B10. On Segmental Factorability in Second Language Learning

Kenneth de Jong and Noah Silbert

Linguistics and Cognitive Science, Indiana University

(kdejong-at-indiana-dot-edu & nosilber-at-indiana-dot-edu; www.indiana.edu/~lsl)

Main Points

→ This paper distinguishes between two models of learning a second language phonological classification system

1) Token-by-token learning starts with individual elements and generalizes to broader featural classification skills later on.

2) Generalized classification starts with featural classification schemes first and fills in specific details later on.

→ Evidence comes in two forms

1) Subjects who are more accurate in distinguishing one pair of segments are also more accurate with other pairs contrasting the same feature.

2) Identification patterns for unfamiliar ('new') segments approximate those expected for featural independence.

→ Tentative Conclusion: listeners create generalized classification schemes first and develop more detailed and robust classification with increased experience

Introduction

Two types of SLA Phonological Models

1) Second language learning is determined by general properties of the two languages
   - models rely on prosodic position, featural contrast, etc., all of which are generalized across segments
   - predominates linguistically oriented literature
2) Second language learning is determined by specific peculiarities of allophonic variants of each segment
   - the basic segmental classification is incorporated into SLM and other models
   - predominates experimental literature

→ Previous work on L2 vowel classification shows little generality across segments (e.g. data in Strange *et al*, 1998)

Two types of Segment Identification Models

1) Segment identification relies on independent classification of a sound on the basis of general features distinguishing it from other sounds
   - Nearey (2003) & Boothroyd & Nittrouer (1988) develop models of perceptual classification of syllables on the basis of (independent) classification of each segment
2) Segment identification relies on map of individual tokens, such that each segment has its own independent signature.

→ Previous work on vowel classification finds little generality in identification patterns (Nearey, 2003)

Current Research Question

Is there evidence for generality in early second language acquirers' identification of consonants in the second language?

OR … Do acquirers simply construct segment-specific identification skills?

Methods

Listeners

• Each listener was a student in their early 20s at Kyonggi University near Seoul, Korea
• None had traveled extensively in English speaking countries
• 41 listeners grouped into 4 blocks

Stimuli

Speakers

• 4 Speakers of American English from the Northern Midwest
• All in late 20s

Corpus

• Obstruent consonants in non-words
• Symmetrically structured inventory in a 2x2x2 matrix

	Coronal		Labial	
	Voiced	Voiceless	Voiced	Voiceless
Stops	**/d/**	**/t/**	**/b/**	**/p/**
Fricatives	**/ð/**	**/θ/**	**/v/**	**/f/**

• The Stops are similar to Korean stops, though the voicing contrast is somewhat different
• The Fricatives are new. Korean has no anterior non-sibilant fricatives
• All consonants paired with vowel /a/
• Consonant in four prosodic environments

	Intervocalic	At Edge
Pre-stress	**/ ∂ `pa/ ('uhpah')**	**/pa/ ('pa')(=initial)**
Post-stress	**/ `a p∂/ ('oppa')**	**/ap/ ('op')(=final)**

•Korean does not have stress, and voicing contrasts are neutralized at a final edge

Procedure

• Speakers cued orthographically
• Each stimulus was produced in isolation

Procedure

• English identification task presented in counterbalanced order with various other tasks for other studies
• Quasi open-set paper response sheets with following options

dog	tell	thin	that	fall	vase	sit	zip	pin	ball	rain	law	hall	wood	yes	
d	t	θ	ð	f	v	s	z	p	b	r	l	h	w	y	other ( )

Analysis 1. Interlistener variation

Rationale

• If each segment has its own identification signature, the development of abilities for each segment need not parallel those for any other segment
• In generalized featural models, all segments which share a featural contrast should develop in parallel

Test

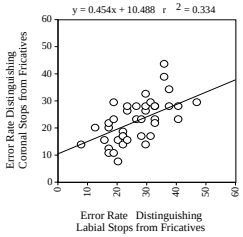
• Regress accuracy in a pair of segments for each listener against their accuracy in other pairs contrasting the same feature
• Featural models predict that listeners who are better at the contrast in one segment will be better at the contrast in other segments, yielding a strong positive correlation

Results 1.

