Sibilant fricatives in English and Japanese: different in production or perception?

Benjamin Munson[1] · **Fangfang Li (presenter)[2]** · Kiyoko Yoneyama[3] · Kathleen Currie Hall[2] · Mary Beckman[2] · Jan Edwards[4] · Yuki Sunawatari[3]

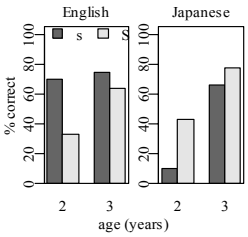
[1]Dept of Speech-Language-Hearing Sciences, University of Minnesota
[2]Dept of Linguistics, Ohio State University
[3]Dept of Foreign Languages, Daito Bunka University, Japan
[4]Dept of Communicative Disorders, University of Wisconsin – Madison

Fricatives and fricative development

- Both English and Japanese have alveolar /s/ contrasting with post-alveolar coronal:
  - English: /ʃ/
  - Japanese: /ɕ/
- Large-scale norming studies show opposite patterns of acquisition, and different errors.
  - English: /s/ is mastered earlier than /ʃ/ and target /ʃ/ is often transcribed as [s]. (Smit et al. 1991)
    e.g. *shoe* *safe* **English: Fronting error**
  - Japanese: /ɕ/ is mastered earlier than /s/ and target /s/ is often transcribed as [ɕ] (Nakanishi et al., 1972). **Japanese: Backing error**
    e.g. *Shukurimu* "cream puff" *semi* "cicada"

The παιδολογος project

- Project goal: to examine cross-linguistic difference in children's phonological development.
- Database: recorded children and adults speech of words with consonant-vowel initial sequences. The languages included are English, Japanese, etc.



- Our native speaker transcription of children's fricative productions for English and Japanese were consistent with the opposite error patterns described in the large norming study.

Research goals

To examine whether the apparent asymmetry in child acquisition is a consequence of…

- Fine differences between languages in production of the contrast;
- Differences in the perceptual criteria that listeners in the two languages use to perceive /s/ and its postalveolar counterpart;
- Or differences in both.

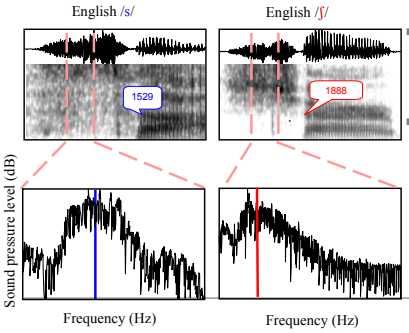
Cross-language subtle differences in fricatives

- For the postalveolar variant, this is less remarkable.
  - English /ʃ/: lamino-palatal alveolar; rounded
  - Japanese /ɕ/: lamino-dorsal alveolopalatal; unrounded

/ʃ/ /ɕ/

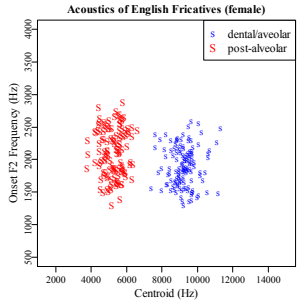
- For the /s/, this difference is rather remarkable, as these represent what is ostensibly the 'same' category.
  - Japanese /s/: laminal alveolar; less sibilant
  - English /s/: apical-alveoar; more sibilant

Production experiment: Acoustic parameters from Li, Beckman & Edwards (in prep.)



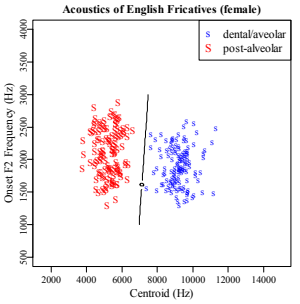
- Centroid frequency: weighted mean frequency (Forrest et al. 1988).
- Onset F2 frequency: second formant frequency taken at the vowel onset following the fricative (Funatsu, 1995; Halle & Stevens 1997).

Acoustics of English /s/ vs. /ʃ/



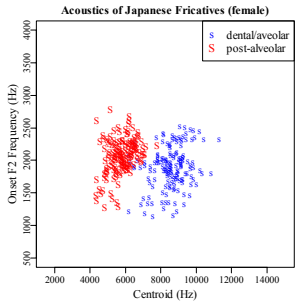
- /s/ and /ʃ/ can be completely separated in the centroid dimension.
- Stepwise multiple logistic regression also confirmed that centroid was the only parameter needed to separate out /s/ from /ʃ/.

