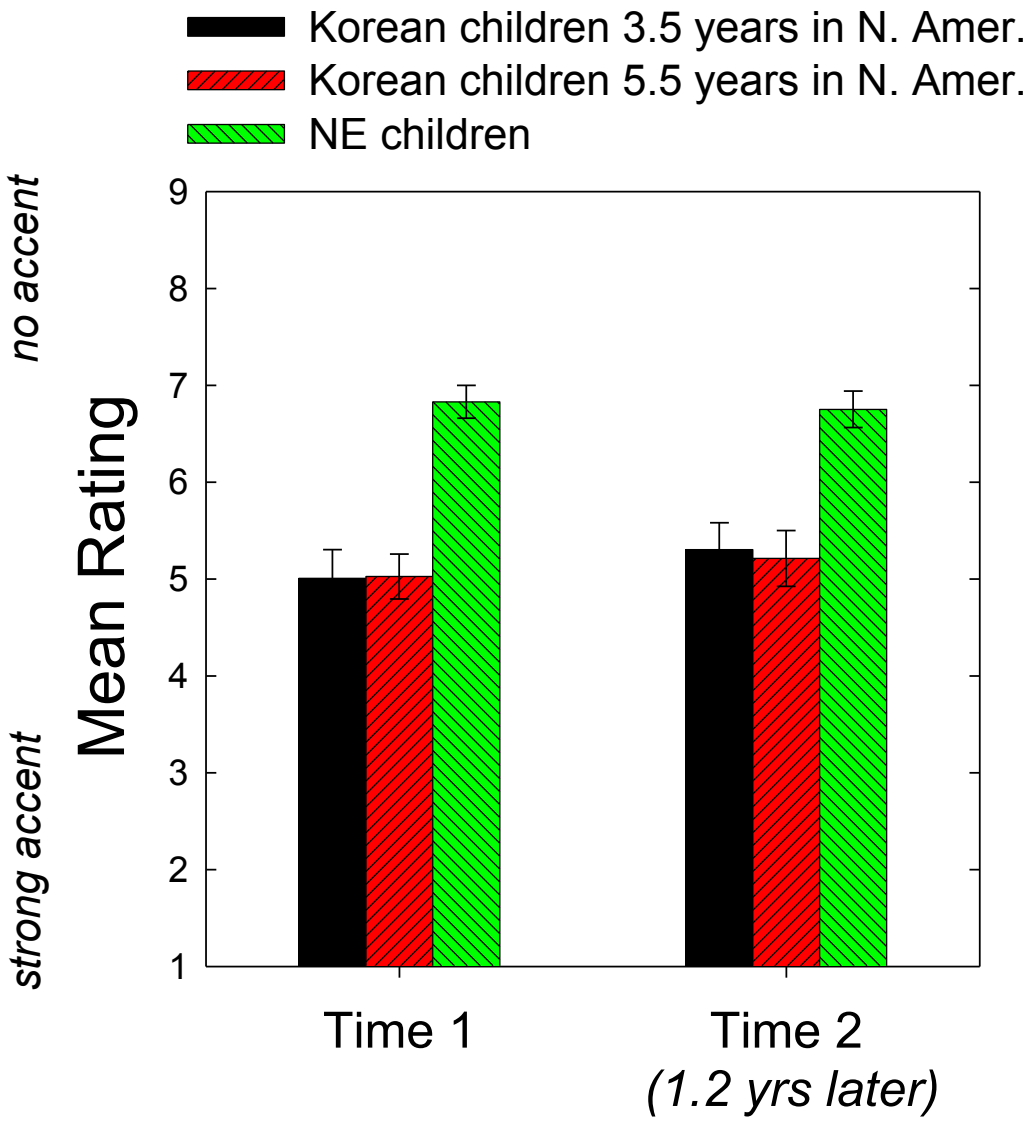
% of immigrants (n = 24 per AOA group) who spoke English w/o foreign accent



Doubts re: CPH

- Flege et al. (in press)
- Evidence of foreign accent in children currently learning an L2

The results obtained by Flege et al. (in press). The Ss were native Korean children who had been living in North America for either 3.5 or 5.5 years as well as age-matched native English speaking children who were born and raised in an English-speaking community. FA in English sentences was rated using a 9-point scale. All Ss were tested twice.



Flege et al. (in press)

An important finding of this study was that the Korean children produced the English sentences with a detectable foreign accent (FA)

<u>Analysis of subgroups</u>

- 10-year-old Koreans who had arrived in North America four years earlier, that is, at the chronological age of 6 years, received significantly lower ratings than did 10-year-old native English children

Conclusions re: CPH

- The presence of FA in children is not something that is likely to be observed casually. The FA measure used here was fairly fine-grained.
- This and similar findings convinced us that many, perhaps even most children who learn an L2 will speak it with a detectable foreign accent, even following years of immersion
- This finding is not something one would expect if the presence of a FA was the result of having passed a "critical period"
- The presence of FA in many adult L2 learners of L2 is of course consistent with CPH. However, work by Bongaerts et al. has demonstrated that some adult L2 learners manage to speak without FA

Contrastive Analysis (CA) hypothesis

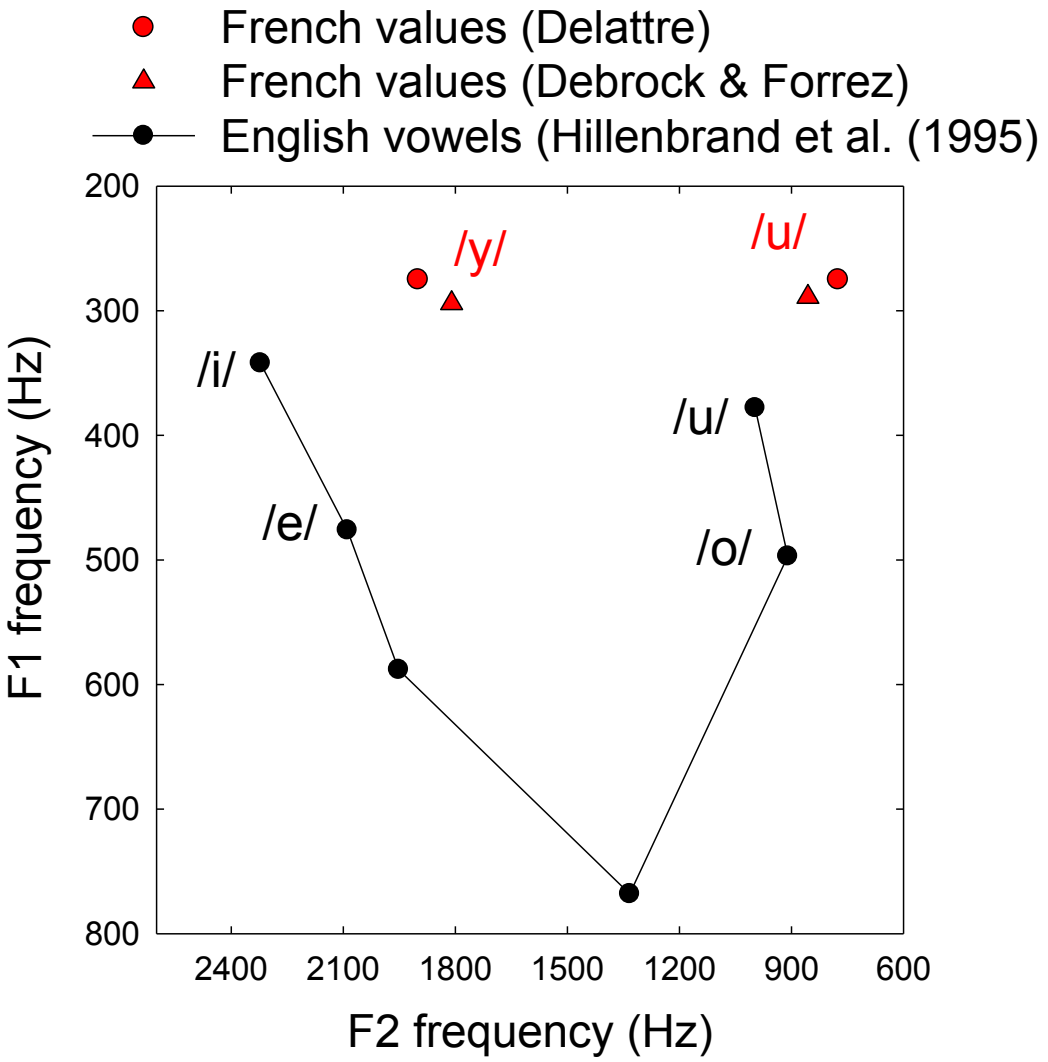
- In 1979-1984, most L2 research was framed in terms of CAH, which claimed that
  - L2 phonemes that are similar to L1 phonemes will be “easy” to produce;
  - L2 phonemes that are different from L1 phonemes will be “hard” to produce.
- Interference was seen as the major cause of most learning problems: what you already know in L1 will sometimes help you, but it may hurt you as well, depending on the structure(s) involved

Doubts re: CAH

- Flege (1987) examined English speakers' production of French /y/ and /u/
- Both vowels heard as English /u/ (Rochet) even though
  - English /u/ is more fronted than French /u/ (higher F2)
  - French /y/ is "radically different" from any English vowel, and so "new" (Delattre, 1964, p. 83)

French-English vowels

Comparison of acoustic values for two French vowels (red symbols) and six English vowels. The mean F1 and F2 values (in Hz) were drawn from the studies of Delattre and Debrock & Forrez. The values for English vowels were drawn from a study by Hillebrand et al. (1995)



Flege (1987)

Tested 3 groups of native English speakers, all women

Group B: Consisted of American college students who had just returned to Chicago after a 9-month academic program in Paris, France.

Group C: Consisted of somewhat older native speakers of American English who all had obtained advanced degrees in French, taught French at American university, and had lived for at least a brief period in France

Group D: Consisted of Americans who lived in Paris for *M* = 10 years

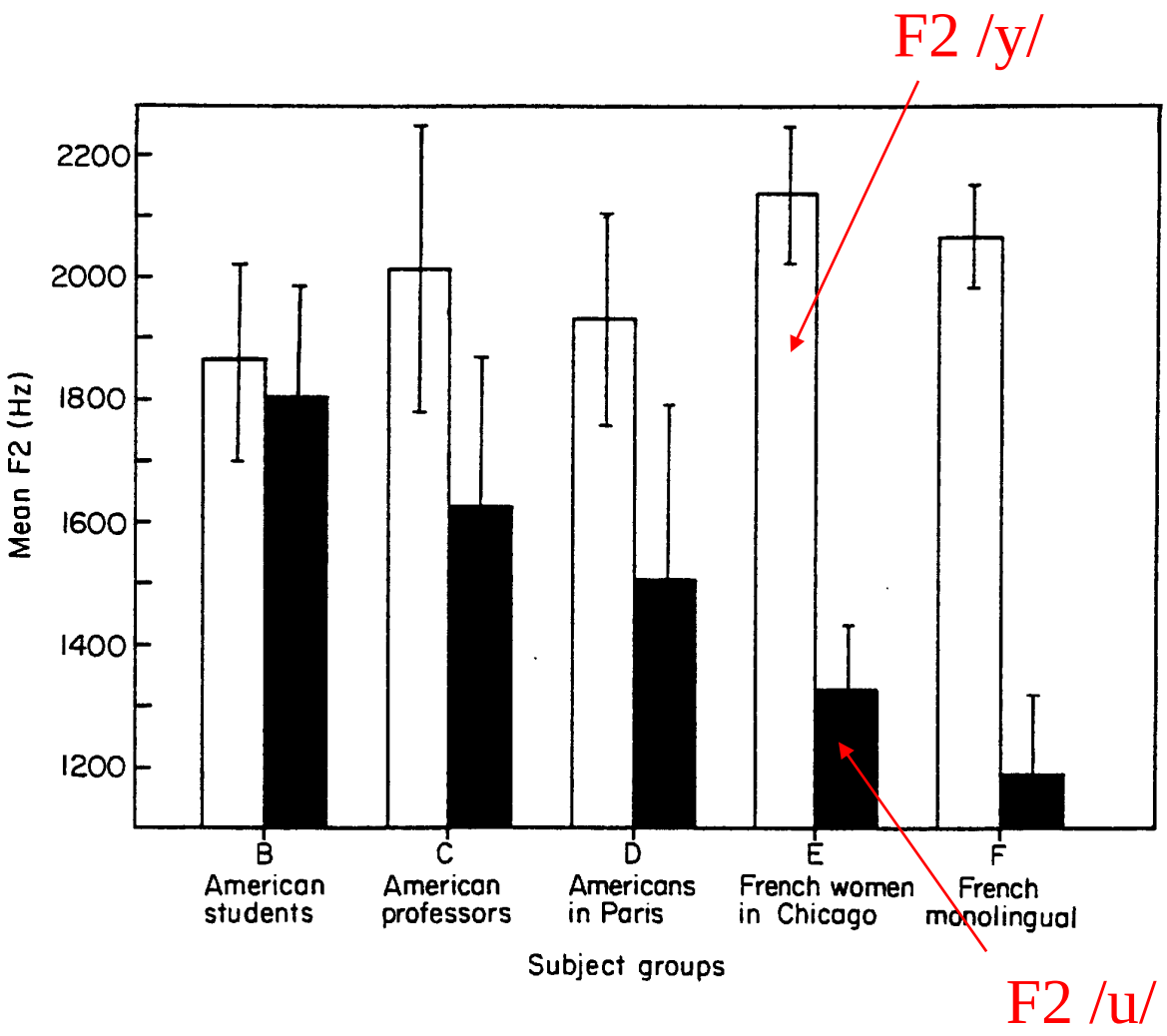
Flege (1987)

… and 2 groups of native French women

- Group B:
- Group C:
- Group D:
- Group E: All were living in Chicago when tested and had lived there for an average of 10 years
- Group F: French monolinguals living in Paris, France

Mean acoustic values for French vowels

Mean F2 values reported by Flege (1987) for the high French vowels /y/ (white bars) and /u/ (black bars) as produced by five groups of participants. The French monolinguals produced the largest F2 differences between /y/ (a front rounded vowel) and /u/ (a back rounded vowel)



Flege (1987)

- All three groups of native English-speaking women (i.e., groups B, C & D) produced French /u/ with significantly higher F2 values than did French monolinguals. This means that they were producing "fronter" variants of the French vowel, presumably as the result of the influence of the "similar" American English vowel /u/
- However, none of the three native English groups differed significantly from the French monolinguals when producing the "new" French vowel, /y/.
- Possible interpretation: they could produce French /y/ accurately because there was no influence from a vowel in their L1 phonetic inventory

Conclusions re: Contrastive Analysis Hypothesis (CAH)

- Comparisons of acoustic values suggested that French /y/ may be more dissimilar from the closest English vowel than is French /u/. This needs to be verified in a formal perceptual test.
- The Flege (1987) results suggest that adult learners of an L2 may be more successful at producing a “new” vowel in the L2 than an L2 vowel that is more similar to something already found in the L1
- This is the opposite of what one might expect from the Contrastive Analysis approach, which specifies that “similar” is easy whereas “new” is hard to learn

Abstract linguistic analyses

- In the period 1979-1984 it was widely believed by students of L2 acquisition that the phonologies of an L1 and L2 come into contact at an abstract phonemic level of analysis
- On this view, learners perceive the sounds of an L2 through the grid of their existing L1 “phonology” (see work by Trubetzkoy)
- As a result, learners perceive (hear) and produce L2 words as if they were concatenations of L1 phonemes (“new wine in old bottles”)

Abstract linguistic analyses

- Phonemes were freely commutable. This property ensured that large lexicons could be generated with a small set of phonemes.
- The phonemes themselves were regarded as bundles of commutable distinctive features (most of which were also highly abstract, even though some were associated with specific acoustic and articulatory dimensions
- It was generally accepted that it is not possible for learn to use “new” features that are not already deployed in the service of phonemic contrasts in the L1

Doubts re: abstract analyses

- Flege & Port (1981) evaluated these assumption by examining the production of English /p/ by several groups of native speakers of Saudi Arabian Arabic.
- All were young men who had come to the U.S. on scholarship to study at Indiana University

Flege & Port (1981)

- Arabic was chosen as the target L2 in this study because the Arabic phonemic inventory has /b/, /d/, /t/ and /k/, but no /p/ or /g/
- Logically, Arabic must have [voicing] and [place] features for stops
- The authors reasoned that if learning to produce an L2 occurs at a phonemic level, their Saudi participants should be able to produce English /p/ successfully by re-combining the Arabic [voicing] feature needed for /t/ and /k/ with the Arabic [place] feature needed for /b/

Flege & Port (1981)

- Acoustic phonetic measurements indicated that the native Arabic participants produced with English /p/ with temporal properties appropriate for a bilabial stop,but with the glottal pulsing that is characteristic of a phonologically voiced stop.
- Not surprisingly, the Saudis' English /p/ productions were often heard as /b/ by native English-speaking listeners
- Conclusion: The Saudi Ss did not re-combine abstract features
- Conclusion: the Saudis' difficulty is best described as learning to produce a new phonetic segment (speech sound) not a new phoneme

Doubts re: abstract analyses

- McAllister, Flege & Piske (2002) evaluated L2 learners' ability to learn to use a new distinctive feature
- The authors examined the use of a phonemic [length] feature that is needed to produce and perceive vowel distinctions in Swedish
- The nonnative participants were speakers of English and Spanish who had lived for more than 10 years in Stockholm, Sweden
- A phonemic [length] feature is not used to distinguish vowels in either English or Spanish. The nonnative Ss had to learn it in Swedish if they were to produce and perceive Swedish vowels adequately

McAllister, Flege & Piske (2002)

McAllister et al. examined four Swedish long-short vowel pairs that differed according to the [length] feature:

- Two pairs of mid-vowel contrasts. The vowels in these pairs were relatively similar in vowel quality (formant frequencies). The differences resided almost entirely in terms of duration, which is measured in msec a phonetic level of analyis
- One pair of high vowels, and one pair of low voewls. These contrasts were based on both duration and spectral quality

McAllister, Flege & Piske (2002)

Participants in the McAllister et al. study: Twenty native speakers each of:

- Swedish,
- English
- Spanish
- Estonian

The Estonians served as a control group. Their L1 has vowel contrasts based on [length]. Their difficulty in producing or producing Swedish vowels, if any, could be attributed to general difficuly learning an L2 and not to the need to learn a new abstract feature

McAllister, Flege & Piske (2002)

The stimuli used in a perceptual experiment were highly frequent words spoken by an adult male native speaker of Swedish. All stimuli contained either a phonological long or short vowel

- Half of the stimuli were real Swedish words, produced correctly
- The remaining half of the stimuli were non-words created by producing a word having a long vowel with the corresponding short vowel or vice-versa

McAllister, Flege & Piske (2002)

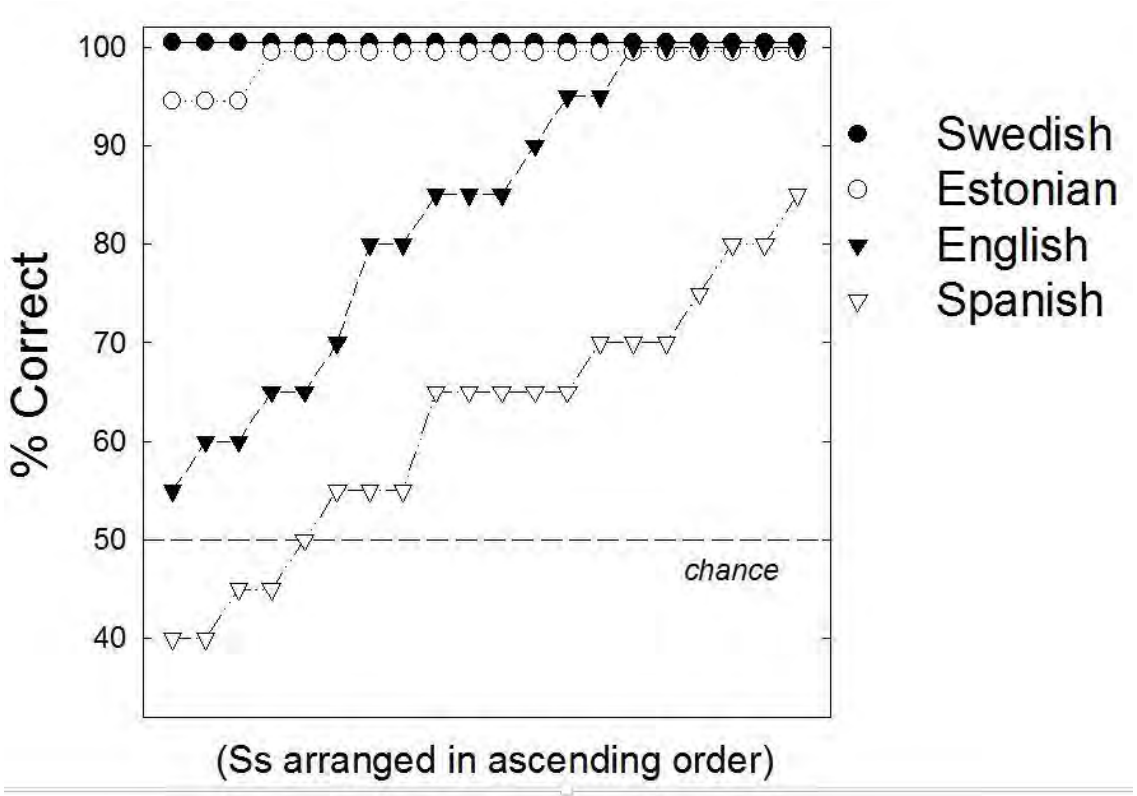
- The task of participants in the perception experiment was to indicate if each stimulus had been produced “correctly”
- As expected, the task was very easy for native speakers of Swedish
- An analogous task in English would be to ask if

  /fiʃ/ was a correctly produced English word.