Manner Identification

Coronals vs. Labials

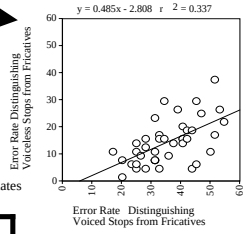
• Listeners tend to be more accurate with labial segments
• Accuracy for coronal fricatives correlates with accuracy for labial fricatives



Next: voicing affects acoustics of manner contrast more directly…

Voiced vs. Voiceless

• Listeners tend to be more accurate with voiced segments
• Accuracy for voiced fricatives correlates with accuracy for voiceless fricatives



Across prosodic position

• Listeners' accuracy in manner contrasts correlates across all four prosodic positions

	Initial	Pre-stress	Post-stress	Final
Initial	1.000	**0.283**	**0.302**	**0.196**
Pre-stress		1.000	**0.384**	**0.241**
Post-stress			1.000	**0.140**
Final				1.000

Not all segmental accuracies correlate, so this is not just due some listeners being better overall than others. Compare the manner results with those for voicing identification:

	Initial	Pre-stress	Post-stress	Final
Initial	1.000	**0.488**	0.074	0.027
Pre-stress		1.000	**0.325**	0.002
Post-stress			1.000	0.034
Final				1.000

• Intervocalic conditions correlate with one another
• Pre-stress conditions correlate with one another
• Voicing contrasts represent different skills depending on the prosodic position
• Korean voicing allophony is conditioned by prosodic position; separate skills likely correspond to allophonic differences

Results 2.
Part-whole Analysis Analytic

Procedures

Rationale

• In generalized featural models, identifying a segment is the product of identifying each feature distinguishing it from others

Test

• Solve for 'J-factor' (Boothroyd & Nittrouer, 1988):
Observed segment accuracy =
(Observed average feature accuracy) raised to **J**th power.
• Current corpus can be neatly described with three binary features, complete independence would yield J = 3.
• Nearey (2003) finds segmental models produce J < 2.

Results Overall

	Initial	Pre-stress	Post-stress	Final	Overall
J-factor	2.627	3.034	2.816	2.750	2.710

• J-factors consistently near 3
• Small reductions may be due to familarity effect (Benki, 2003).
• J's reduced in initial position, which is most similar to Korean

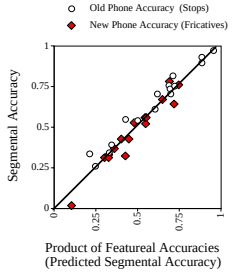
Results Split by Segment

Below we plot average observed accuracy for each segment in each prosodic position against the accuracy predicted as the product of the accuracy of each feature

• Correlation between accuracy and predicted accuracy is close (Stops: $r^2 = 0.969$, Fric.: $r^2 = 0.944$)

• Average ratio of accuracy to predicted accuracy for fricatives = 1.01

• Stops tend to be identified more accurately than predicted (circles lying above the diagonal)



Summary

• Experiment reveals evidence for generalized featural structure in second language consonant identification
  - Cross-subject differences in manner accuracy generalize across consonants and prosodic locations
  - Identification accuracy conforms to independence models with J-factor values close to the number of features being contrasted
• Experiment also provides evidence for specific criterial differences that are less general than features
  - Voicing identification is specific to prosodic location, possibly due to allophonic effects in the native language
  - Segments associated with similar segments in native language exhibit a small reduction of J-factors, suggesting a non-generalized aspect to their identification
• Results suggest a model in which second language learners begin with generalized labeling criteria which get individually tuned with familiarity with individual segments. The native language provides such familiarity, making criteria for 'similar phones' more specific and less general

Acknowledgements

Work supported by the NIDCD (grant #R03 DC04095-01A2) and by the NSF (grant #BCS-9910701). We also express our appreciation to Prof. Mi-Hee Cho for help in collecting the data reported here and to Kyoko Nagao and Hanyong Park for their work on the design and processing of the data presented here.

References Cited

BENKI, J. (2003). Quantitative evaluation of lexical status, word frequency, and neighborhood density as context effects in spoken word recognition. *Journal of the Acoustical Society of America*, **113**: 1689 - 1705.

BOOTHROYD & NITTROUER, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. *Journal of the Acoustical Society of America*, **84**: 101 - 114.

NEAREY, T (2003). A parameter of syllabification for VstopV and relative timing invariance. *Journal of Phonetics*, **16**: 299 – 326

STRANGE, W., R. AKAHANE-YAMADA, R. KUBO, S.A. TRENT, K. NISHI, & JJ. JENKINS (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, 26: 311 – 344.