Acoustics of English /s/ vs. /ʃ/



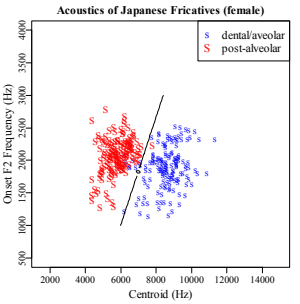
- /s/ and /ʃ/ can be completely separated in the centroid dimension.
- Discriminant prediction plot showed that onset F2 offered very limited predictive power.

Acoustics of Japanese /s/ vs. /ɕ/



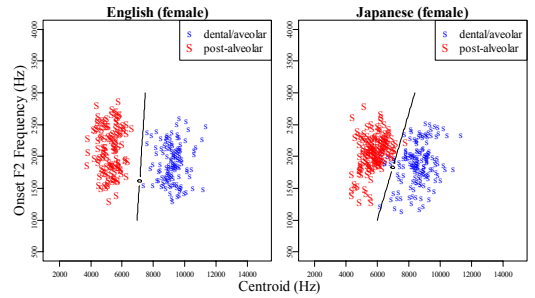
- /s/ and /ɕ/ cannot be completely separated in the centroid dimension.
- Stepwise multiple logistic regression also confirmed that centroid, together with onset F2 Frequency were needed to separate out /s/ and /ɕ/.

Acoustics of Japanese /s/ vs. /ɕ/

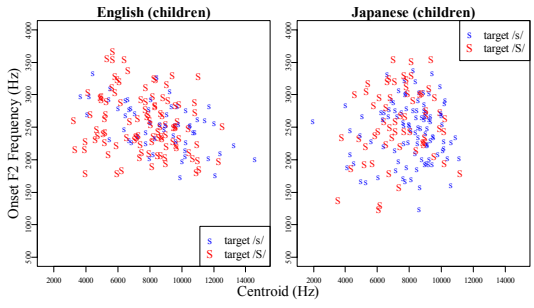


- /s/ and /ʃ/ cannot be completely separated in the centroid dimension.
- Discriminant prediction plot showed that onset F2 offered more predictive power than the English /s-ʃ/ pair.

English vs. Japanese



English vs. Japanese (children)



English vs. Japanese (children)

- Children's productions are very similar across languages, with no good separation between the two target categories and many intermediate tokens.
- However, the similar productions were parsed differently in different languages, with English listeners identifying more /s/ as correct production and the Japanese listeners identifying more /ɕ/ as correct productions.
- The perceived acquisition pattern might reflect a somewhat complex interaction between the children's productions and what the adults accept as a correct production.

Perception experiment

- Research questions:
  - What are the cross-linguistic perceptual differences in perceiving the two fricatives in the two languages?
  - To what extent is the apparent cross-linguistic asymmetry due to differences in perceptual norms?
- Prediction:
  - Given the production differences, we would expect that adult native speakers of English and Japanese would parse the multidimensional acoustic space differently.
- Subjects:
  - 19 English-listeners from Minneapolis, MN, US.
  - 20 Japanese-listeners from Tokyo, Japan.

Perception experiment

- Stimuli:
  - Initial CV in words produced by English- and Japanese-acquiring children (from the παιδολογος project database.)
  - Words selected by excluding stopping errors and other fricative substitution errors with /f/ or /θ/.
- Task:
  - Each listener hears two blocks of the same 400 tokens.
  - They were asked whether each token began with an <s> sound in one block and an <sh> in the other block.
  - They need to answer by pressing "Yes" or "No" button.
  - Naïve listeners didn't know they were listening to multiple languages.