  In this illustrative example, the English vowel /i/ substitutes the vowel /ɪ/, making a non-word in English, which should be labelled “incorrect”

A Swedish long-short contrast: /ø/ vs /ø:/

Mean % correct scores. The Swedish and Estonian Ss obtained high scores bcause they could determine whether familiar words contained a phonemically long or short vowel. The English Ss obtained lower scores. Some Spanish Ss performed at or below chance indicating that they didn't know if words in their Swedish lexicon had a long or a short vowel. These Ss had clearly not learned to use a new feature in their L2



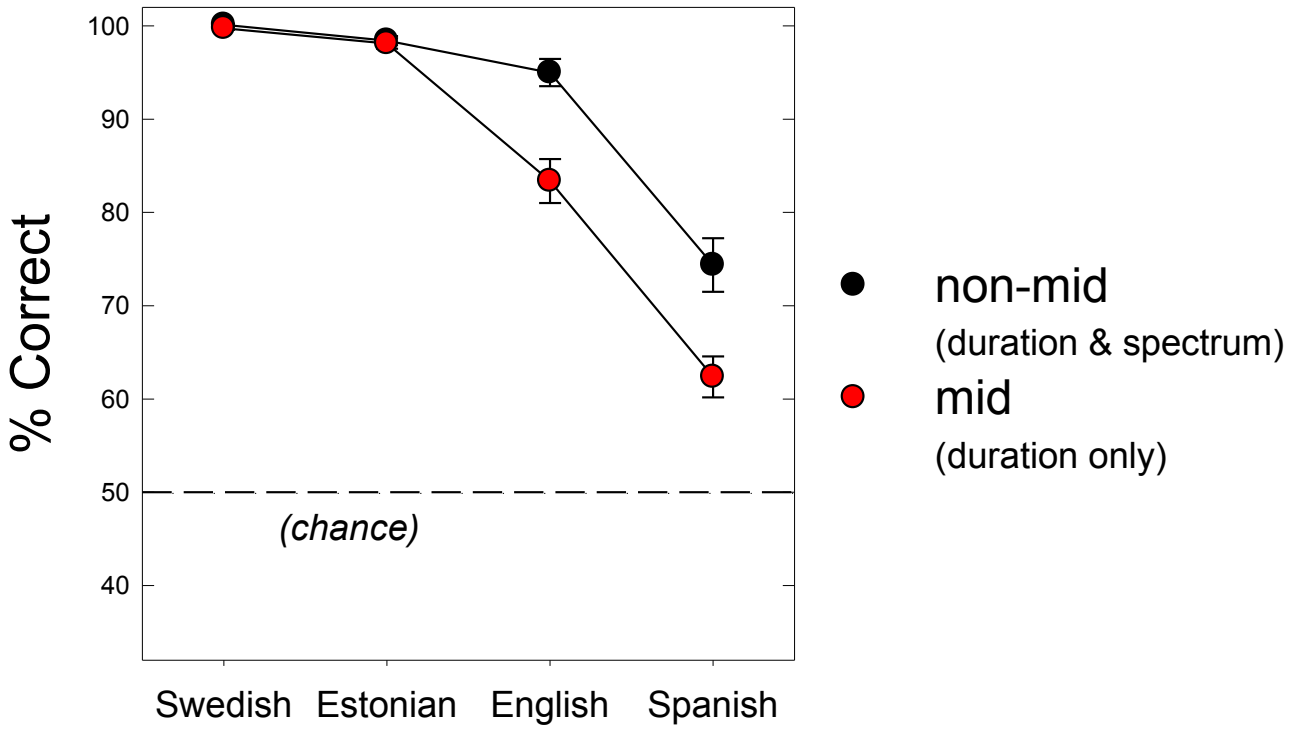
McAllister, Flege & Piske (2002)

The authors carried out separate analyses for

- 2 mid-vowel pairs: these vowel contrasts are based almost entirely on duration; the opposing members of the pairs differ little in terms of spectral quality
- 2 non mid-vowel pairs (1 high, 1 low): these contrasts are based on based on both duration & spectral quality differences.
- We expected better performance on the high- and low-vowel pairs. If an L2 learner were oblivious to the duration differences they could always rely on the spectral differences

McAllister, Flege & Piske (2002)

The native English and Spanish Ss obtained higher scores for non-mid than for mid-voweel pairs because, for the former, they could make use of both temporal (duration) and quality (spectral cues whereas for the mid-vowel pairs they were forced to rely on temporal cues. No such differeence was observed for the Ss whose L1 makes use of a [length] feature



McAllister, Flege & Piske (2002)

Some conclusions

1. Estonians benefited from the presence of a [length] feature in their L1. Use of a phonological feature (sensitivity to an acoustic phonetic dimension?) is likely to be re-deployed in an L2
2. Native English & Spanish Ss continued to rely on spectral cues (features?) when they acquired new Swedish words. Some showed evidence of little or no sensitivity to duration contrasts between Swedish vowels.
3. Results indicate that it is (a) difficult to acquire a new, abstract feature, or (b) difficult to acquire sensitivity to an acoustic phonetic dimension – duration – that is not used or else used differently in the L1

McAllister, Flege & Piske (2002)

However

- It is important to note the very large differences between individual Ss, both between languages (Spanish, English) and within languages
- Most native English Ss and some native Spanish Ss showed significantly above-chance performance, which indicates at least some sensitivity the presumably didn't exist before the Ss were exposed to Swedish. (More research needed!)
- The excellent performance of some native English Ss could be cited as evidence that new features can be acquired in an L2 … at least by some

Doubts re: abstract analyses

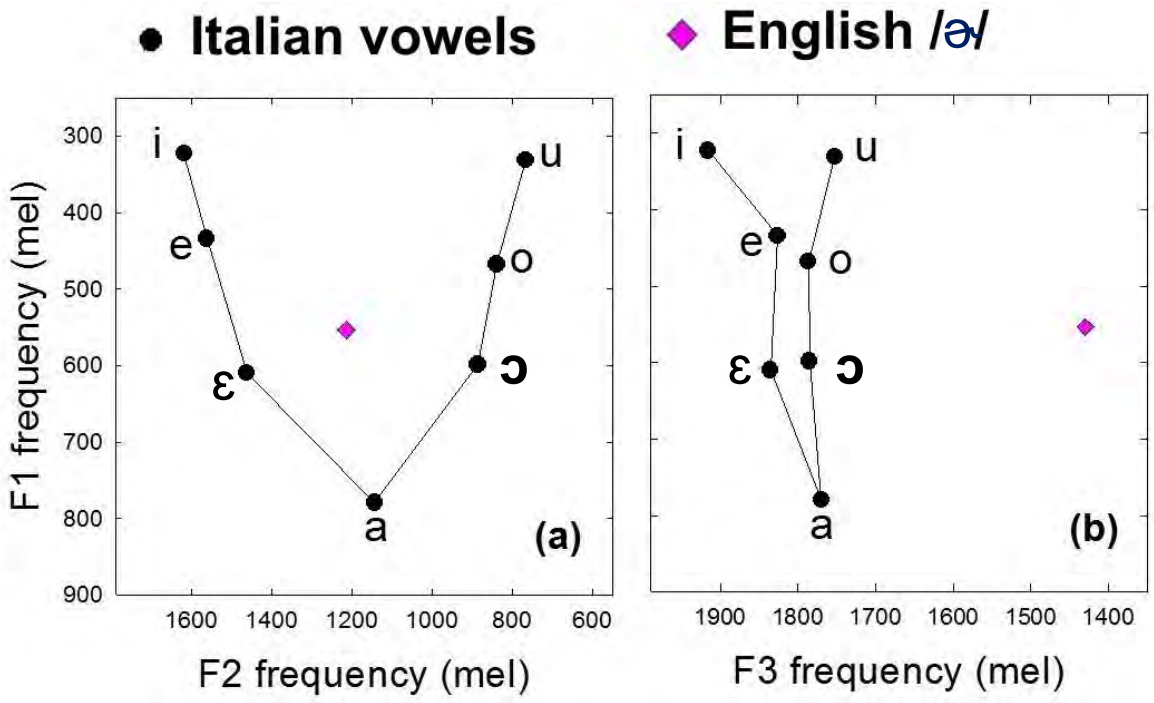
Now let's consider the results of three studies that examined production of the rhotic English vowel /ɚ/ (as in "bird", "heard") by native speakers of Italian:

- Munro, Flege & MacKay (1996)
- Flege, Schirru & MacKay (2003)
- Flege & MacKay (2004)

Doubts re: abstract analyses

- English /ɚ/ very dissimilar acoustically and perceptually from any vowel found in Italian
- When Italians mimic an American accent in Italy they rhotacize vowels in an exaggerated manner, indicating they are aware of this acoustic phonetic property and can, under the right circumstances, produce it
- The perceptual dissimilarity of English /ɚ/ from any Italian vowel was demonstrated in a perceptual assimilation experiment carried out by Flege & MacKay (2004)
- This corroborated acoustic analyses comparing /ɚ/ to Italian vowels

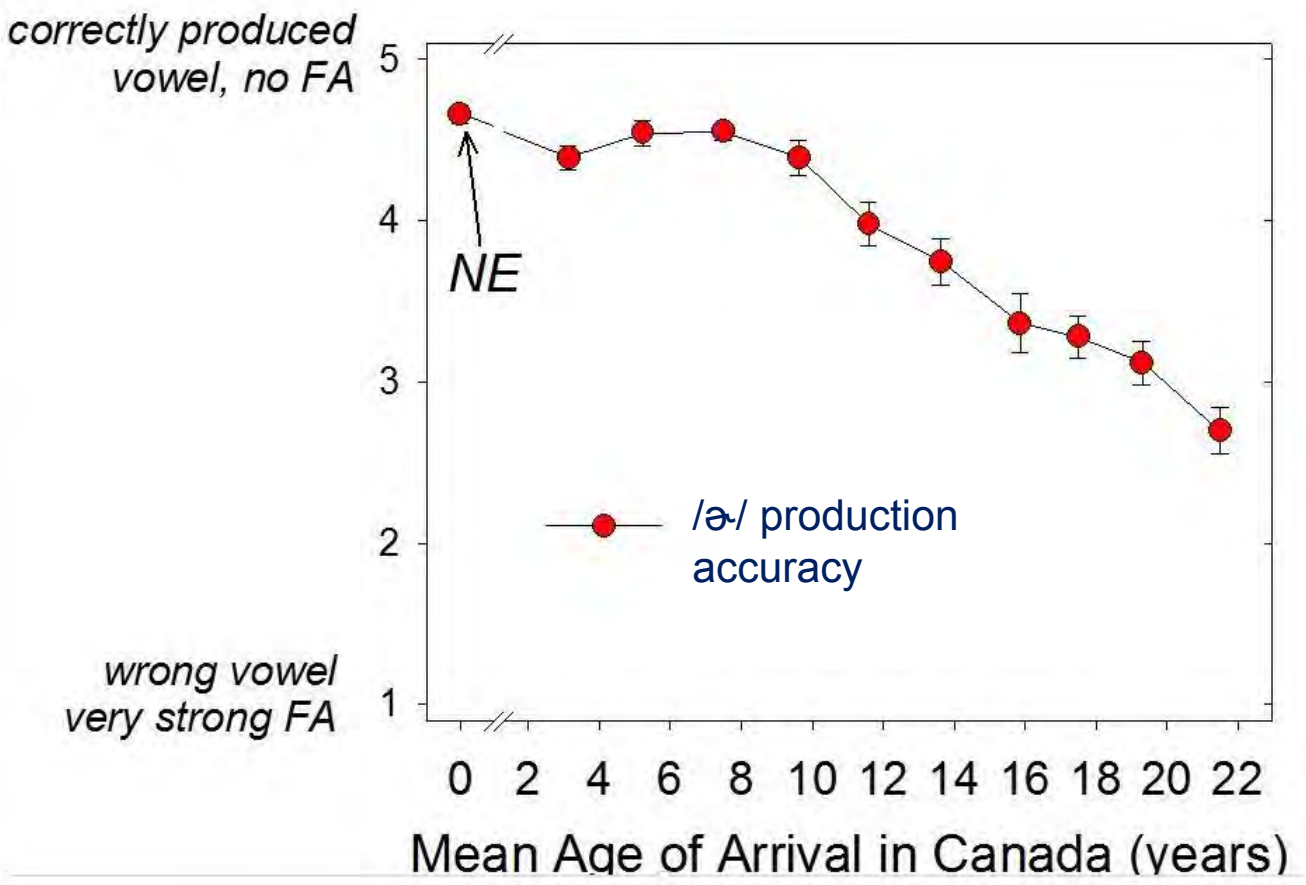
Acoustically, the /ɚ/ of English differs substantially from any vowel in the Italian inventory, both in terms of (a) F1 and F2 formant frequencies and (b) the frequency of the third formant, F3



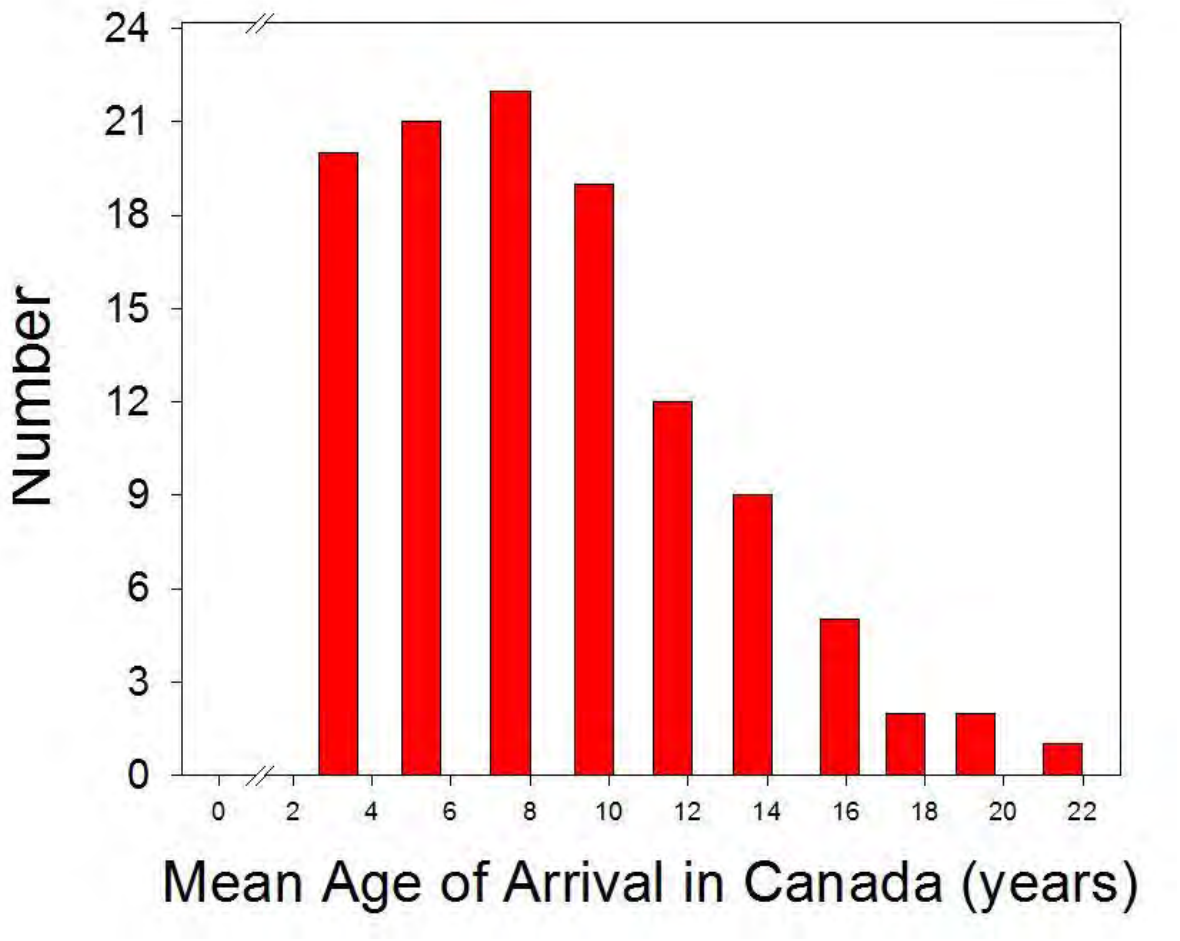
Munro, Flege & MacKay (1996)

- Examined a total of 240 native speakers of Italian who had immigrated to Canada. The participants differed, among according to
  - Their original age of arrival (AOA) in Canada
  - Length of residence (LOR), in years Canada
- Production samples were obtained having the Ss repeated English words after a filled delay (delayed repetition task)
- Vowels producing accuracy was assed by having native English listeners rate the vowels (5-pt scale)

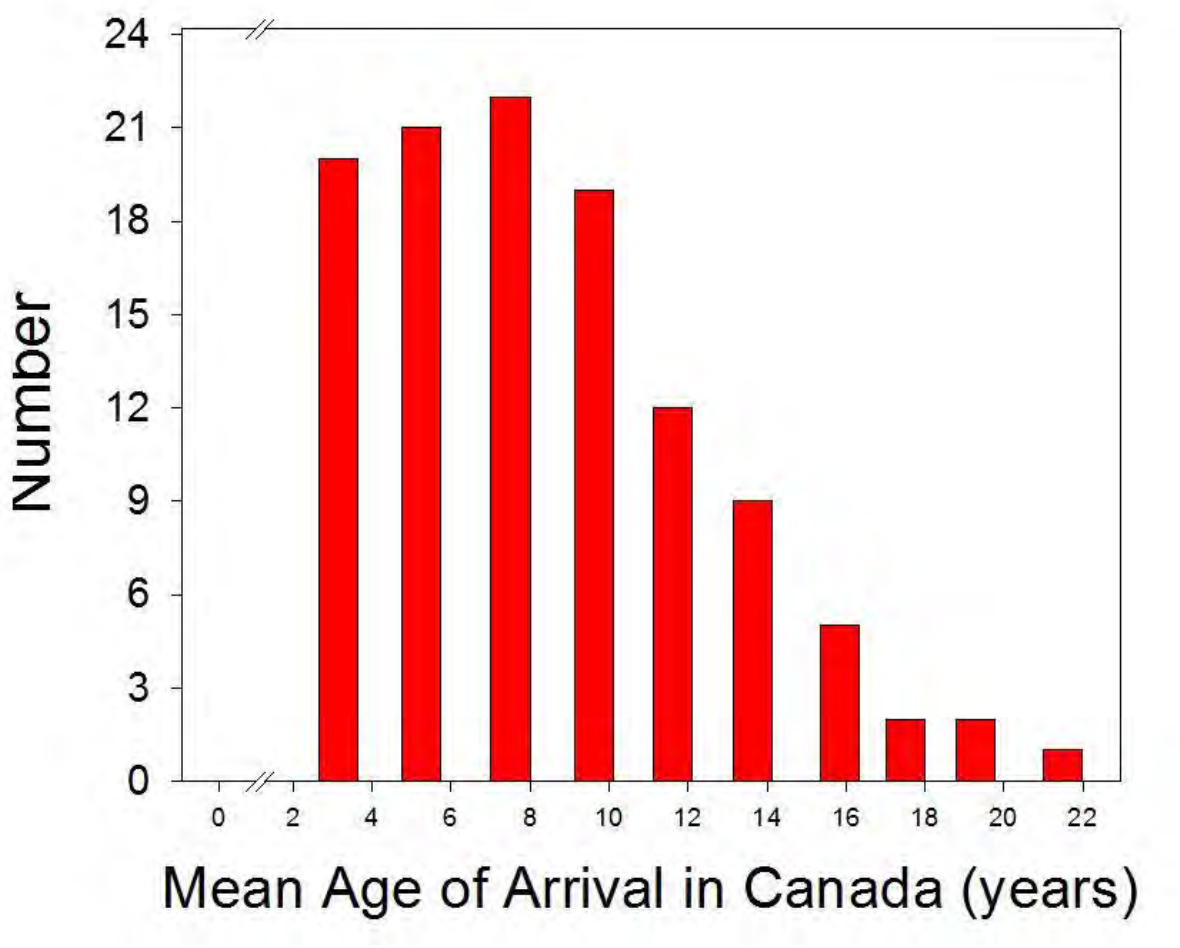
Munro et al. (1996). Mean ratings of /ə-/ production accuracy for groups of native Italian Ss differing in age of arrival (AOA) in Canada



Munro et al. (1996). Number of native Italian Ss per group (n = 24) whose /ə˞/ productions received a rating within 2 SDs of the mean rating obtained for production of this vowel by native English speakers



Munro et al. (1996). Most Italian Ss who arrived in Canada before the age of 12 years produced /ə-/ well, but few of those who arrived later in life did so. Why? A limitation on production? Or on perception?

