Cross-linguistic perception of velar and alveolar obstruents: A perceptual and psychoacoustic study

Tim Arbisi-Kelm[1,2], Jan Edwards[2], Benjamin Munson[3], and Eun Jong Kong[2]

[1]Department of Communication Sciences and Disorders, Augustana College; [2]Department of Communicative Disorders, University of Wisconsin, Madison, [3]Department of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis

Introduction and Rationale

- Velar stop consonants coarticulate more with the following vowel than do stops at other places of articulation.
- This coarticulation may explain why velar stops tend to be acquired relatively later than stops at other places of articulation across many languages.
- The fine phonetic detail of this coarticulation is highly language specific (Arbisi-Kelm et al., 2008). For example:
  - Velar stops before front vowels are more front in Greek than in English.
  - Velar stops before back vowels are more back in Greek than in English.
- Anecdotally, this is reported to influence perception.
  - Greek speakers hear very front /ki/ as [ki] while English speakers hear it as [ti].
- Two purposes to this study:
  - Investigate systematically how these cross-linguistic differences in production influence naïve speaker perception of place of articulation.
  - Relate these cross-linguistic differences in perception to differences in production, as quantified by psychoacoustic properties of the stop bursts.

Method: Perception Experiment

- Participants:
  - 20 naïve adult speakers of American English (tested in Minneapolis)
  - 20 naïve adult speakers of Greek (tested in Thessaloniki, Greece)
- Stimuli:
  - CV sequences produced by 2- to 5-year-old English- and Greek-speaking children and adults participating in the παιδολογος project
  - English stimuli were /d/- and /g/-initial sequences.
  - Greek stimuli /t/- and /k/-initial sequences.
  - These stops are voiceless and unaspirated in both languages.
  - There were six categories, based on native speaker transcriptions.
- Procedure:
  - VAS procedure was used: see Figure 1.

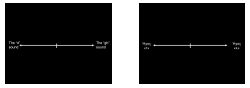

Figure 1
VAS rating display for English (far left) and for Greek (immediate left).

ASA conference, April 2010, 4aSC22

Results: VAS ratings as a function of transcription category

- English stimuli:
  - For both listener groups: VAS ratings for all of the transcription categories were significantly different from each other except for correct /d/ and [d] for /g/ substitutions.
  - English listeners perceived English productions as more velar-like than did Greek listeners.
- Greek stimuli:
  - English listeners perceived only two distinct transcription categories: velar and "not-velar".
  - Greek listeners perceived three transcription categories: velar, intermediate, and alveolar.
  - Greek listeners perceived Greek productions as more velar-like than did English listeners.

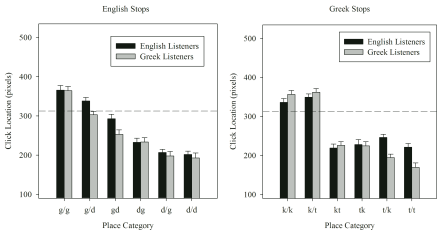

Figure 2
Average VAS ratings for English stimuli (left) and Greek stimuli (right) for the six transcription categories for the two listener groups.

Psychoacoustic Analysis

- We generated spectral slices across a 10-ms Hamming window, centered at the burst, to obtain a frequency distribution of the burst energy.
- We calculated specific loudness (SL) against equal rectangular bandwidth (ERB), as in Moore et al. (1997).
- We used two measures, peak ERB and compactness index.
- In previous work with adult productions across languages (Arbisi-Kelm et al. 2008):
  - Peak ERB differentiated velar and alveolar stop consonants in back vowel contexts
  - Compactness index differentiated velar and alveolar stop consonants in front vowel contexts.

More information on the παιδολογος project, as well as a downloadable version of this poster is available at:
http://www.ling.ohio-state.edu/~edwards

Acknowledgements
This research was supported by T32 grant DC005359 to Tim Arbisi-Kelm, NIH grant R01 DC02932 and NSF grant BCS 0729140 to Jan Edwards, and NSF grant BCS 0729277 to Benjamin Munson.

Psychoacoustic Measures

- Loudest peak (Peak ERB)
  - Represents the point of highest specific loudness (measured in ERB)
  - Higher peak ERB = shorter front cavity (alveolar or front velar)
  - Lower peak ERB = longer front cavity (velar)
- Compactness index (CI)
  - Proportion of overall loudness focused within a 3-ERB band centered at loudest peak.
  - Higher value = compact peak (velar)
  - Lower value = diffuse peak (alveolar)

Analysis: Mixed effects models

- Two mixed effect models were calculated separately for the two stimulus languages: one for front vowel and one for back vowel contexts.
- Dependent variable = proportional pixel location.
- Four independent variables = (1) peak ERB, (2) compactness index, (3)-(4) interaction of each of (1,2) with listener language (Greek or English).

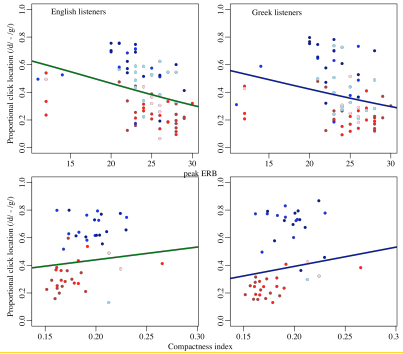

Figure 3
Proportional click locations as a function of peak ERB for English stimuli in back vowel context (top two panels) and of compactness index for Greek stimuli in front vowel context (bottom two panels) for English (left) and Greek (right) listeners. Color coding is by transcription category for alveolar (red) and velar (blue) tokens.

Results

•English stimuli:
- Significant effect of peak ERB only for back vowel context.
- Significant effect of both peak ERB and compactness index for front vowel context.
- No significant interactions with listener language for either vowel context.

•Greek stimuli:
- Significant effect of peak ERB and compactness index for both vowel contexts.
- No significant interactions with listener language for either vowel context.

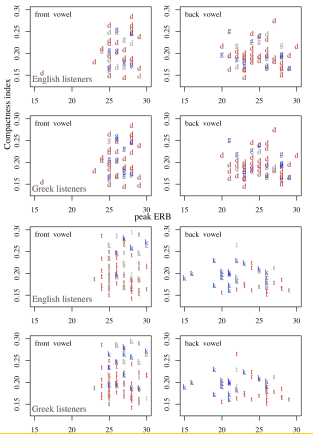

Figure 4
Compactness index plotted against peak ERB for English listeners (top 2 panels of each set) and Greek listeners (bottom 2 panels of each set) in front and back vowel contexts. Top four panels are for English stimuli and bottom four panels are for Greek stimuli. Color coding is by VAS results for alveolar (red) and velar (blue)

Results

- Language difference
  - Only Greek listeners appear to use compactness index in differentiating velar and alveolar stops in front vowel context (especially for Greek stimuli).
  - English listeners have a bias toward alveolar stops in front vowel context.

Discussion

- Peak ERB and compactness index worked well to explain listeners' judgments of place of articulation of alveolar and velar stops, especially for Greek stimuli.
- Previous psychoacoustic analysis of a larger set of English-speaking children's productions found that they did not consistently use compactness index to differentiate place of articulation until age 5.
- The observation that Greek (but not English) listeners rely on compactness index in differentiating place of articulation in front vowel contexts may be related to the fact that velar stops in front vowel contexts are more front in Greek than in English.
- Greek listeners' greater reliance on the compactness index in front vowel contexts may explain the anecdotal observation that Greek listeners perceive very front /ki/ as [ki], while English listeners perceive it as [ti].