Procedures of calculating group responses

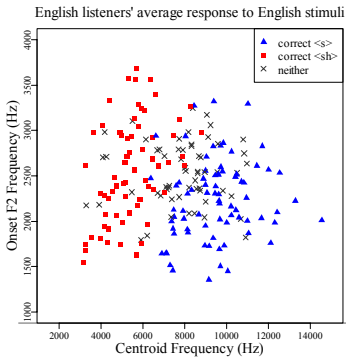
Stimuli	Word	Question	Naïve listeners' responses		Average response	Group acceptability
	"sofa"	Is it <s> ?	Listener 1: "Yes"	1	**0.84**	**1**
			Listener 2: "No"	0		
			Listener 3: "Yes"	1		
			Listener 4: "Yes"	1		
			Listener 5: "Yes"	1		
			Listener 6: "Yes"	1		
			Listener 7: "No"	0		
			….	…		
			Listener 20:"Yes"	1		

bigger than 0.7

= Mean (1,0,1,1,0,1,0…1)
= 0.58

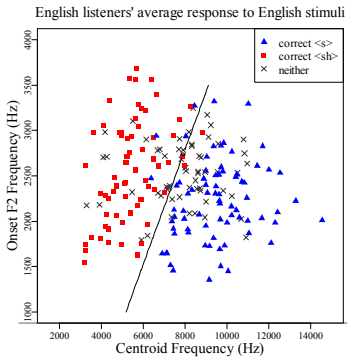
- Criterion of average response to be above chance: "0.7"

Cross-linguistic perception difference



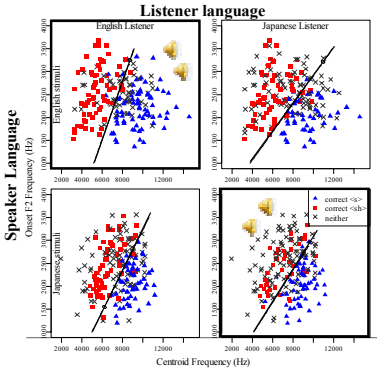
- Correct <s>: tokens with averaged listeners' responses above 0.7 for the question 'Is it <s>?'
- Correct <sh>: average response above 0.7 question 'Is it <sh>?'
- "neither": tokens with averaged listeners' responses below 0.7 both questions.

Cross-linguistic perception difference



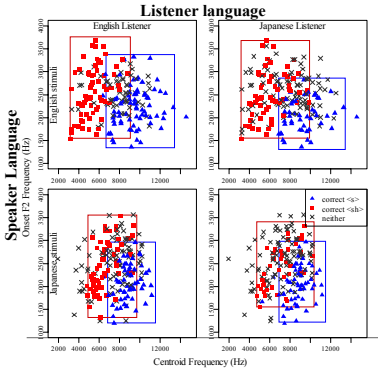
- Discriminant analysis
  - English listeners, similar to their productions, rely centroid frequency more heavily than on the onset F2 freuqency.

Cross-linguistic perception difference



- English: centroid is weighted more than onset F2 frequency.
- Japanese: centroid and onset F2 frequency are nearly equally important.
- The perception of a target fricative in a particular language is constrained by the competing fricatives which share the same space.

Cross-linguistic perception difference



- The English listeners' acceptable range for /s/ is bigger than for /ʃ/, and is bigger than Japanese listeners acceptable range for /s/.
- The difference in Japanese listeners' acceptable range for /s/ and /ɕ/ is the opposite.

Summary

- The analogous fricative categories {s/S} in the two languages differ in both production and perception.
- Perception differences track production differences.
- English listeners have a bigger centroid*onsetF2 range for correct <s> than for correct <sh>.
- Japanese listeners have a bigger centroid*onsetF2 range for correct <sh> than for correct <s>.

Discussion

- Cross-linguistic studies are at a loss to explain order of acquisition asymmetries if all they use are coarse alphabetic transcriptions.
- Categories' phonetic instantiations differ—an /s/ in one language isn't the same as an /s/ in another, even if we confidently transcribe both as /s/.
- The difference in the acceptable centroid range can potentially account for the opposite error patterns documented in acquisition literature in the two languages.

Conclusion

- If we examine fine phonetic detail in production and perception, we get a clearer picture of acquisition.
  - Categories that are acquired early are ones in which there is great permissible variation in the adult language.
  - This greater variation leads adults to accept a wider variety of kids' productions of that sound as correct.
- Future research: Whether the bigger variability in adult norms…
  - Makes it easier for child to hit the target in production.
  - Makes it easier for child to extract the category in perception.

  More work needs to be done to tease apart all three possibilities.

Acknowledgement

Work supported by…

- NIDCD grant 02932 to Jan Edwards.
- McKnight Foundation, through a University of Minnesota Presidential Fellowship to Benjamin Munson.
- OSU Cognitive Science Center 2006 Interdisciplinary Summer Fellowship to Fangfang Li.