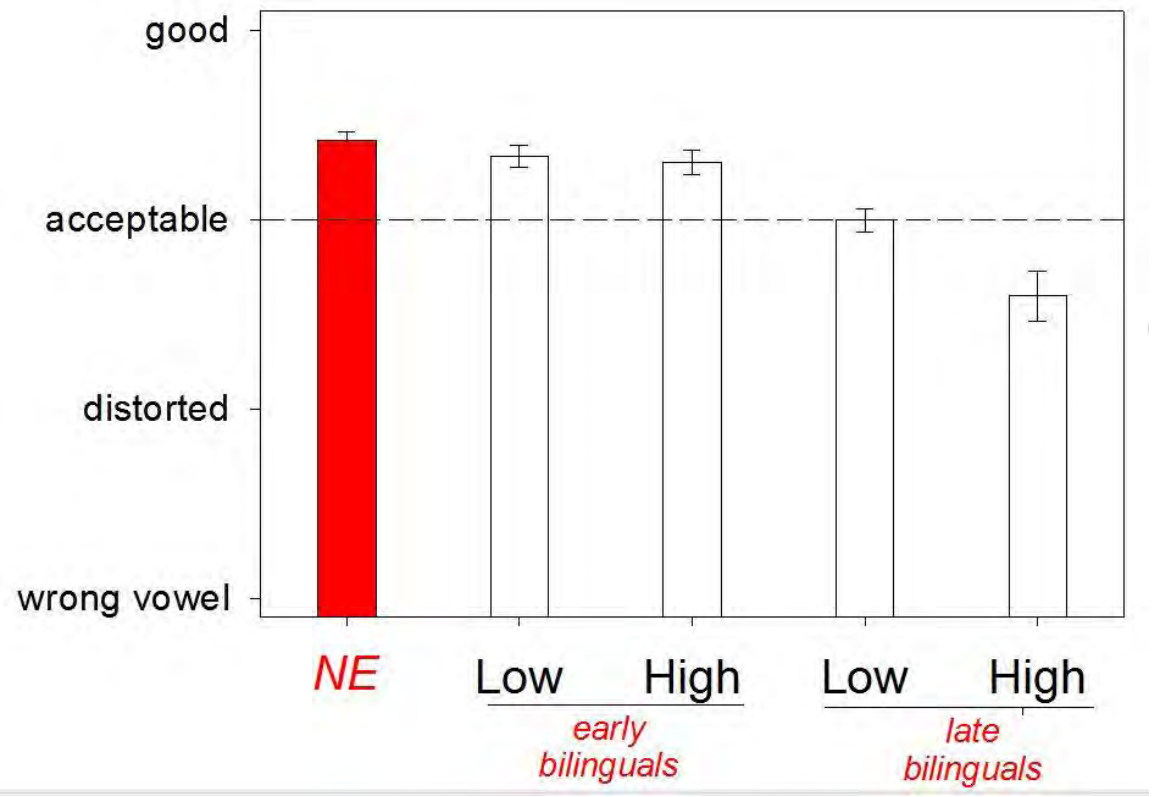


Flege, Schirru & MacKay (2003)

This study examined production of /ə˞/ by 5 groups of 18 participants each:

- Native speakers of English (NE)
- 2 groups of early learners who arrived in Canada from Italy as children but differed according to average self-reported use of their L1, Italian use (means = 7% vs. 43%)
- 2 groups of late learners who arrived in Canada later in life, subdivided according to self-reported use of Italian (means = 10% vs. 53%)

Flege et al. (1993) *NE-listeners used 4 labels to classify English vowels, presented in separate blocks. Tokens were considered "accurate" if classified "good" or "acceptable". Non-parametric tests evaluated the number of Ss in each group (max = 18) whose vowels were produced accurately. Results below are for /ə˞/*



Flege et al.(2003) results

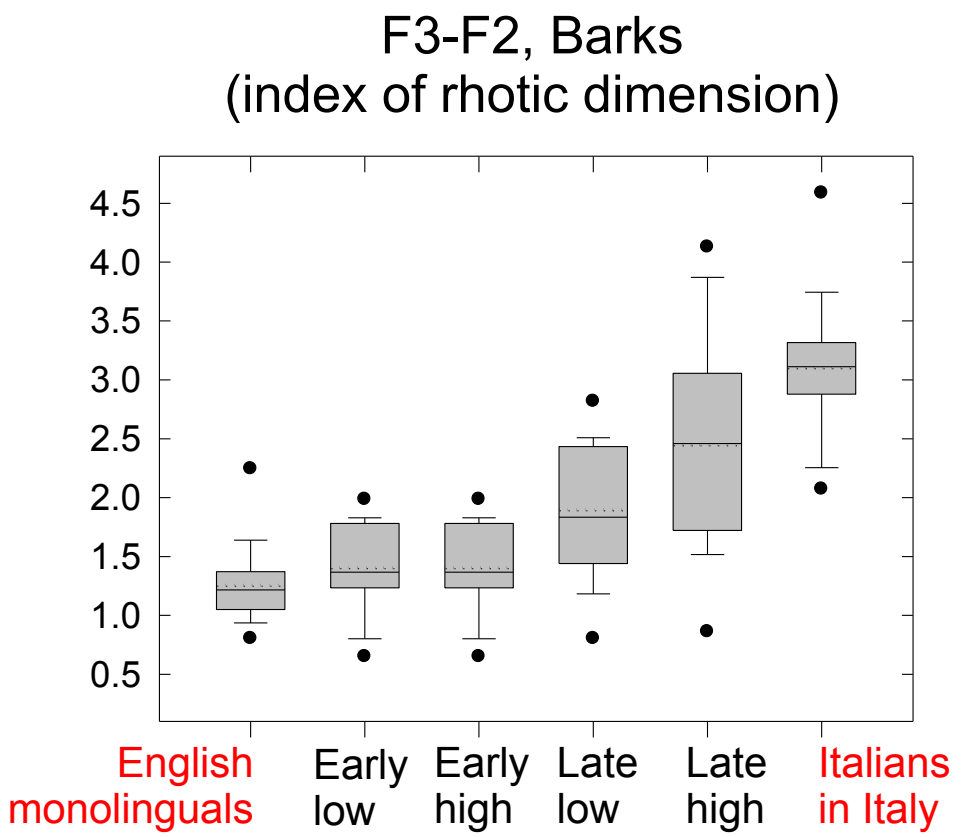
- The non-parametric (also parametric) analyses revealed that native speakers of English produced /ɚ/ more accurately than both groups of late native-Italian learners of English (late-low use of Italian, late-high use of Italian) but did not differ significantly from either group of early learners
- Perhaps a difficulty/inability to learn a new “feature” ([rhotic]) or acoustic phonetic dimension applies just to late learners

Flege et al.(2003) results

- However, detailed acoustic analyses suggested that the Italian late learners of English do acquire sensitivity to the [rhotic] feature
- The analysis examined Bark-transformed F3-F2 differences, which provides a perceptually relevant index of the [rhotic]

Flege et al.(2003) results

The early learners were found to have produced significantly smaller F3-F2 differences in the */ə-/ tokens than did the late native Italian learners of English. In other words, the early learners were more successful than the late learners in producing the "rhotic" feature (property)*



Flege, Schirru & MacKay (2003)

However …

- Late bilinguals produced significantly smaller F3-F2 differences
- Some late bilinguals seem to have acquired sensitivity to [rhotic] feature; in fact, some produced native-like F3-F2 values
- Regression analyes indicated, for all 11 vowels examined, that the more Italian continued to be used, the less accurate was the production of English vowels

Conclusions re: abstract analyses

- Both early and late learners are able to gain access to features (acoustic phonetic dimensions not used to contrast L1 phonemes
- Late learners are less likely to do so than early learners, although some late learners produced /ə̃/ successfully.
- Amount of continued L1 use as a predictor of success in producing L2 vowels. It will be important in future research to understand better the souces of variation in L2 learning, especially among late learners

Categorical perception (CP)

- The CP paradigm applied to most cross-language perception studies carried out in the period 1979-1984
- Indicates significantly better discrimination of stimuli assigned to the same category than two stimuli belonging to the same category
- For example; English voiceless stop tokens having VOT values of 25 and 65 msec will not be discriminated if both are identified as the phoneme /t/
- The audible acoustic differences between two such tokens were said to be “filtered out”

Categorical perception (CP)

- The sound system of two languages differ in terms of number & kinds of categories, and also in the phonetic specification of those categories
- If audible L1-L2 phonetic differences are filtered out before they reach a conscious level of perception, and so cannot be used in constructing representations for words being learned in the L2, then L2 learners are condemned to forever producing and perceiving thte sounds of an L2 inaccurately
- I call this the “doom scenario”

Doubts re: applying tenets of CP to L2 speech learning

- In 1979-1984 people who were researching L2 acquisition weren't generally cognizant of the speech perception literature.
- Abundant evidence existed even then that when monolinguals hear stimuli drawn from their L1 within-category phonetic information is available to them "although the retrieval of this information … will depend on the level of processing" (Pisoni & Tash, 1974)
- There is no reason to suppose that it will be otherwise when the same monolinguals are exposed for the first time to an L2

Doubts re: applying the CP paradigm

Let's consider the results of two studies which suggest that applying the "phonological grid" of the L1 to sounds encountered on the phonetic surface of an L2 via categorical perception does not cause L2 learners to filter out audible (at a sensory level) cross-language phonetic differences

- *Flege (1984)*
- *Flege & Hammond (1982)*

Flege (1984)

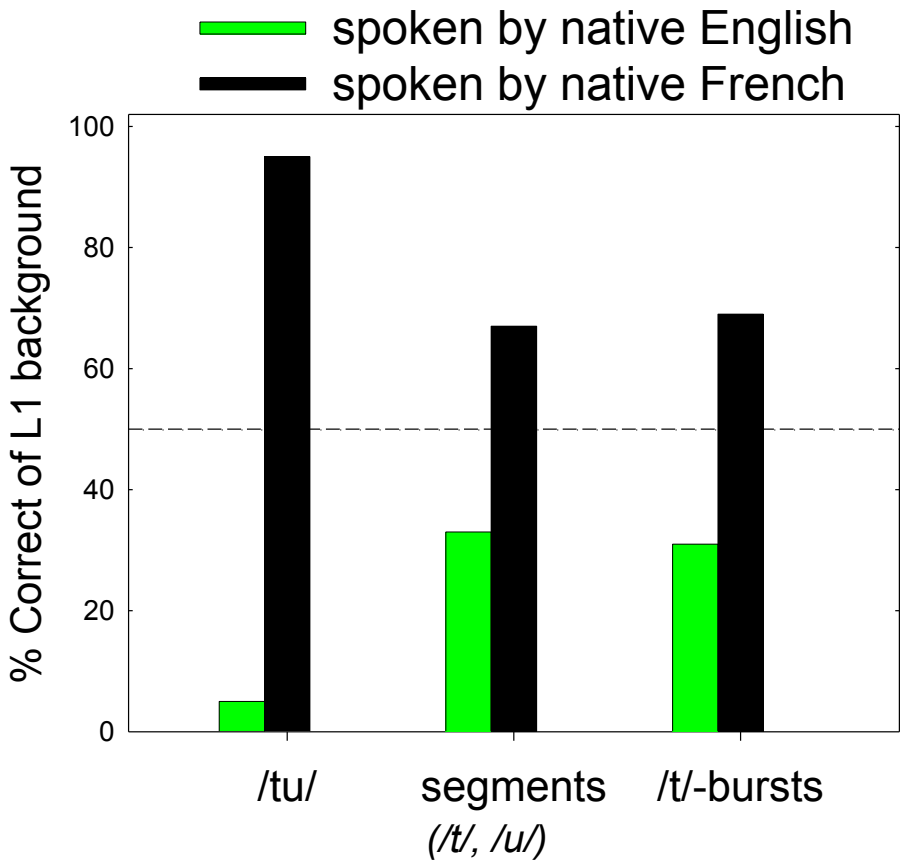
- Examined speech samples produced in English by adult native speakers of English and French. The stimuli consisted of:
  - Unmodified tokens of the syllable /tu/ edited from phrases such as *two little girls*)
  - Hybrid /tu/ tokens created by splicing instances of /t/ or /u/ edited out of original syllables and then cross-spliced. One segment of the hybrid syllables was a a single segment produced a NE speaker, the other segment varied
  - The final set of stimuli consisted of the first 30-ms of /t/, essentially a release burst plus a bit of aspiration

Flege (1984)

- The English stimuli presented in pairs, of which one member was produced by a native speaker of English and the other member consisted of at least a segment produced by a *native speaker of French*
- The listeners' task on each trial was to decide which of the two stimuli was "foreign"
- The listeners were given no training or feedback on the task.
- The task was unspeeded; however, a response was required on every trial before listeners could move on to the next trial.

Flege (1984)

The NE listeners had no difficulty deciding which member of a pair was was foreign-accented when both segments in the syllable was free to vary (the original /tu/ syllables. Accuracy decreased while remaining well above change when just one segment or part of a segment was produced by a non-native speaker



Flege (1984)

These findings demonstrated that NE listeners could detect specific expected aspects of French accent in English, including

- Production of /u/ as a “backer” (non-front) vowel than is typical for English, thereby resembling French /u/
- Production of /t/ with tVOT values that were too short (resembling values typical for French /t/)
- A tendency to realize /t/ with a dental rather than alveolar place of articulation

Flege (1984)

- Contrary to those who would apply the CP paradigm to L2 learning, the result of this experiment showed that English monolinguals' could readily detect "within" category variation arising from cross-language phonetic interference
- Next we ask: is this kind of performance possible outside the laboratory where listeners must pay attention to meaning?

Flege & Hammond (1982)

- These authors tested 50 native English-speaking students who were enrolled in first-year Spanish classes at the University of Florida (Gainesville)
- All participants were familiar with Spanish-accented English having been born in raised in Florida and being enrolled in a Spanish class that was being taught – mostly in English! – by native Spanish instructors who spoke English with fairly strong Spanish accents

The Ss' inserted a set of English words in a carrier phrase (*The__is on the__*). They were asked to produce the sentences, recorded for later analysis, with a "Spanish accent". No coaching, explanation or training was provided

Lexical Items	Substitute	Frequency
nose, cheese, hose	s/z	141 (47%)
vice, veil, vase	b/v	129 (43%)
fig, pig, wig	i/I	127 (42%)
book, hook, crook	u/U	61 (20%)
shell, sheet, sheep	č/š	49 (16%)
bean, phone, bone	ŋ/n	1 (0.3%)
tape,tube, toad	d/t	0 (0.0%)

Flege & Hammond (1982)

- The authors later noted all segmental substitution in the target words. The Ss were subdivided into group according to how often they produced the kinds of substitutions typical for Spanish-accented English (“most” vs. “least”)
- The authors reasoned that those who produced many expected substitution had more experience with Spanish accented English than those who produced fewer such substitution

Flege & Hammond (1982)

- The authors also measured VOT in 6 /t/ tokens for each participant.
- This acoustic phonetic dimension was of interest because
  - Native speakers of Spanish often produce English /t/ with VOT values that are intermediate to the those of Spanish monolinguals (short-lag) and English monolinguals (long-lag)
  - When VOT is shortened sufficiently, an intended /t/ may be heard as /d/

Flege & Hammond (1982)

Findings

- The Ss imitating a Spanish accent in English were never heard (transcribed by 2 phoneticians) as having substituted an English /d/ for the /t/ of Spanish accented English (SAE)
- This could mean either that the Ss had (a) had failed to detect the very typical shortening of VOT in SAE, or (b) were unable themselves to shorten VOT in their imitations
- The VOT measurements revealed that interpretation “b” was correct

Flege & Hammond (1982)

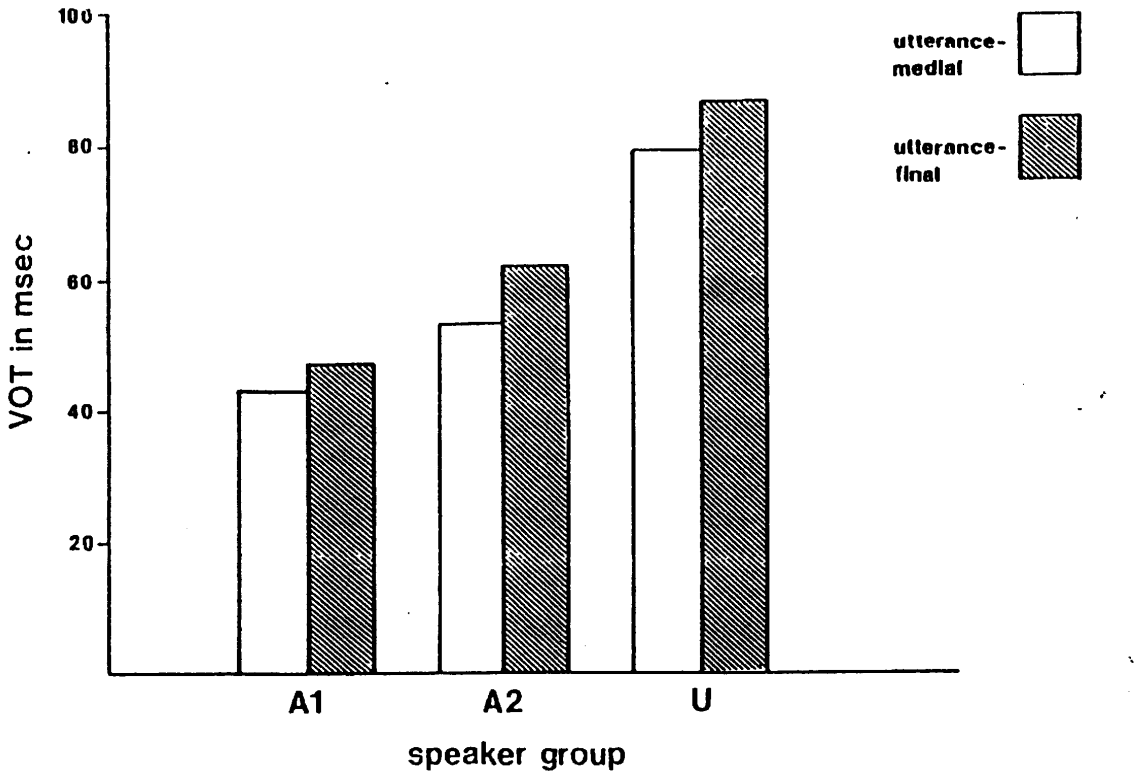
Mean VOT (msec) of key words inserted in the sentences frame “The X is on the Y”

Groups (n = 10 each)

A1 – produced the largest number of expected segmental substitutions

A2 – produced by fewest segmental substitution

U – control group consisting of Ss who simply read the speech materials without trying to produce a Spanish foreign accent



Flege & Hammond (1982)

Histogram representing the distribution of VOT values (in msec) measured in keywords produced by two groups of native English students asked to imitate a Spanish foreign accent in English as well as a control group (indicated by the arrow) of members of a control group who produced the speech materials (of the form: "The X is on the Y") without trying to imitate a Spanish accent

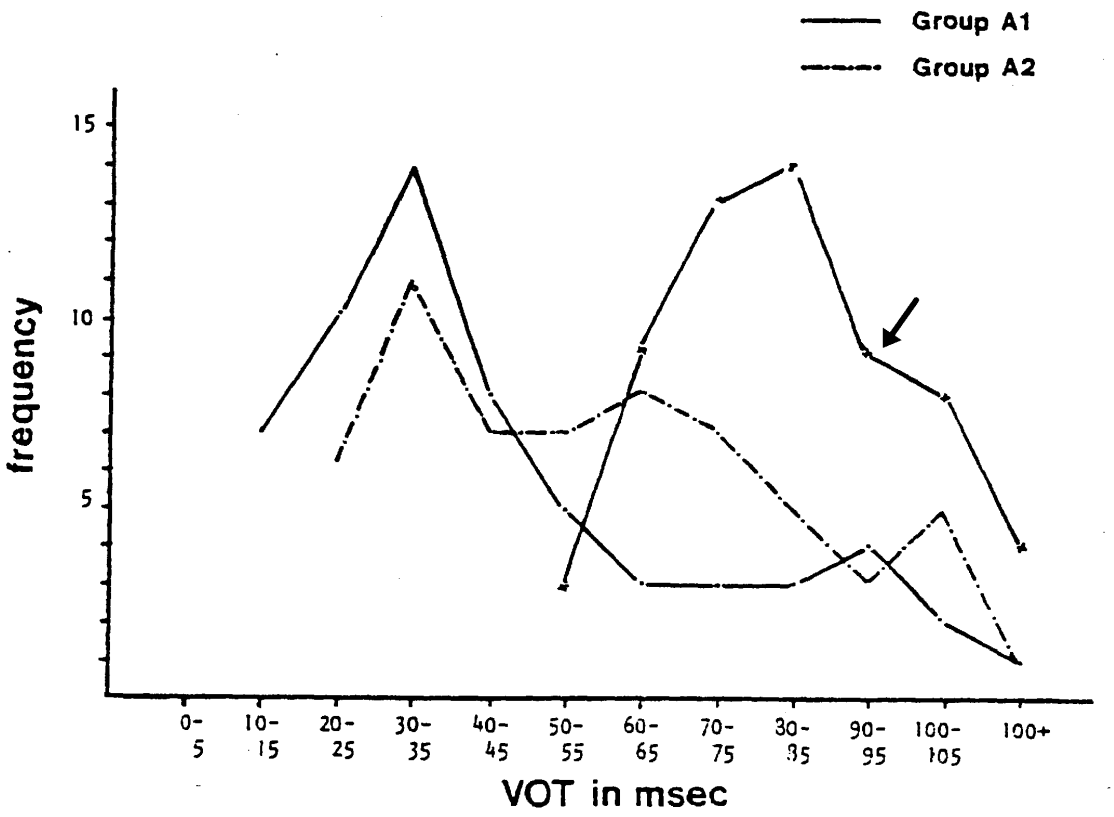


Figure 1. Frequency of VOT values measured in stops produced by three groups of speakers.

Flege & Hammond (1982)

Conclusions

- This last analysis indicated that the mean shortening observed in the Ss' imitations of SAE was not of artifact of averaging values for Spanish short-lag stops with English long-lag stops
- Therefore: as knowledge of SAE developed through direct exposure to English spoken with a Spanis accent, the native English Ss stored information in long term memory. These representations necessarily contained information that represented within-category phonetic variation

Conclusions re: applying the CP paradigm

- Applying the categorical perception (CP) paradigm or the “phonological grid filter” paradigm to second language speech acquisition doesn’t seem to be well founded
- Adults who are exposed to a foreign or second langauge can (eventually) detect within category cross-language phonetic differences, and store this information in long-term memory representations

Unidirectional L1→L2 interference

- When I was writing my PhD dissertation at Indiana University in 1979, students from W. Africa told me that when they returned home after years of absence their family & friends made fun of them for “affecting” an American accent in their L1
- I inferred that learning an L2 (in this case English) might influence production of the L1
- In the period1979-1984 “interference” of the L1 on the L2 was well documented but there was no interest or discussion of L2→L1 interference

Unidirectional L1→L2 interference

This lack of attention to something later recognized as "obvious" was probably the result of a number of assumption that were widely held prior to 1979

- Bilinguals "switch" between separate self-contained L1 & L2 phonological systems. When the L1 is "on" the L2 must necessarily be "off" and so cannot possibly influence the L1
- Errors in an L2 occur because it has not been properly or full learned, something not possible for an L1
- What is learned early on in the L1 "stays learned" (see work by Roman Jakobson)

Doubts re: unidirectional interference

Now let's review the results obtained in research that provides evidence of "L2 effects on L1", that is to say, reverse interference

- Yeni-Komshian et al. (2000)
- Yeni-Komshian & Flege (unpubl.)

Doubts re: unidirectional interference

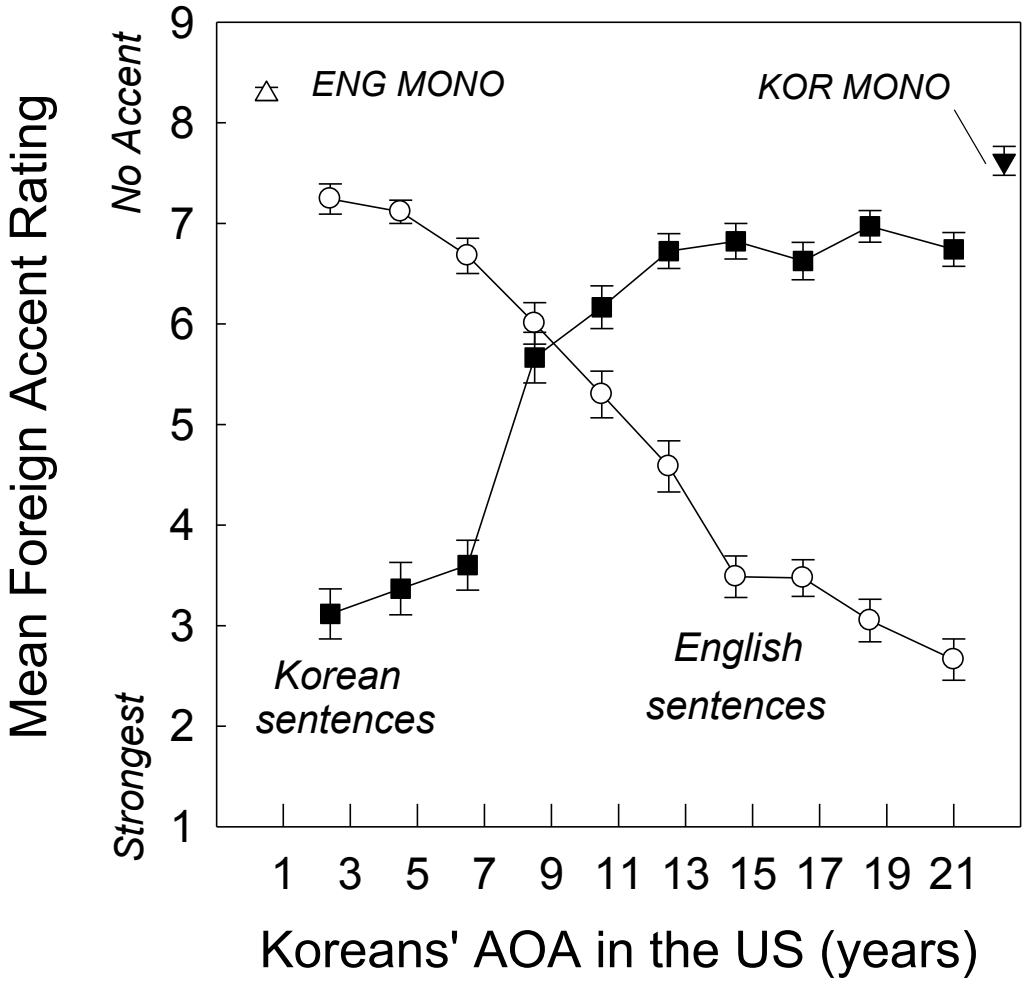
Yeni-Komshian et al. (2000) tested 240 Korean adults living in US

- English sentences were produced by native Korean and native English speakers.
- The native Korean Ss, all long-term residents of the U.S., were selected on the basis of their age of arrival in the US
- The Korean Ss and a Korean monolingual control group, also produced Korean sentences
- The English and Korean sentences were rated for overall degree of foreign accent by English & Korean monolinguals, respectively

Mean foreign accenet ratings obtained for Korean and English sentences by Yeni-Komshian et al. (2000)

Isolated symbols indicate the mean ratings obtained for sentences produced by English monolinguals (unfilled triangle) and Korean monolinguals (filled triangle)

The Korean adults who arrived in the US between the ages of about 2 and 8 years produced Korean sentences with what seems to have been an American foreign accent



Doubts re: unidirectional interference

- Yeni-Komshian & Flege (unpubl.) elicited isolated Korean words beginning with the consonants /s/, /s'/, /t^h/, /t'/
- After being digitally prepared, the stimuli produced by 240 Koreans living in the U.S. were randomly presented in separate blocks to native Korean-speaking listeners.
- Production of the word-initial consonants was judged to be:

  4 very good

  3 okay

  2 distorted

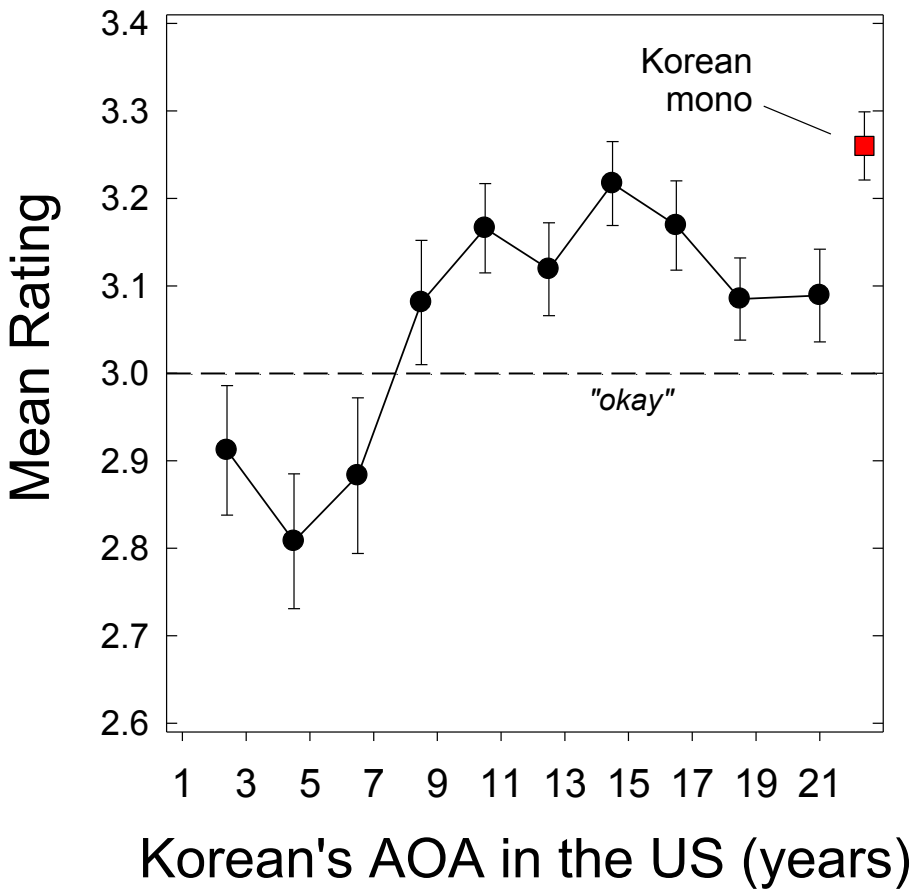
  1 wrong consonant

Results obtained by Yeni-Komshian & Flege (unpublished).

The brackets enclosed +/- 1 S.E.

Korean consonants produced by nearly all of the 240 Korean adults living in the U.S. received lower ratings than did consonants produced by Korean monolinguals in Korea.

Only consonants produced by Korean adults who had arrived in the U.S. prior the age of 8 years were judged, on average, to be less than adequate



Doubts re: unidirectional interference

- Flege & MacKay (unpublished) the production of Italian words spoken by 80 native speakers of Italian who were long-time residents of Canada and by monolingual native speakers of Italian recorded in Padova, Italy
- The voiced stops of English (/b d g/) can be realized with pre-voicing but are usually realized as stops having short-lag VOT values.
- In Italian, on the other hand, /b d g/ are realized with lead VOT values, that is, the closure before release is produced with glottal pulsing (“pre-voicing”)

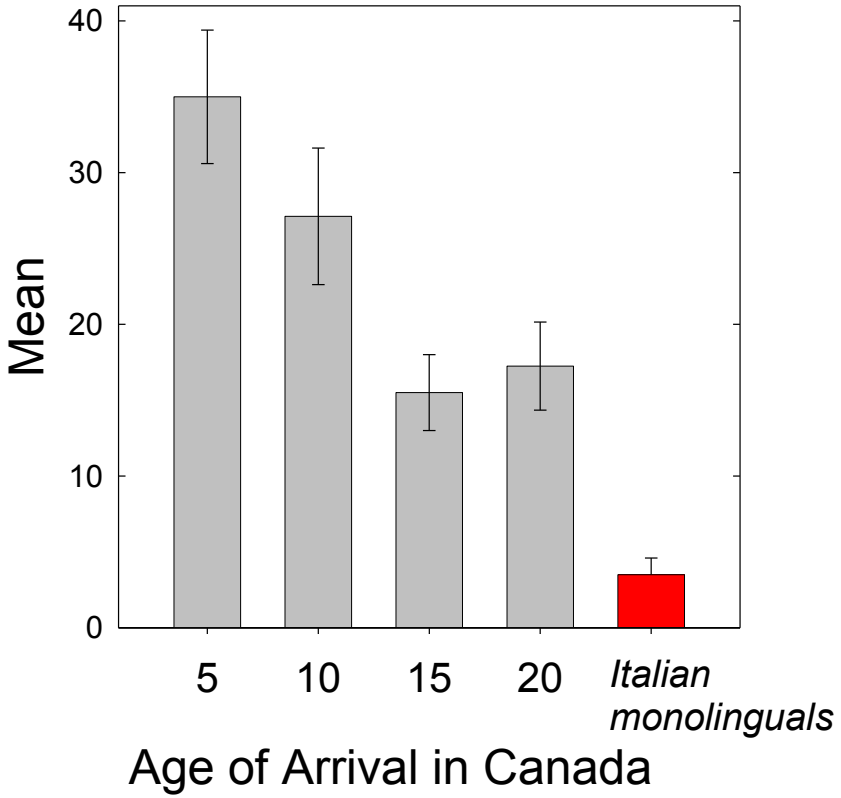
Results obtained by Flege & McKay (unpublished).

The brackets enclosed +/- 1 S.E.

Nearly all of the 80 Italians living in Canada realized Italian /b d g/ tokens as short-lag stops at least some of the time, whereas this was observed seldom in the speech of Italians living in Italy

The earlier in life the Italian-English bilinguals had immigrated to Canada the more often they produced Italian /b d g/ in an English-like fashion, that is, as short-lag stops

Mean % of Italian /b d g/ tokens (n = 20) realized as short-lag stops



Conclusions so far

The picture of L2 speech learning that began to emerge in our research differed substantially from what was generally assumed and/or believed in the period 1979-1984

- The L1 and L2 systems do not exist in splendid isolation
- Learning an L2 does affect on the L1 is produced
- The affect of L2 on L1 seems to be strong for early than late learners of an L2

Outline

1. Purpose of the Speech Learning Model (SLM)
2. Historical background
3. **Core aspects of the SLM**
4. Some predictions generated by the SLM
5. How to falsify the SLM
6. Future directions

The Speech Learning Model (SLM)

The SLM was developed to make sense of the empirical results we had begun to obtain. It rests on several basic premises:

- L2 learners can, in time, veridically perceive the phonetic properties of L2 speech sounds
- As in L1 development, L2 speech learning (a) takes time, and (b) is influenced importantly by the nature of input received
- As in L1 development, production is guided by perceptual representations stored in long-term memory

SLM

Further, the SLM proposes that

- The processes and mechanisms that guide successful L1 speech acquisition—including the ability to form new phonetic categories—remain intact and accessible across the life span
- The phonetic elements that make up the L1 and L2 phonetic subsystems exist in a "common phonological space", and so mutually influence one another

SLM hypotheses

- The greater the perceived dissimilarity of an L2 sound from the closest sound of the L1, the more likely a new category will be formed for the L2 sound
- Category formation for an L2 sound becomes less likely through childhood as representations for neighboring L1 sounds develop
- When a category is not formed for an L2 sound because it is too similar to an L1 counterpart, the L1 and L2 categories will assimilate, leading to a "merged" L1-L2

(illustrations/examples to follow)

SLM hypotheses

- By way of illustrating the hypotheses just stated, let's imagine the vowel spaces of an L1 and L2
- Let's also make some simplifying assumptions to facilitate the discussion

Illustration: perceived dissimilarity

Illustration of “perceived phonetic dissimilarity”

Let’s imagine that there are 5 vowels in the L1, depicted here by ellipses in a 2 dimensional high-low vs. front-back vowel space

Our imaginary language is similar to real languages such as Spanish.

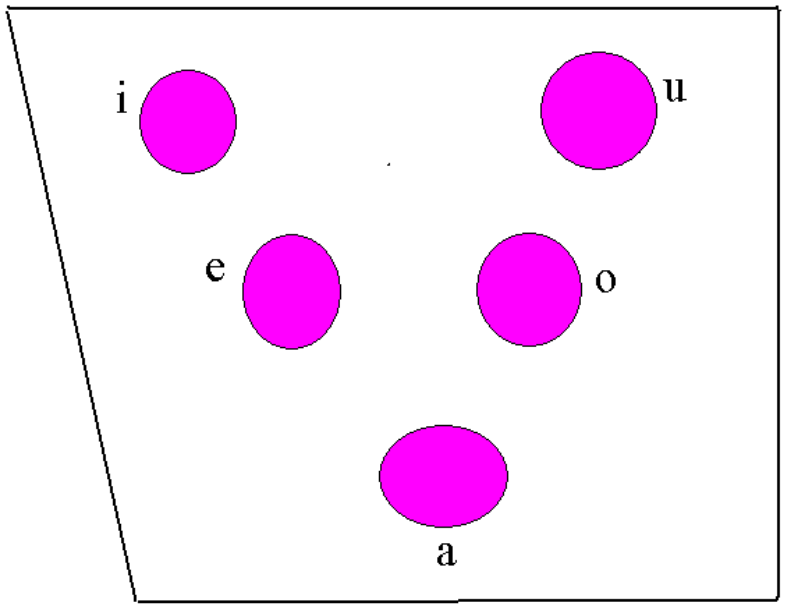
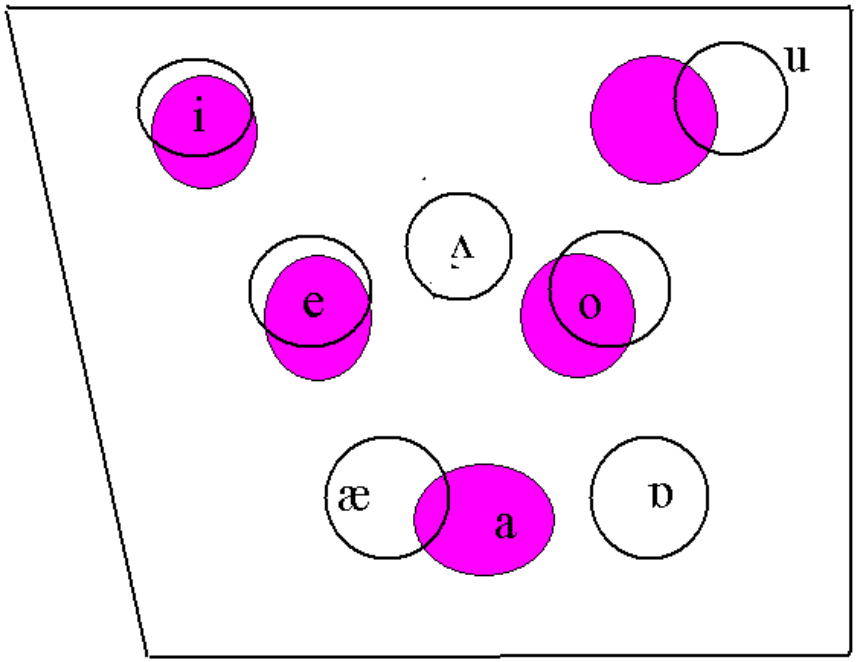


Illustration of “perceived phonetic dissimilarity”

Let’s suppose that the L2 has 7 vowels and that perception of vowels of the L2, like those of the L1, are based entirely on center formant frequency values (no role of either duration or formant movement patterns)

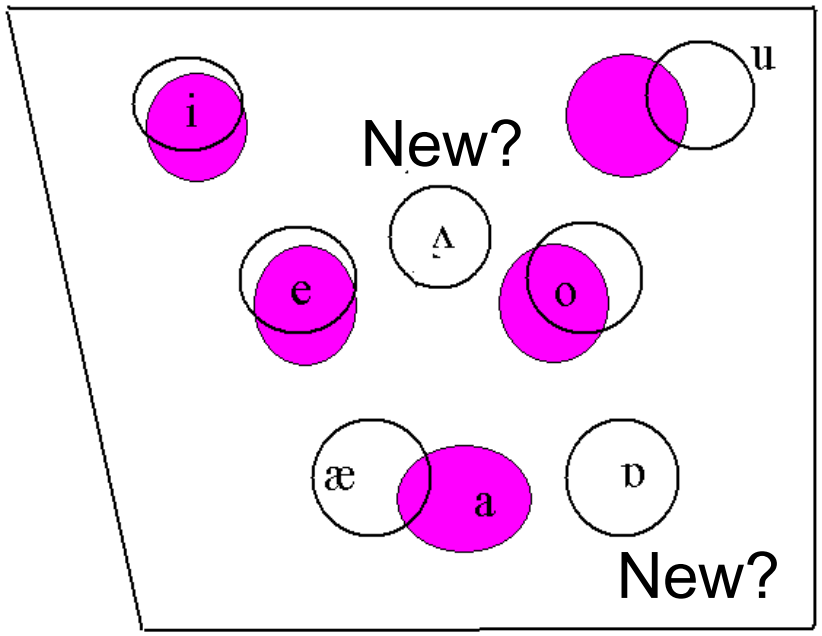
Here we see varying degrees of overlap in the acoustically defined vowel space between 5 L2 vowels and the 5 vowels of the L1. Two L2 vowels occupy space not exploited in the L1



Many researchers would immediately conclude that the two non-overlapping L2 vowels will be treated as “new”

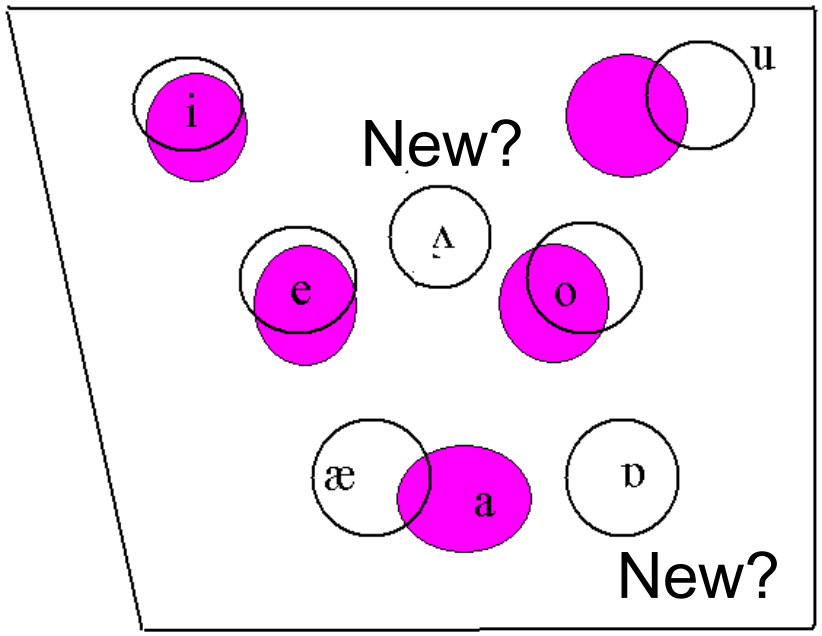
But wait. In the period 1984-1993 the SLM proposed that that vowels in an L2 could be classified as identical, similar, or new. This tri-partite division was abandoned for several reasons in 1994, over a decade ago.

Whether L2 learners will treat a vowel in the L2 as “new” will emerge over time. This determination can not be made by looking at plots of acoustic data.



The SLM regards perceived cross language phonetic dissimilarity as a continuum that must be measured in a perceptual experiment. (Basic technique: have listeners rate pairs of stimuli made up of one L1 vowel token and one L2 vowel token.)

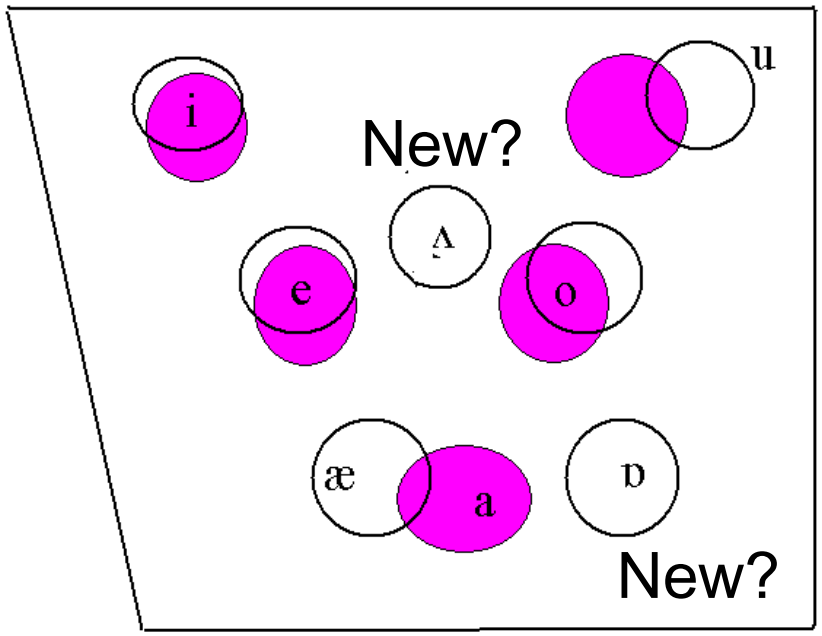
The L2 /ɒ/ would probably be at the high end of the dissimilarity continuum, and the L2 /i/ and /e/ at the low end of the continuum. However, this must be established empirically.



The SLM generates several predictions.

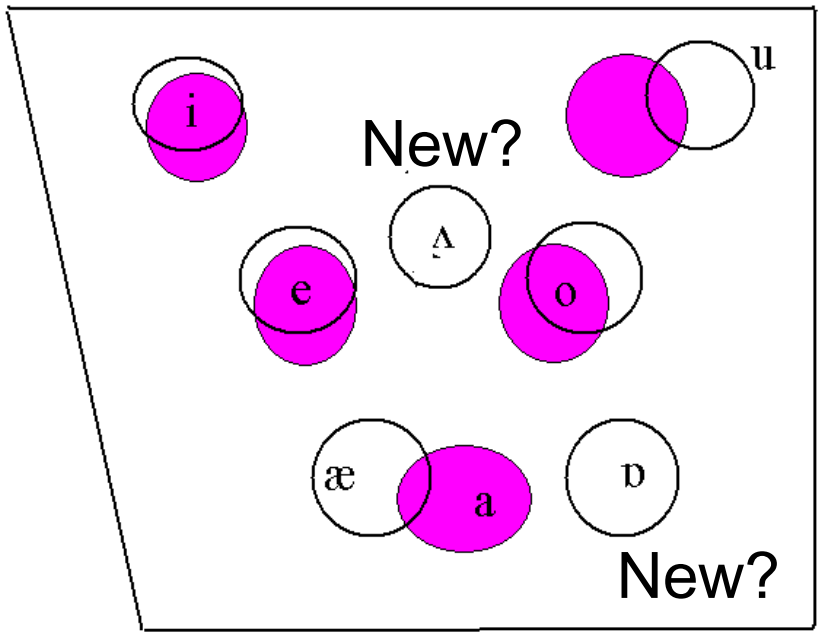
First, L2 vowels rated a phonetically similar to an existing L1 vowel will be produced fairly well in early stages of L2 acquisition. They are said to have gotten a “free ride”

Second, L2 vowels rated as very dissimilar from the closest L1 vowel might be produced poorly in the earliest stages of L2 learning. Perhaps they will be substituted using one or more L1 vowels that are adjacent to the L2 vowel.



However, in the “long run” (e.g., decades of predominant L2 use) such vowels should be produced more accurate than vowels that are less dissimilar from the closest L1 vowel. This is the expected outcome when new phonetic categories are established for certain vowels in the L2.

Note that without a time dimension – which might be simulated through groups differing in L2 experience – the SLM cannot generate testable predictions

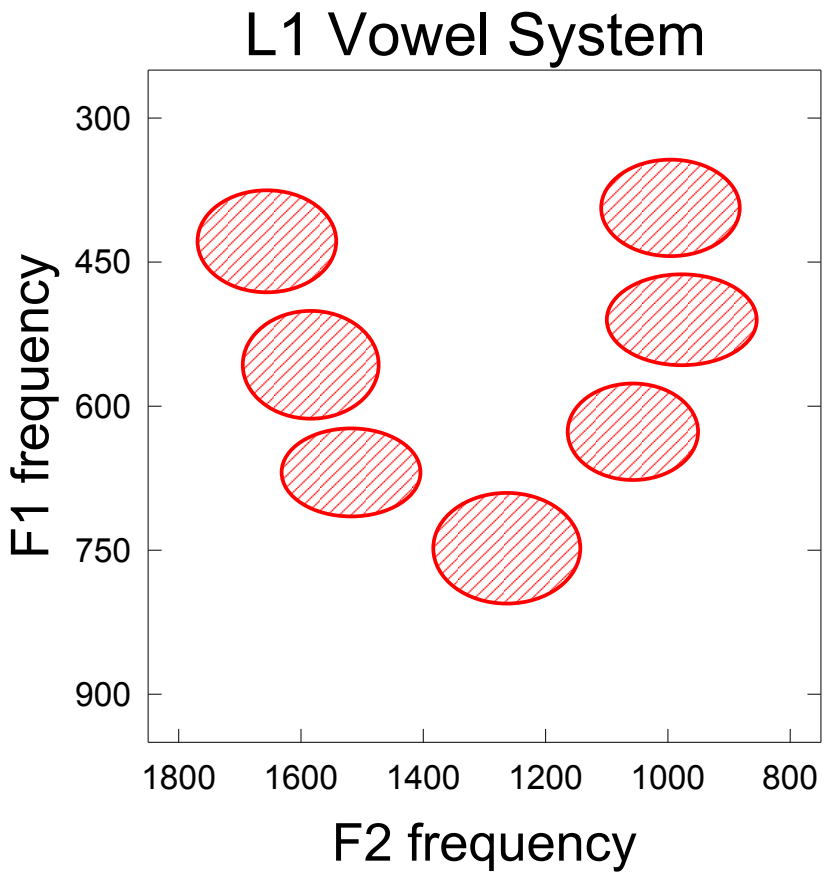


SLM hypotheses

- According to the SLM, degree of perceived cross-language phonetic dissimilarity exerts an import role in determining how successfully vowels in an L2 will eventually be produced.
- This is because perceived L1-L2 dissimilarity is the key to understand if new categories will or will not be established.
- However, the SLM posits that interactions between vowels in the combined L1-L2 vowel space also play a role.
- Once again, let's make some simplifying assumptions to facilitate the discussion

Illustrating assimilation, dissimilation

Lets imagine a 7-vowel L1 vowel system …



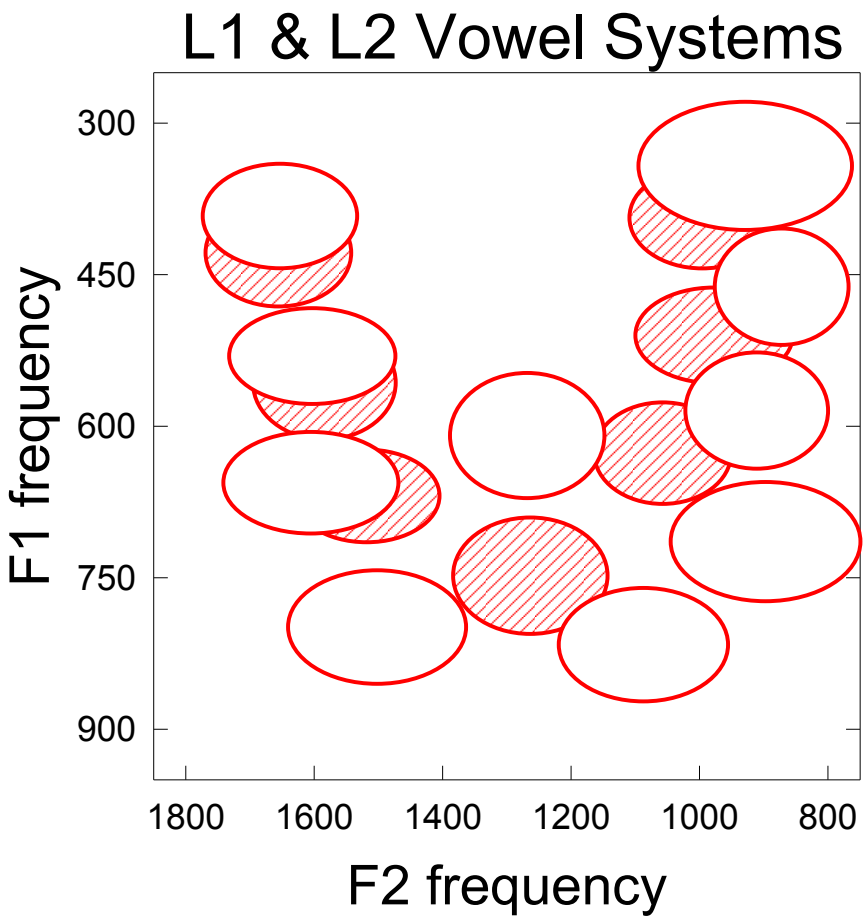
assimilation dissimilation

… and a 10-vowel L2 vowel system (unfilled ellipses)

The SLM proposes that when learners are unable to create a new category for an L2 vowel because it is too similar to an existing L1 vowel, the two vowels will eventually form a composite – coming to resemble one another

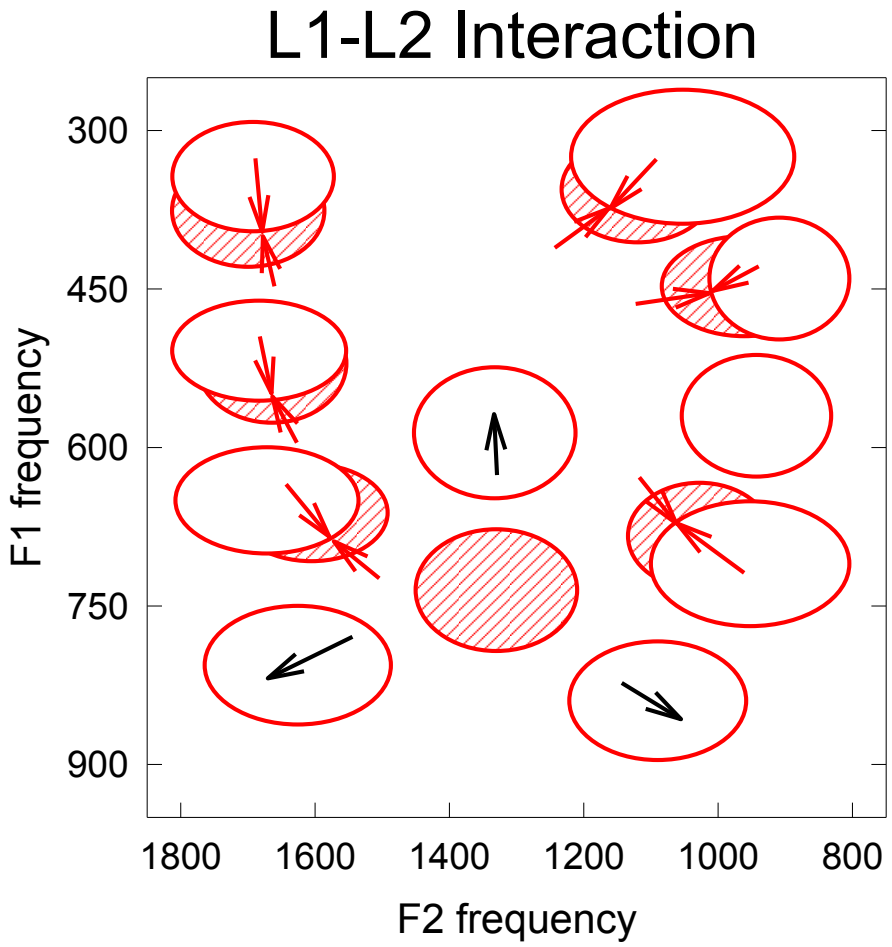
When categories ARE created for an L2 vowel, it and the closest L1 vowel are predicted to dissimilate in order to minimize perceptual confusions in the combined L1-L2 vowel space.

(Recall that fluent bilinguals often insert L1 materials into the L2, and vice versa.)



Observing assimilation and dissimilation processes in L2 acquisition requires a lot of data and time. These processes are probably observed best over real rather than apparent time, i.e., in longitudinal research rather than research comparing groups differing in (presumed) L2 input and use.

The figure at the right illustrates what might happen when learners of the L1 (7 vowels) learn the hypothetical L2 (10 vowels): dissimilation of three L2 vowels from neighboring L1 vowels, and assimilation of the remaining seven L1 vowels



Outline

1. Purpose of the Speech Learning Model (SLM)
2. Historical background
3. Core aspects of the SLM
4. **Testing SLM predictions**
5. How to falsify the SLM
6. What is most needed now?

Testing SLM predictions

Hypothesis: When category formation does not occur – because an L2 sound differs insufficiently from the closest L1 sound, the L2 sound and the closest L1 sound will assimilate

- Productions of the L2 sound will continue to resemble the L1 sound
- Productions of the L1 sound will shift in the direction of the L2 sound

Testing SLM predictions

Flege (1987) examined the production of French and English /t/ by members of two groups

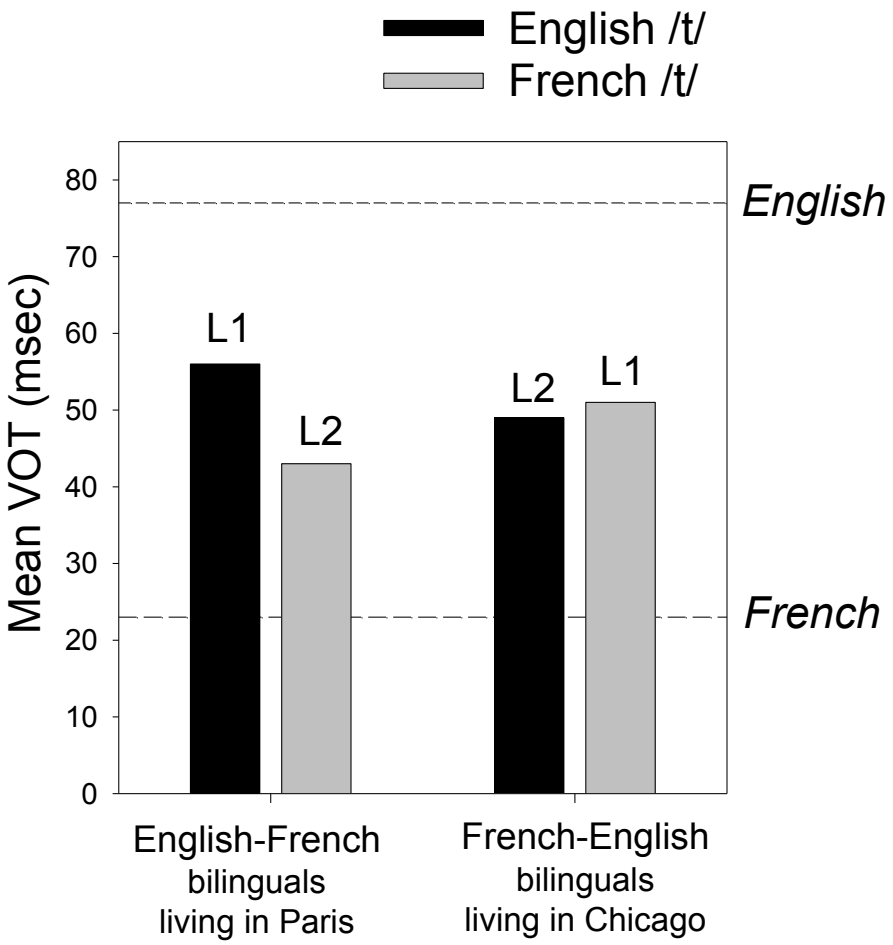
- American women who had lived in Paris for *M* = 10 years
- French women who had lived in Chicago for *M* = 10 years

Flege (1987) measured the VOT of /t/ produced in the initial position of words in English and French by two groups of bilinguals as well as the productions of French and English monolinguals.

The dashed lines indicate the mean value of stops produced by the monolinguals in English, French

Left: mean VOT of productions of the English-French bilinguals in English (L1) and French (L2)

Right: production of the French-English bilinguals in English (L2) and French (L1)



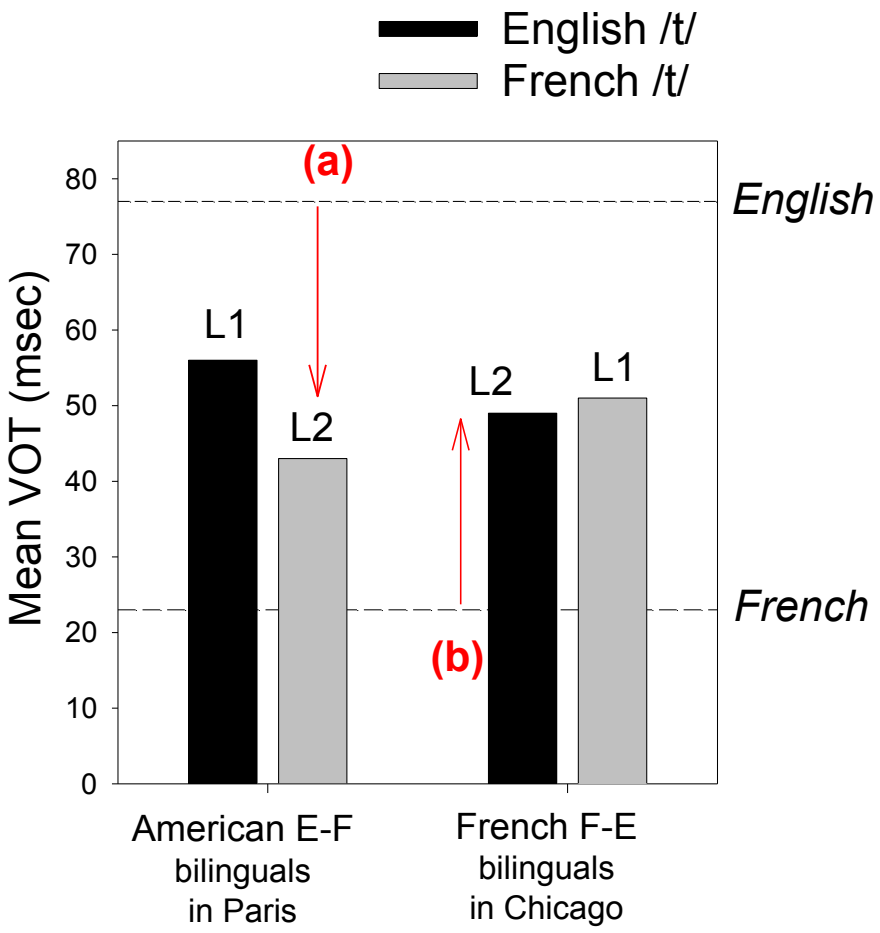
Flege (1987)

The red arrows indicate productions in the L2

(a) For the American women in Paris, French /t/ is produced with shorter VOT (English-like) values than would be typical for French

(b) For the French women in Chicago, English /t/ produced with longer VOT values than would be typical for English

Both groups learned something in the L2, making modifications in the right directions, but neither group produced stops in the L2 accurately



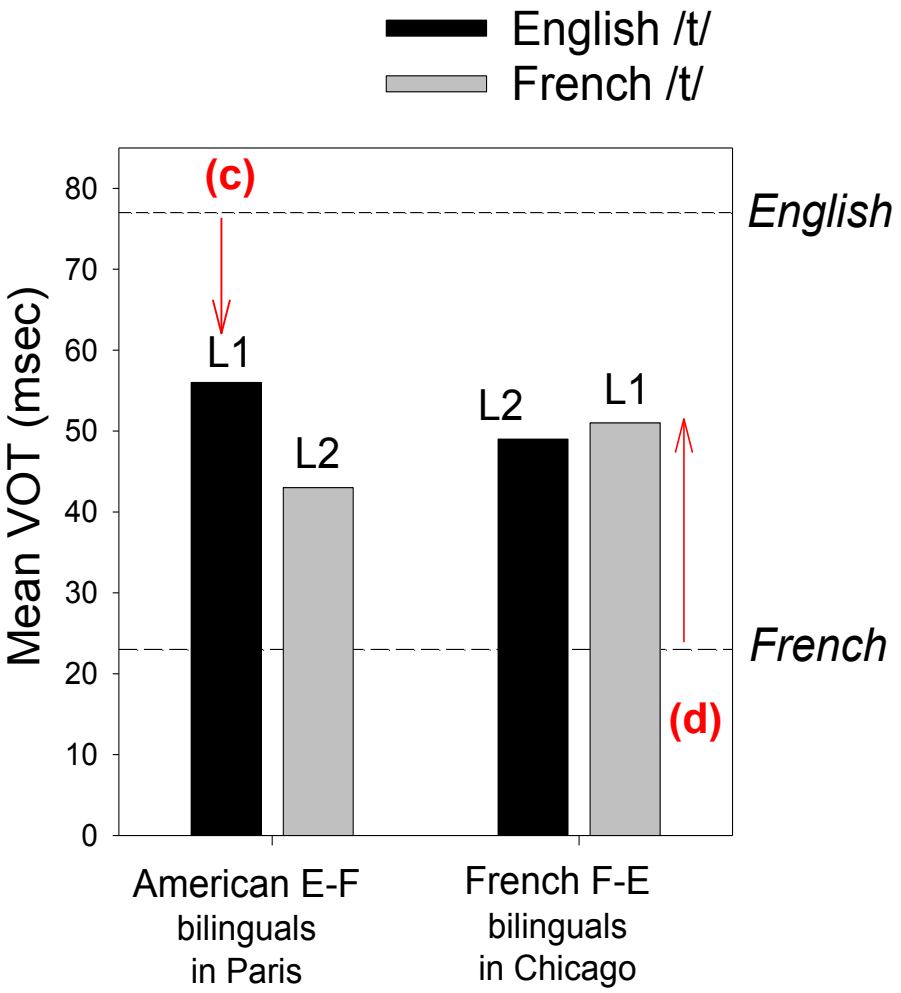
Flege (1987)

The red arrows here indicate productions in the L1

(c) For the American women in Paris, English /t/ was produced with shorter (French-like) VOT values than would be typical for English

(d) For the French women in Chicago, French /t/ was produced with longer (English-like) VOT values than would be typical for French

This supports the hypothesis that L1 and L2 phonetic elements exist in a common space and mutually influence one another



Testing SLM predictions

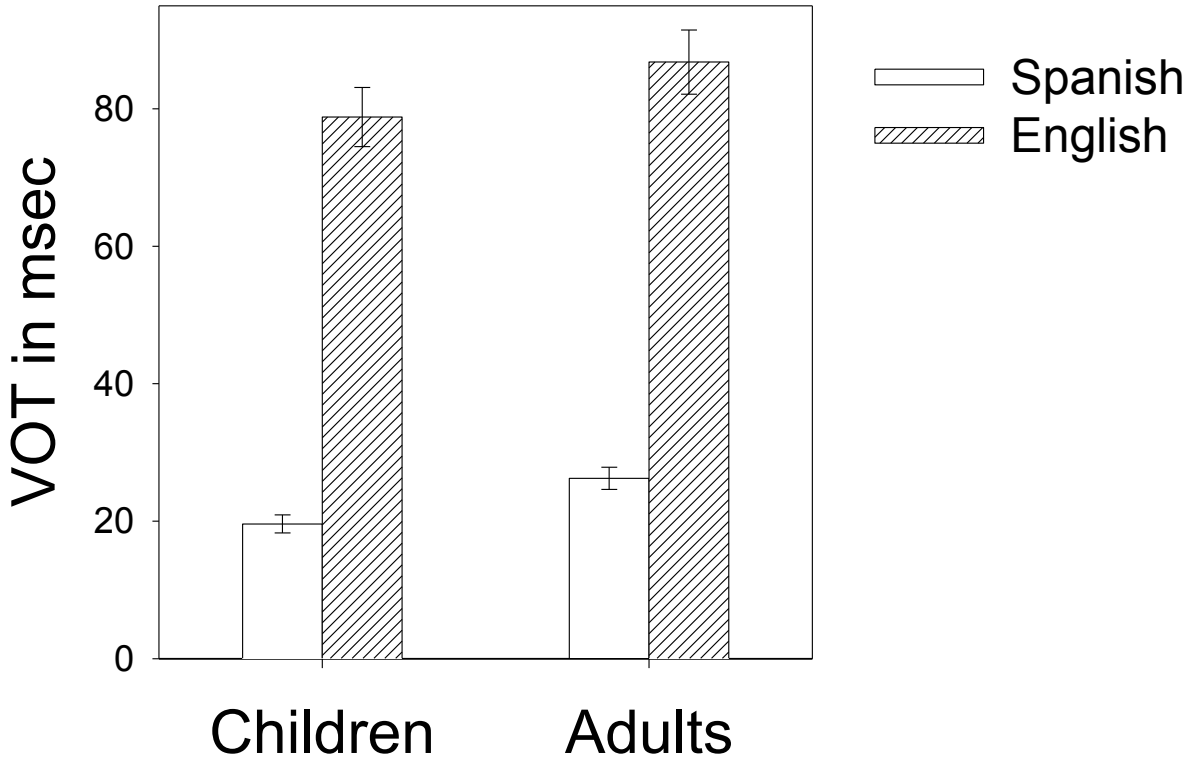
Prediction: When a new category is formed for an L2 sound, it and/or the nearest L1 sound may dissimilate so that they are more distant from one another in phonetic space. This process may make production of one or both phonetic elements less accurate from the point of view of normative values obtained from monolinguals.

This renders concrete the adage (see work by F. Grosjean) that a bilingual cannot be “two monolinguals in one head”

Testing SLM predictions

- Flege & Eefting (1986, 1987) examined production of phonologically voiceless stops, /p t k/, in Spanish and English words.
- They recorded four groups of participants
  - Spanish monolinguals: adults & children living in Puerto Rico
  - English monolinguals: adults & children in Alabama

The results of Flege & Eefting (1986, 1987) replicated findings obtained in previous cross-language phonetic research, showing that phonologically voiceless stops are realized differently at the phonetic level. Spanish monolinguals (both adults and children) were found to produce /p t k/ as short-lag stops having substantially shorter VOT values than did stops produced by English monolinguals (both adults and children)

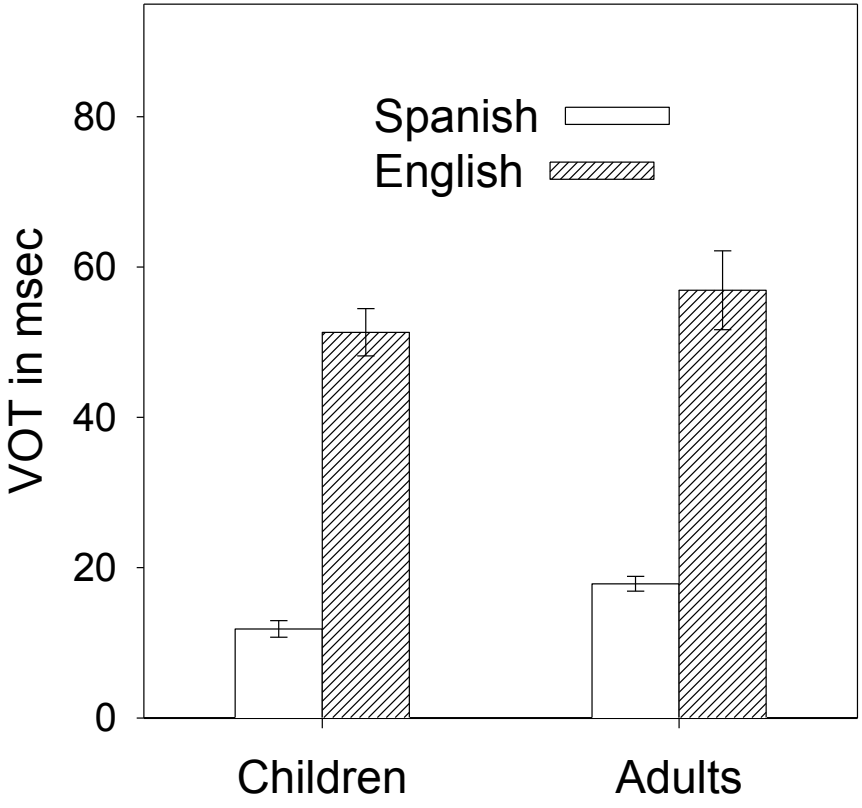


Testing SLM predictions

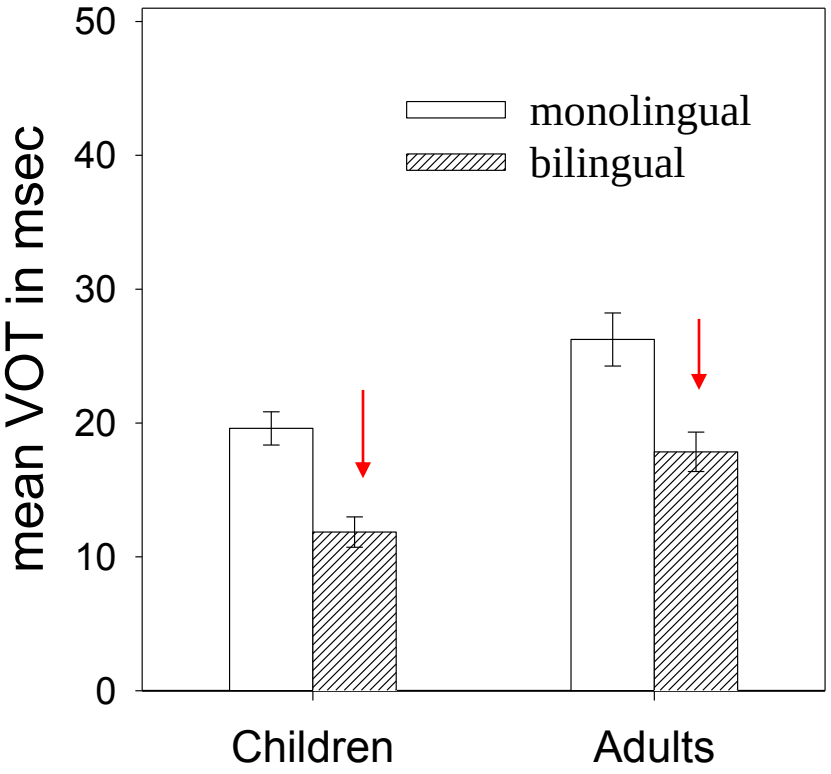
Flege and Eefting (1986, 1987) also recorded two groups of early bilinguals

- Native Spanish adults who had begun learning English when attending a bilingual school in an essentially monolingual Spanish community in Puerto Rico
- Native Spanish children who were currently enrolled in the same bilingual school in Puerto Rico where English was used as the language of instruction in most classes

Flege & Eefting (1986, 1987) . Both groups of Spanish-English bilinguals produced the phonologically voiceless stops /p t k/ with substantially longer VOT values in English L2 than in Spanish L1. This is a positive outcome for those who endorse second-language instruction at an early age



Flege & Eefting (1986, 1987) . Here we compare productions of Spanish /p t k/ by two groups of Spanish monolinguals (children, adults) and by the two groups of early Spanish-English bilinguals (children, adults). As expected, the monolinguals produce Spanish /p t k/ as short-lag stops having VOT values less than 30 msec. Both groups of bilinguals produce voiceless Spanish stops with even shorter VOT values. This provides evidence of dissimilation.



Testing SLM predictions

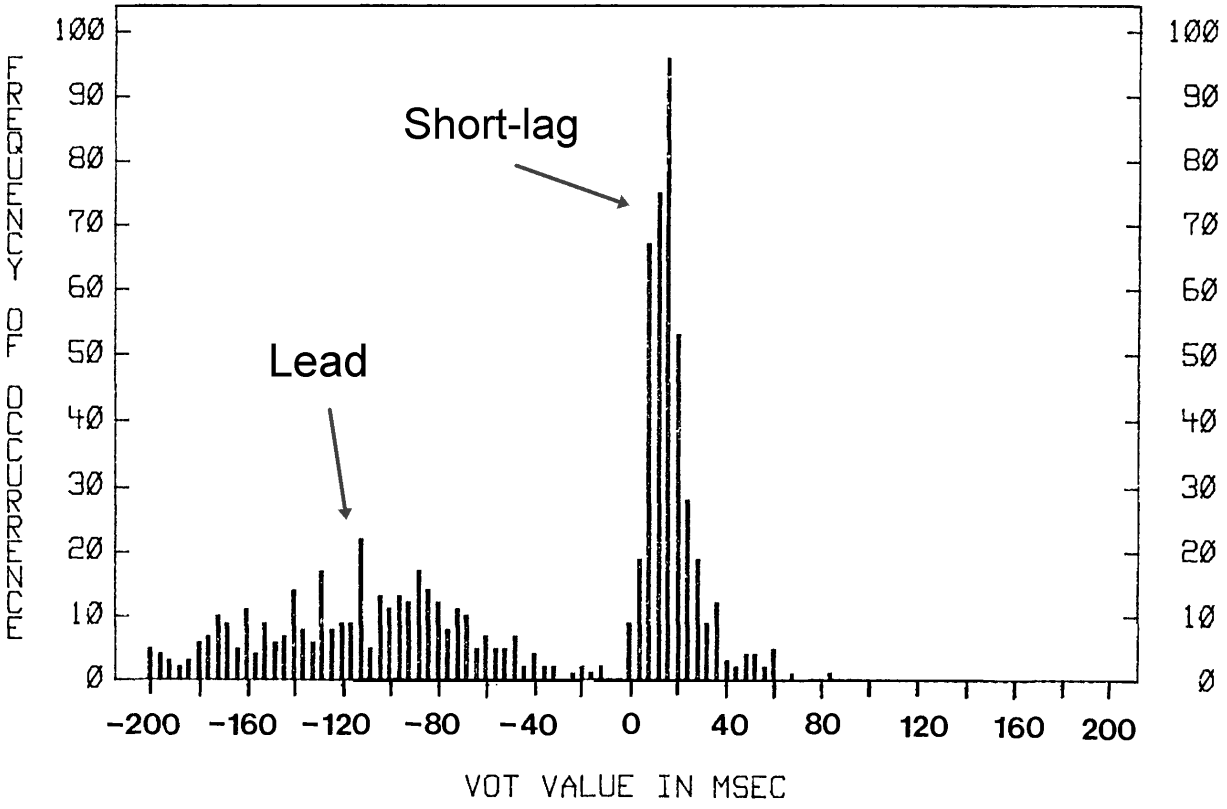
According to the SLM, children are more likely to form phonetic categories for L2 sounds than adults because their L1 categories are less fully developed and represent weaker “attractors” for sounds encountered on the phonetic surface of an L2

The SLM maintains, however, that the processes and mechanisms subserving the completely successful acquisition of an L1 by monolinguals are used by persons who acquire an L2, even in adulthood. Thus, according to the SLM, even adults retain the capacity to form new categories for L2 sounds if given the right kind of input and the opportunity (time) needed to do so.

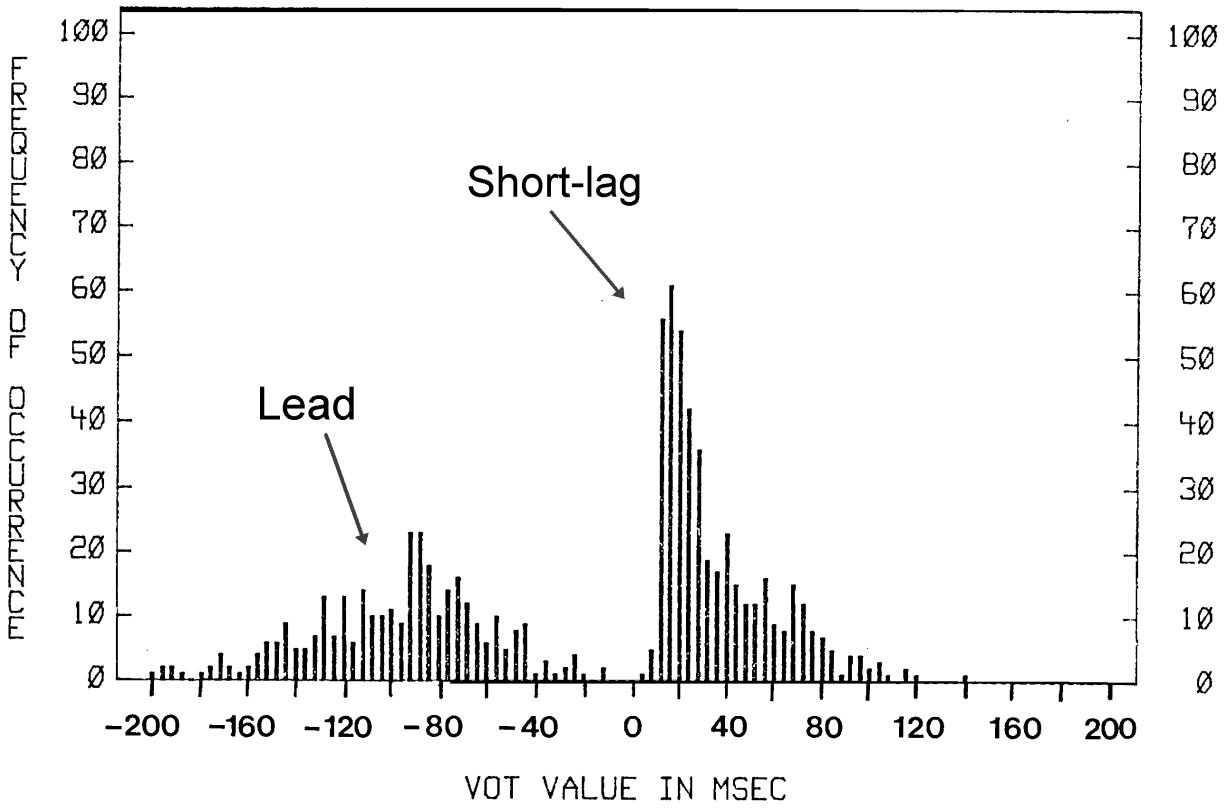
Testing SLM predictions

- Flege & Eefting (1988) provided evidence of category formation for /p t k/ by early Spanish-English bilinguals
- These authors tested
  - Spanish-English bilinguals
  - Spanish & English monolinguals
- The participants' task was to imitate members of a synthetic /d/ to /t/ continuum made up of stimuli differing in VOT, which ranged from lead (pre-voiced) values to long-lag values (an aspirated [t^h])

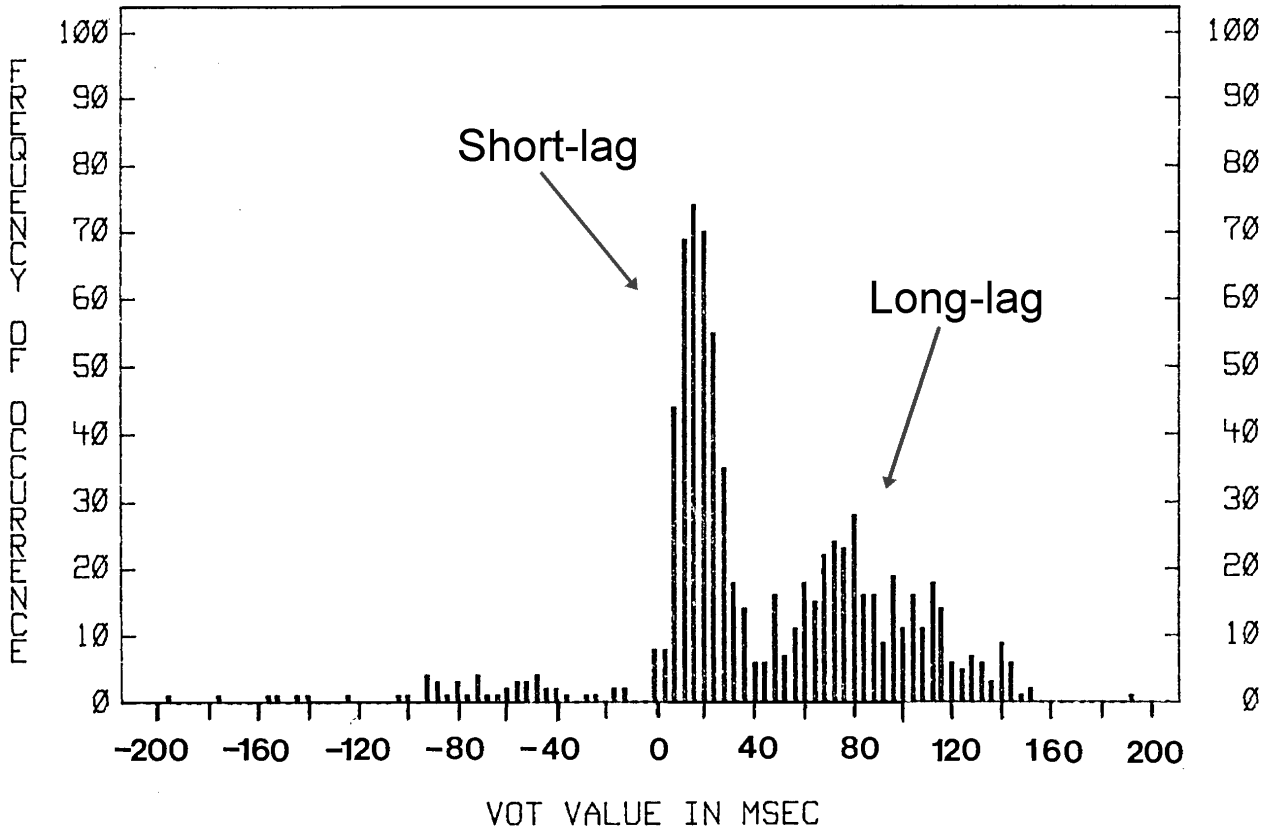
Flege & Eefting (1988) . This figure shows the distribution of VOT values in the 900 imitations of members of a VOT continuum by Spanish monolingual children. The children did not accurately reproduce the VOT values present in the stimuli. They instead tended to produce stops having VOT values in the lead (pre-voiced) range or with VOT values in the short-lag range. Both are typical for Spanish.



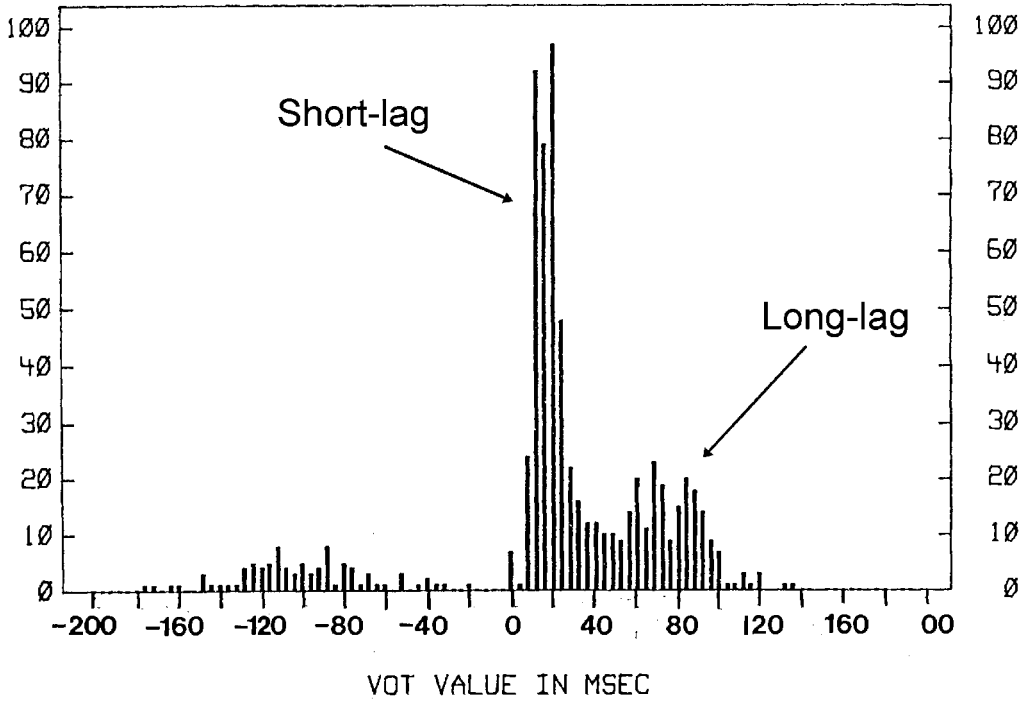
Flege & Eefting (1988) . Here are the results for the monolingual Spanish adults, who did not accurately imitate VOT values in the stimuli. We again see two distributions of VOT values in the imitation responses, one typical for the Spanish /d/, the other for the /t/ of Spanish. The authors concluded that Ss rapidly classified the initial stops in the perceptual stimuli in terms of phonetic categories established in L1 acquisition and then produced them according to the VOT value specified by the phonetic categories



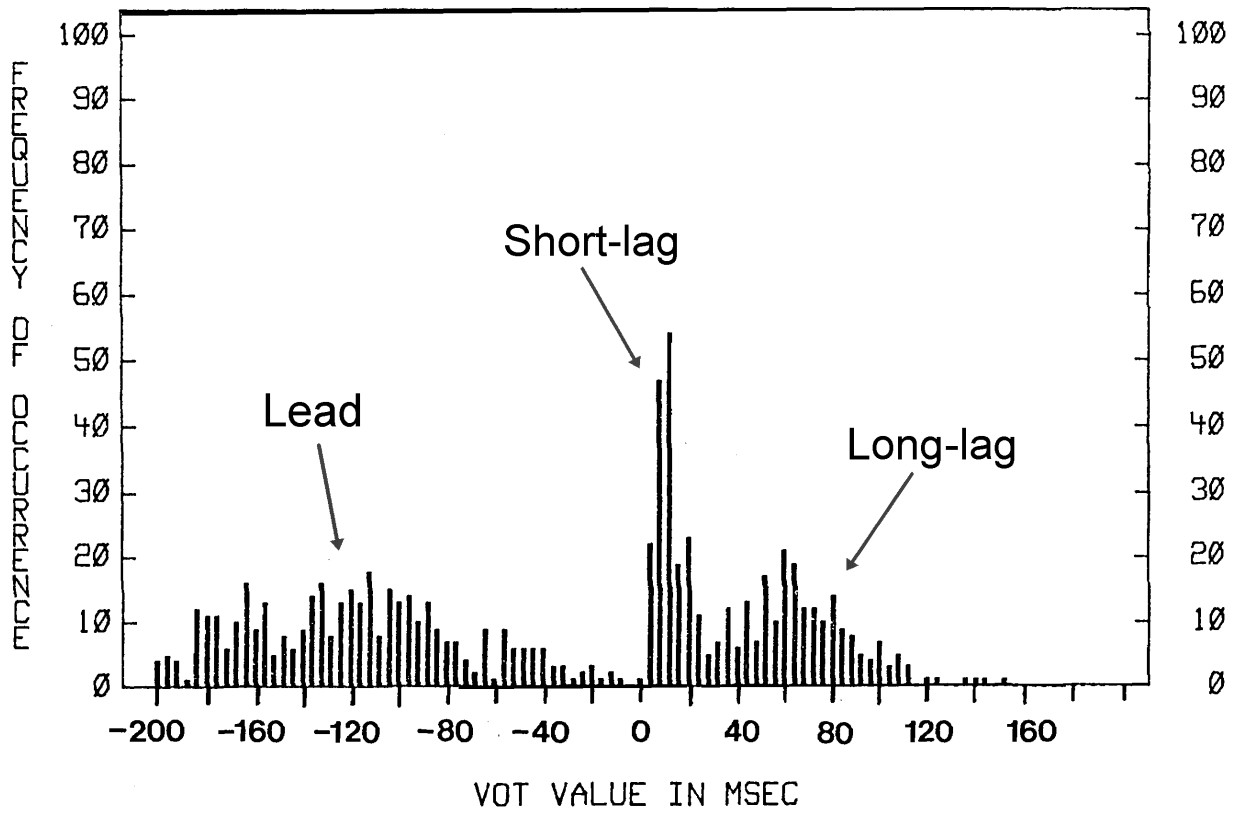
Flege & Eefting (1988) . Dramatically different results can be seen in the distribution of the 900 imitations of the same VOT continuum by monolingual English children. These children produced few pre-voiced (lead VOT) stops. They showed two distributions in the lag VOT region. The short-lag values are typical for English /d/ and the long-lag values are typical for English /t/



Flege & Eefting (1988) . Much the same pattern of results was for the monolingual English adults.

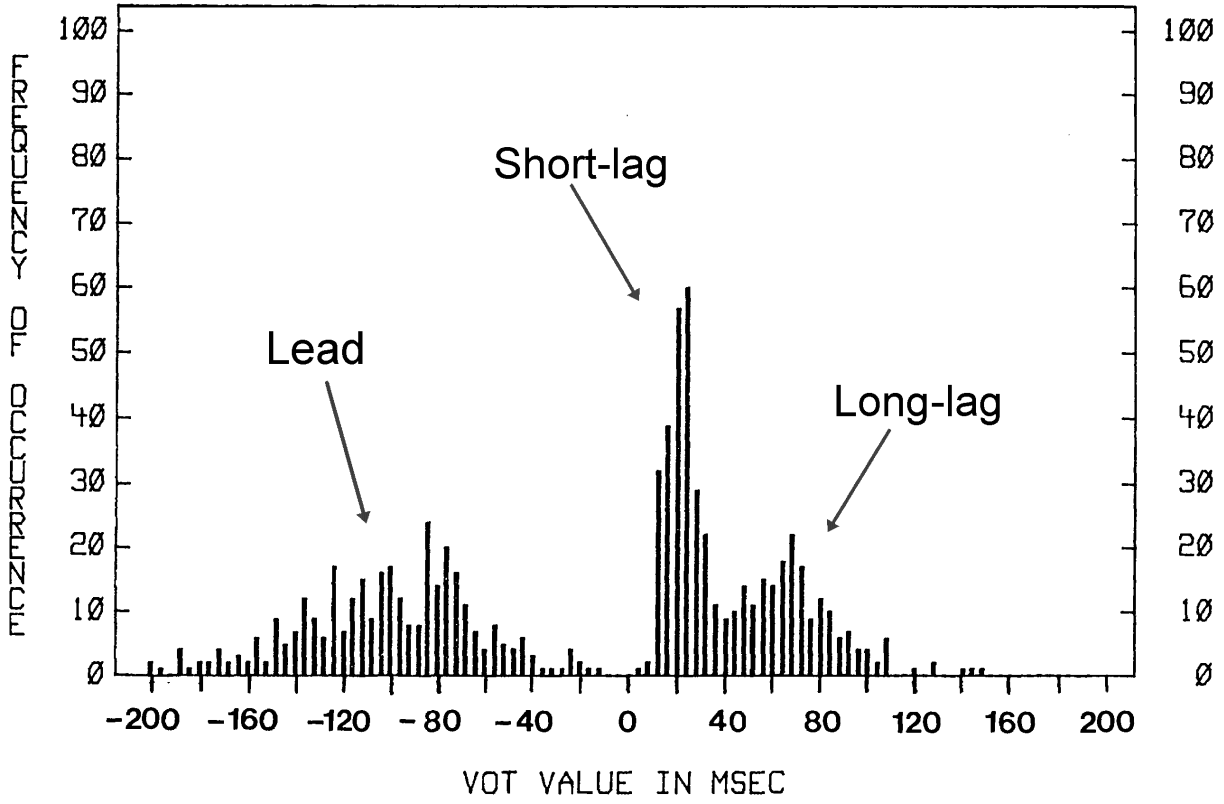


Flege & Eefting (1988) . Distribution of VOT values obtained during imitation of the VOT continuum by Spanish-English bilingual children. These children three distinct distrubutions of VOT values in their imitation responses. The authors interpreted this as evidence for the existence of three distinct phonetic categories



Spanish-English bilingual children

Flege & Eefting (1988) . The conclusion regarding the existence of three distinct phonetic categories was reinforced by the finding obtained for the bilingual Spanish-English adults. The authors concluded that the Spanish-English bilinguals retained two phonetic categories established during L1 acquisition (lead, short-lag) and added a third phonetic category needed for the long-lag stops of English when they acquired English as an L2.



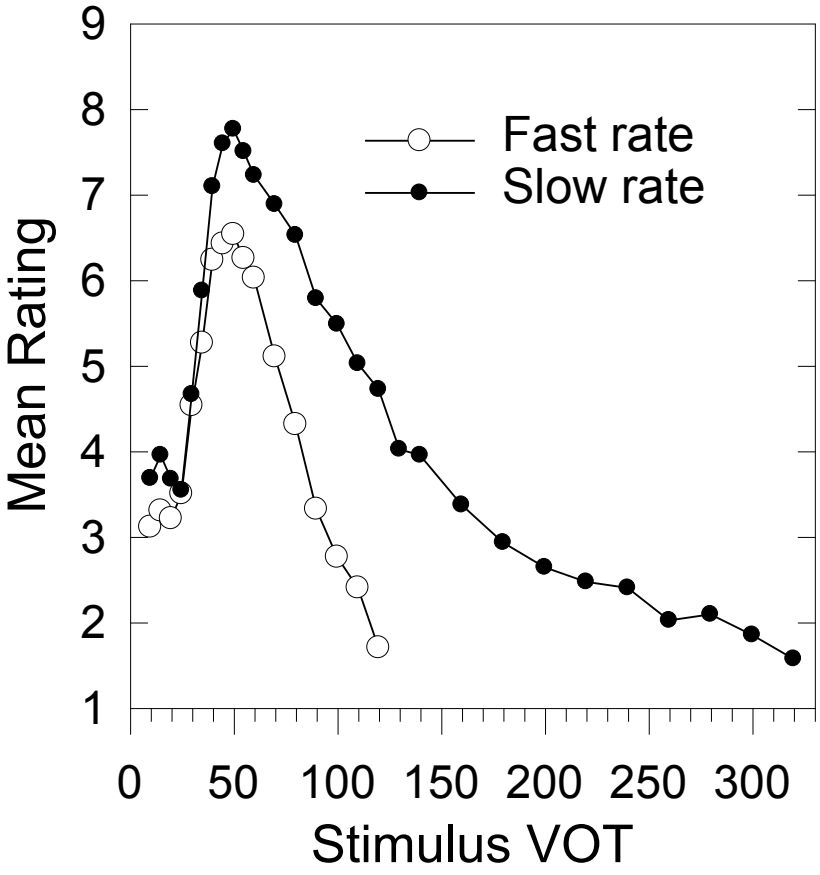
Testing SLM predictions

- Flege, Schmidt & Wharton (1996) and Schmidt & Flege (1995) provided evidence of category formation for dthe long-lag /p/ of by a few late Spanish-English bilinguals
- Participants rated the randomly presented member of a VOT continuum for category “goodness”
- In English, VOT values in word-initial /p/ tokens shorten as speaking rate increases
- The authors created two VOT continua, one that stimulated speech produced at a slow rate, the other at a faster rate.

For both continua, ratings obtained from native English Ss increased (indicating a better perceived "goodness" as instances of the /p/ category) as VOT increased, then a systematic decrease in the ratings as VOT values in the stimuli went beyond values typical for English.

The pattern indicates a match between production and perception. The Ss accepted longer VOT values as "good" in the slow rate continuum accordance with the fact that VOT values are longer in speech produced at a slow rate

Native speakers of English show rate-dependent processing of stops differing in VOT

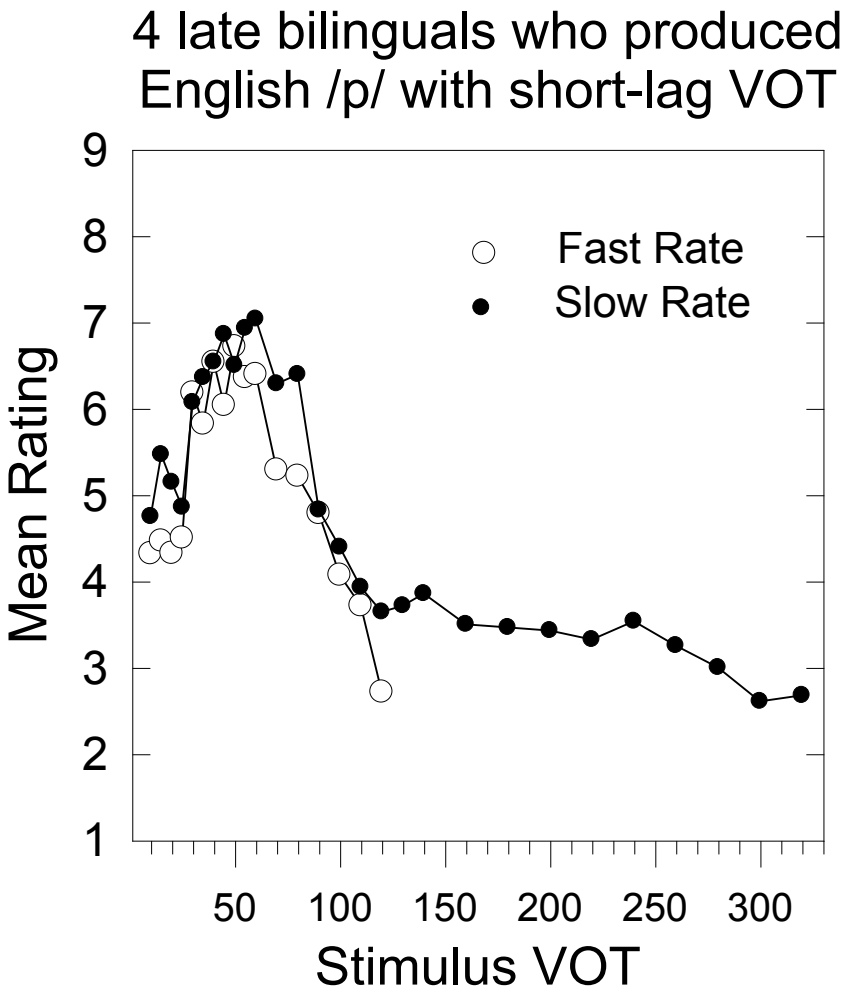


Native speakers of Spanish who had learned English as adults (late bilinguals) were also tested.

There is relatively effect of variations in speaking rate on VOT in Spanish short-lag stops.

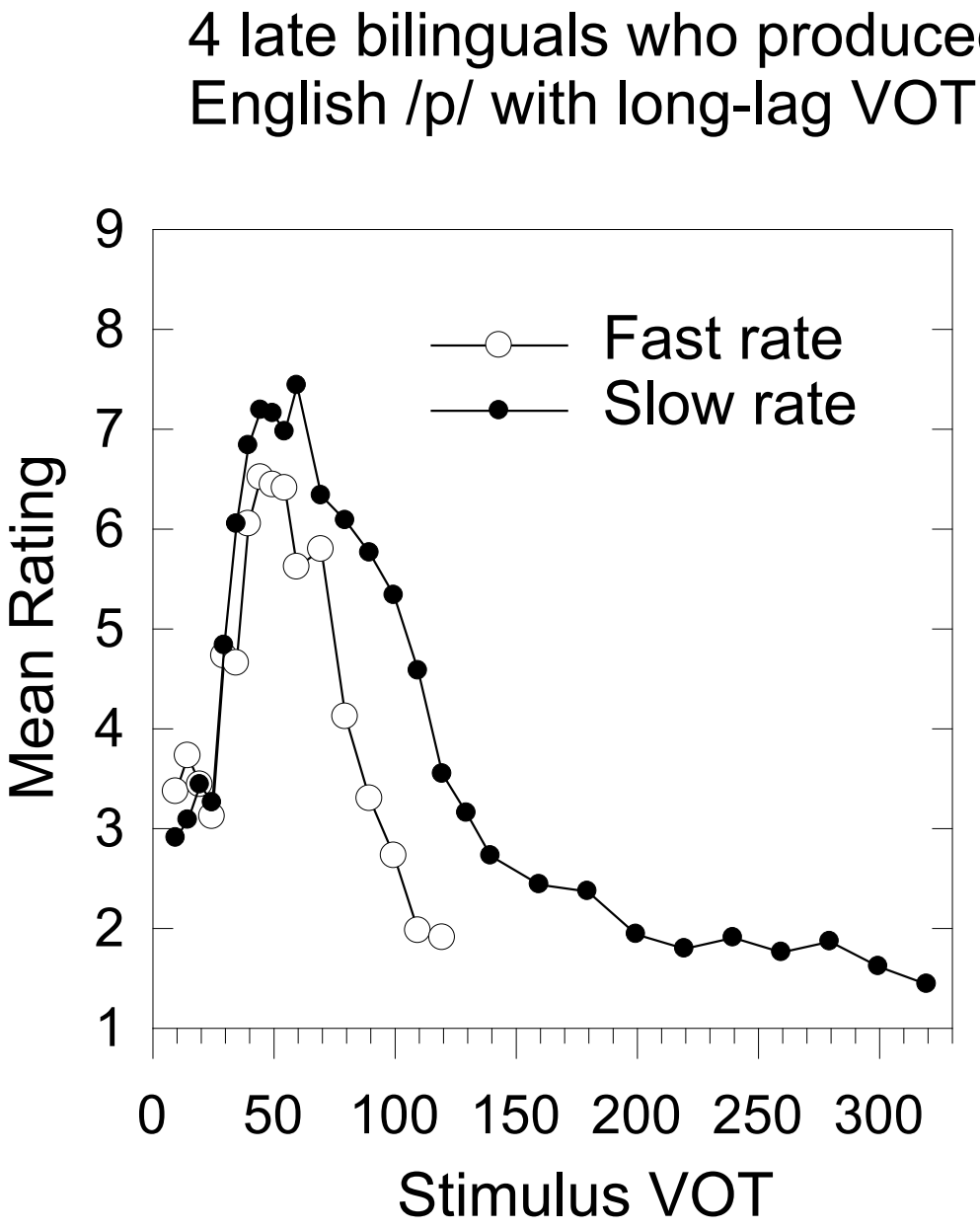
Here we see the results obtained for 4 of 15 late bilinguals tested. They produced English /p/ with short-lag values that are typical for Spanish (range of mean production values = 13 – 18 msec)

The perception data shown here – no indication of a speaking rate effect on the goodness judgements coincides with the production data. These late bilinguals had not created new phonetic categories for English /p/



A very different picture emerged for the 4 (of 15) late bilinguals who managed to produce English /p/ with long-lag VOT values.

The four late bilinguals showed a clear speaking rate effect on the goodness judgements. This, taken together with the production results, suggests that these latte bilingual had created new phonetic categories for English /p/



Outline

1. Purpose of the Speech Learning Model (SLM)
2. Historical background
3. Core aspects of the SLM
4. Some predictions generated by the SLM
5. **How to falsify the SLM**
6. What is most needed now?

How to falsify the SLM

- A theoretical model is interesting and useful only to the extent that it can be falsified, or to the extent that research findings can be obtained that diverge partially from predictions, permitting the model to be improved.
- Here I'll suggest some ways the SLM might be falsified.
- But first, a few comments are necessary regarding the adequacy of testing methods.

Adequate measurement of L2 speech are needed

For examinations of L2 segmental production we need to ask

- Do native speakers of the target L2 hear segments produced by L2 learners as they were intended (categorical judgment)? If so, do the native listeners rate the L2 segments as “distorted” or “foreign accented” (qualitative)
- When measured acoustically, do relevant dimensions in the target L2 segments differ significantly from native speakers’ productions?

Adequate measurement of L2 speech are needed

For examinations of L2 segmental perception we need to ask:

- Do L2 learners correctly identify the segment?
- If so, do they do so as rapidly as native speakers?
- Do they show a greater influence of semantic/lexical context than native speakers?

Appropriate participants must be tested

- It is unreasonable to conclude that L2 learners are unable to accurately produce target L2 sounds if they have only (or mostly) heard inaccurate productions of the L2 target sound.
- It takes children years to learn to accurately produce phonetic segments in the L1. It is inappropriate to conclude that L2 learners are unable produce an L2 segment accurately until they have received at least as much input as is needed by monolingual L1 learning children to learn to produce the segment accurately

Appropriate participants must be tested

To establish upper limits on phonetic learning in an L2, it makes sense to find participants who use the L1 relatively seldom

- L2 learners, especially late learners, tend to continue using the L1 frequently
- For such participants, divergences from native speakers of the target L2 may reflect psycholinguistic difficulty in separating the L1 & L2 subsystems rather than inability to learn phonetically

Assess perceived cross-language dissimilarity

- Most predictions generated by the SLM are based on knowing what it the degree of perceived phonetic dissimilarity from the closest L1 sound(s)
- Acoustic phonetic measurements, especially those derived from published studies of the L1 and L2 employing different techniques, cannot substitute adequate perceptual measures

Assess perceived cross-language dissimilarity

Currently used procedures are inadequate

- Ss hear a specific L2 phone
- classify it in terms of abstract (imagined) L1 category
- Rate it's goodness of fit to the (imagined) L1 category
- How do we know the experiment know that the participant being tested is really "thinking about" the L2 phonetic segment specified by the experimenter?

Assess perceived cross-language dissimilarity

A paired comparison technique is much better suited for obtained perceptually based estimates of cross-language phonetic dissimilarity

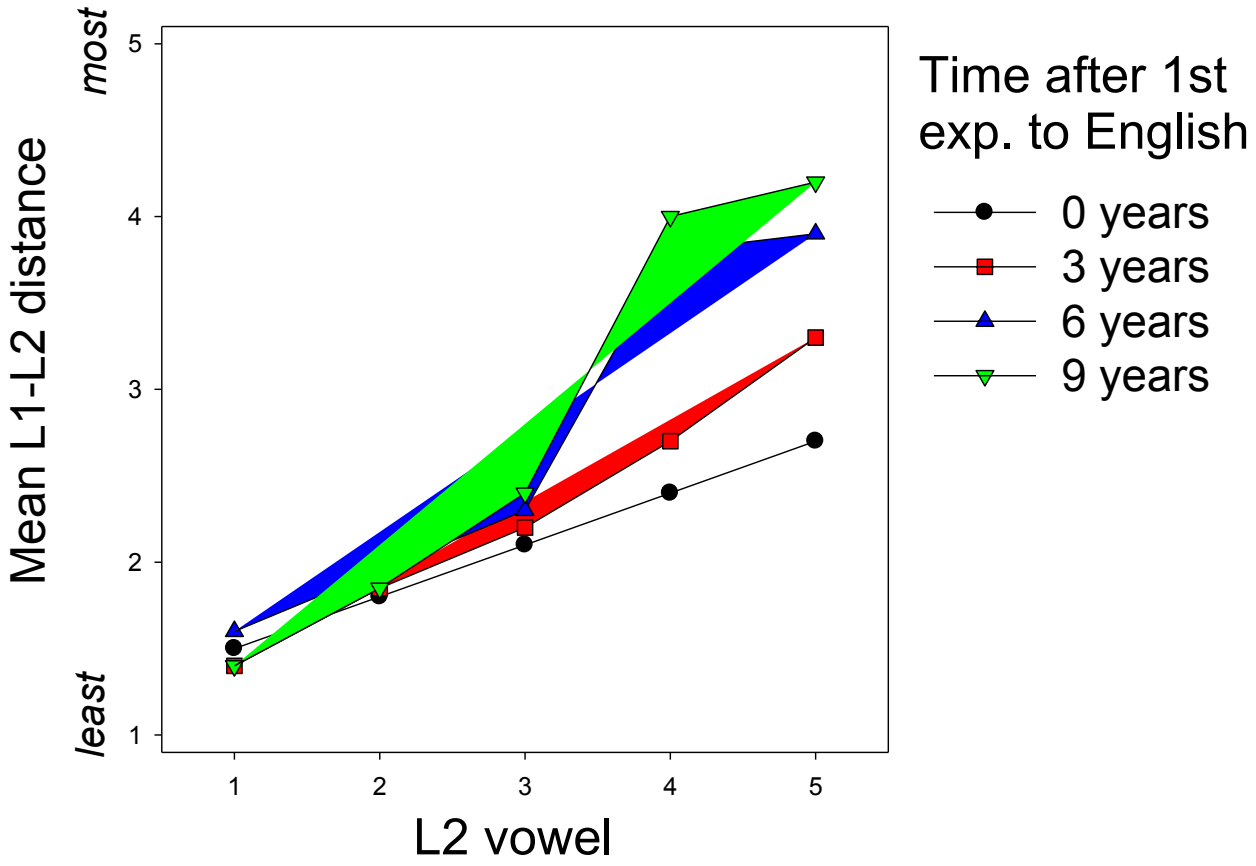
- This requires obtaining samples of L1 and L2 phones produced in similar contexts in the L1 and L2 by monolingual native speakers of those languages
- Each pair of stimuli contains an L1 sound and an L2 sound.
- To ensure that the full range of possible judgments are obtained, some pairs should consist of tokens from the same portion of the phonetic space, while other should be drawn from distant portions of the phonetic space

Assess perceived cross-language dissimilarity

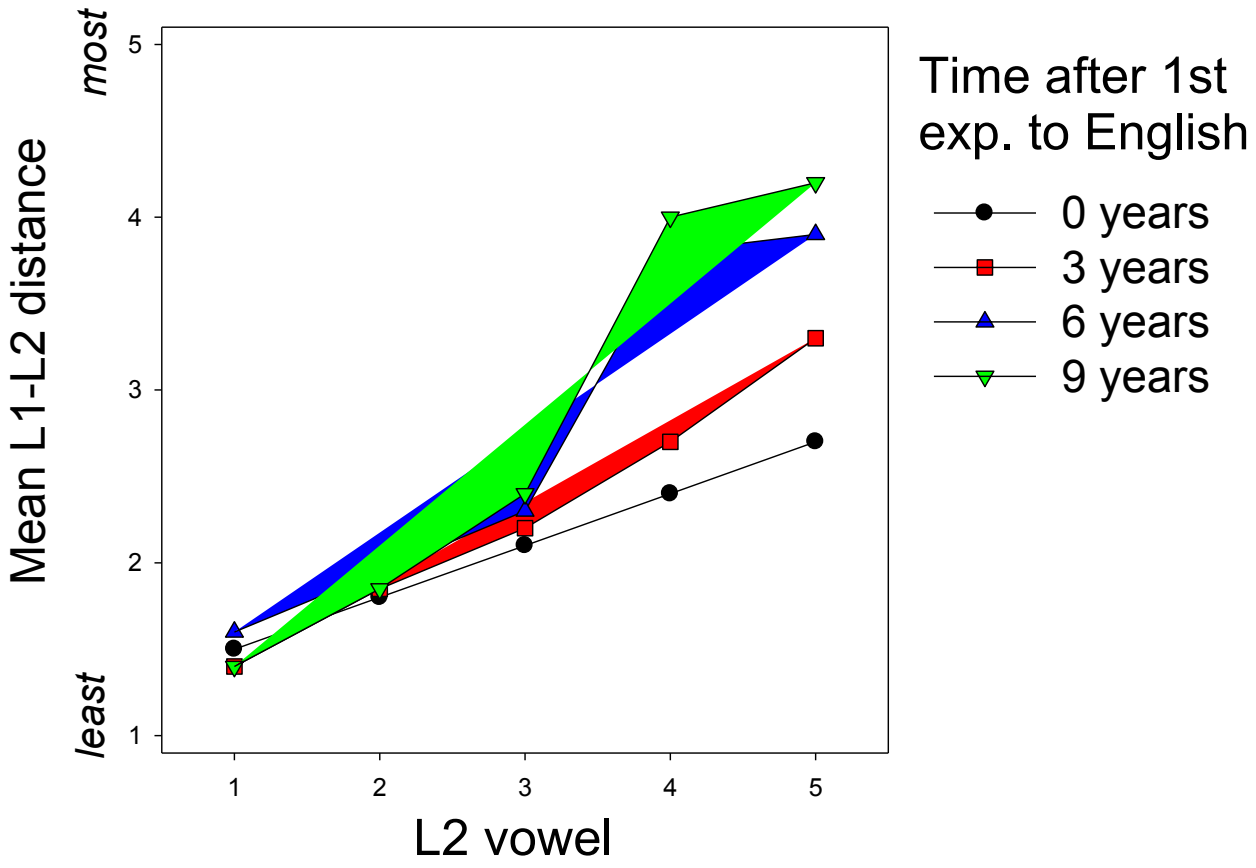
Let's consider the results of a hypothetical experiment which has the aim of measuring the perceived phonetic dissimilarity (distance) of the 5 vowels in an L2 from the closest vowel in the listener's L1.

- Each L2 vowel will have been paired with a variety of vowels in the L1.
- However, let's consider just the ratings obtained for the pairings between each of the 5 L2 vowels and what turns out to be the closest L1 vowel.

Hypothetical cross-language phonetic dissimimilarity ratings. The data for “0 years” indicate the five mean dissimilarity ratings obtained from monolingual speakers of the L1 who have no prior exposure to the L1. Their mean ratings have been rank ordered. The L2 vowel marked “1” is the most similar to the closest vowel in L1, the L2 vowel marked “5” is the most dissimilar



To these hypothetical data, we had cross-language phonetic dissimimilarity ratings obtained from L1 speakers who have been exposed to the hypothetical L2 for varying amounts of time. A marked increase in perceived dissimilarity of one or more L2 vowels with higher levels of exposure would be indicative of L2 category formation.



How to falsify the SLM

- L2 target vowels that show an augmentation in perceived phonetic dissimilarity in real or apparent time do not show a greater improvement in production accuracy than do L2 target vowels who perceived dissimilarity remain constant over real or apparent time

How to falsify the SLM

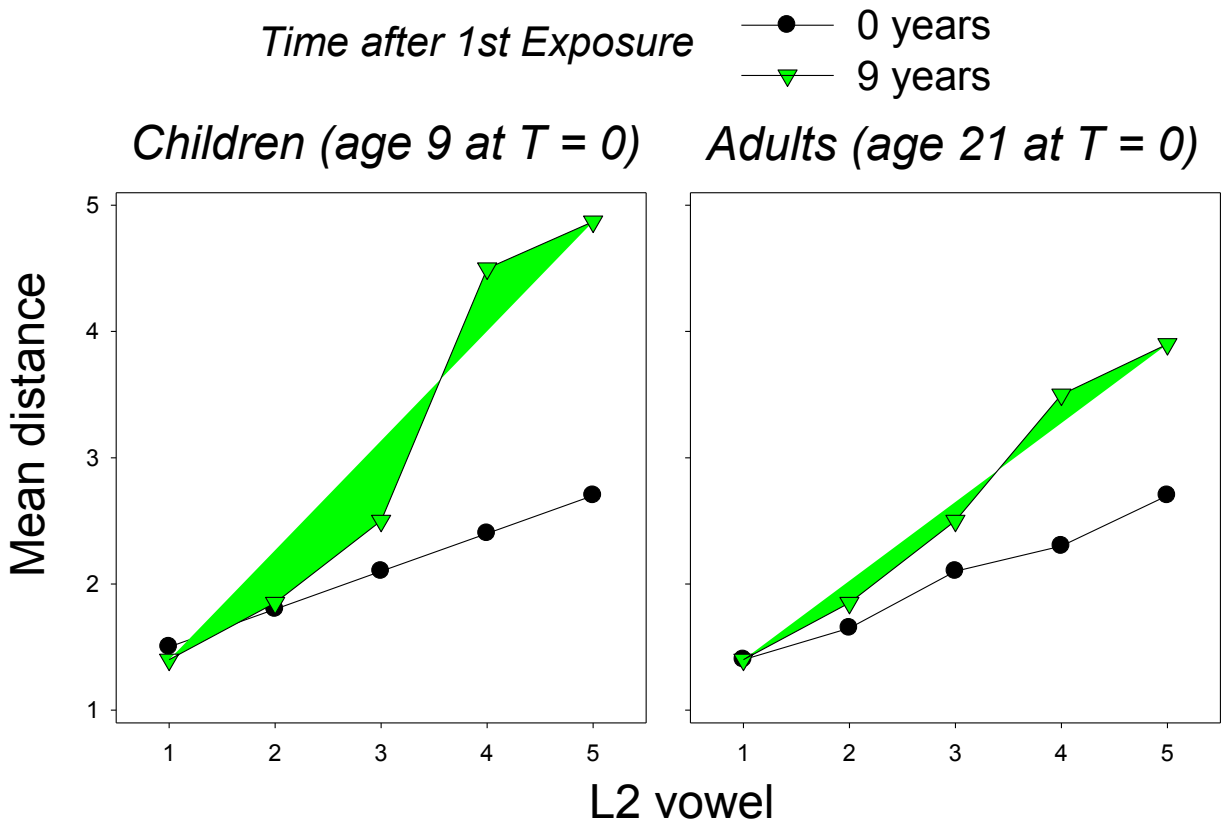
According to the SLM, category formation (CF) becomes less likely as L1 categories develop

- **Hypothesis**: as L1 categories become more robust through childhood, they “more powerful attractors” for L2 speech sounds (see Baker, Trofimovich, Mack & Flege, 2002)

How to falsify the SLM

- Let's imagine another hypothetical experiment that would be suitable to falsify this core hypothesis of the SLM.
- It is a repetition of the earlier experiment, this time with a direct comparison between child and adult learners of an L2
- The two groups differ in age: 9 vs. 21 years.
- Perceived dissimilarity scores are obtained at two sessions separated by 9 years, and so the Ss are 18 and 30 years of age at Time 2

How to falsify the SLM



The SLM predicts that there will be a greater augmentation of perceived cross-language dissimilarity for at least the most distant vowel by the children compared to the adults.

How to falsify the SLM

- The observation of a greater augmentation in perceived L1-L2 distances for early than late learners – or for more of the early than the late learners (considered individually) would support the hypothesis that category formation (CF) for L2 sounds becomes less likely as the L1 system develops
- The conclusion regarding CF would be confirmed if the vowels showing evidence of CF demonstrated greater improvements in production accuracy than did those showing no augmentation of perceived cross-language phonetic dissimilarity

Outline

1. Purpose of the Speech Learning Model (SLM)
2. Historical background
3. Core aspects of the SLM
4. Some predictions generated by the SLM
5. How to falsify the SLM
6. **What is most needed now?**

What is needed most now?

1. More adequate methods of participant selection
2. Standardized measure of perceived L1-L2 phonetic distance
3. More precise measures of L2 input
4. Quantitative methods to model the relation between L2 input and L2 performance
5. Explanations for the large individual differences seen in many study of L2 learning, especially among late learners

What is needed most now?

6. The development of a standardized test of category formation using non-overt responses such as MMN or neural maging)
7. Large scale studies examining multiple measures of production & perception specifically designed to falsify the SLM
8. Patience
9. Hard work
10. Imagination
11. Luck

Thank you

- I thank NIH for it's continued support for this research over the past 20 years.
- Thanks are due my many collaborators, including: K. Aoyama: W. Baker: O-S. Bohn: W. Eefting; S. Fletcher; R. Fox; E. Frieda; S. Guion; R. Hammond; J. Hillenbrand; S. Imai; S. Liu; M. Mack; V. Mann; R. McAllister; D. Meador; M. Munro; T. Nozawa; T. Piske; T. Riney; C. Schirru; A. Schmidt; H. Southwood; Y. Takagi; P. Trofimovich; K. Tsukada; A. Walley; C. Wang; R. Wayland.
- Special thanks are due to **Ian R. A. MacKay, Robert F. Port,** and Grace Yeni-Komshian

References

- Baker, W., Trofimovich, P., Mack, M., & Flege, J. (2002) The effect of perceived phonetic similarity on non-native sound learning by children and adults. In B. Skarabela, S. Fish & A. Do (eds.) *Proceedings of the 26th Annual Boston Univ. Conference on Lang. Development,* pp. 36-47. Somerville, MA: Cascadilla Press.
- Bongaerts, T. (1999) Ultimate attainment in L2 pronunciation: The case of very advanced late learners. In D. Birdsong (Ed). Second Language Acquisition and the critical period hypothesis. Mahwah NJ: Lawrence Erlbaum, pp. 133-160.
  Bongaerts, T., Mennen, S. & van der Slik, F. (2000) Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia Linguistica* 54, 298-308.
- Bongaerts, T., Planken, B. & Schils, E. (1995) Can late learners attain a native accent in a foreign langauge? A test of the critical period hypothesis. In D. Singleton & Z. Lengyel (eds). *The Age Factor in Second Language Acquisiton*. Clevedon, UK: Multilingual Matters, pp. 30-50.
- Bongaerts, T., Summeren, C., Planken, B. & Schils, E. (1997) Age and ultimate attainment in the pronunciation of a foreign language. Studies in Sec. Lang. Acquis. 19, 447-466.
  Delattre, P. (1964) Comparing the vocalic features of English, German and French. *Int. Rev. Appl. Ling.* 2, 71-97.
- Flege, J. (1987) The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification. *J. Phonetics* 15, 47-65.

•Flege, J., Birdsong, D., Bialystok, E., Mack, M., Sung, H. & Tsukada, K. (in press) Degree of foreign accent in English sentences produced by Korean children and adults. *J. Phonetics*.
•Flege, J., Frieda, A. & Nozawa, T. (1997) Amount of native-language (L1) use affects the pronunciation of an L2. *J Phonetics* 25, 169-186.
•Flege, J. & Hammond, R. (1982) Non-distinctive phonetic differences between language varieties. Studies in Second Lang. Acquis. 5, 1-17.
•Flege, J. & MacKay, I. (2004) Perceiving vowels in a second language. *Studies in Sec. Lang. Acquis. 2*6, 1-34.
•Flege, J., MacKay, I. & Meador, D. (1999) Native Italian speakers' production and perception of English vowels. *J. Acoust.I Soc. Amer. 106,* 2973-2987.
•Flege, J. & Port, R. (1981) Cross-language phonetic interference: Arabic to English. *Lang. & Speech* 24, 25-146.
•Flege, J., Schirru, C. & MacKay, I. (2003) Interaction between the native and second language phonetic subsystems. *Speech Comm., 40,* 467-491.
•Flege, J. & Schmidt, A. (1995) Native speakers of Spanish show rate-dependent processing of English stop consonants. *Phonetica* 52, 90-111.

•Flege, J., Schmidt, A. & Wharton, G. (1996) Age affects rate-dependent processing of stops of stops in a second language. *Phonetica*, 53, 143-161.
•Hillenbrand, J., Getty, L., Clark, M. & Wheeler, K. (1995). Acoustic characteristics of American English vowels. J. Acoustical Soc. Amer. 97, 3099-3111
•Lambacher, S., Martens, W., Kakehi, K., Marasinghe, C. & Molholt, G. (2005) The effects of identification training on the identification and production of American English vowels by native Japanese speakers. Appl. Psycholing. 26, 2276-248.
•Lambacher, S., Martens, W., Kakehi, K., Marasinghe, C. & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native Japanese speakers. Appl. Psycholing. 26, 2276-248.
•Lenneberg (1969) Biological Foundations of Language. New York: Wiley.
McAllister, R., Flege, J., & Piske, T. (2002) The influence of the L1 on the acquisition of Swedish vowel quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, 30, 229-258.
•Munro, M., Flege, J. & MacKay, I. (1996) The effect of age of second-language learning on the production of English vowels. *Appl. Psycholing.* 17, 313-334.
•Pisoni, D. & Tash, J. (1974) Reaction times to comparisons within and across phonetic categories. *Percep. & Psychophys.* 15, 285-290.
•Schmidt, A. & Flege, J. (1995) Effects of speaking rate changes on native and non-native production. *Phonetica*, 52, 41-54