

# On the categorical nature of coarticulation and other interpolative gestures

Daniel Silverman  
University of Illinois  
dan@uiuc.edu

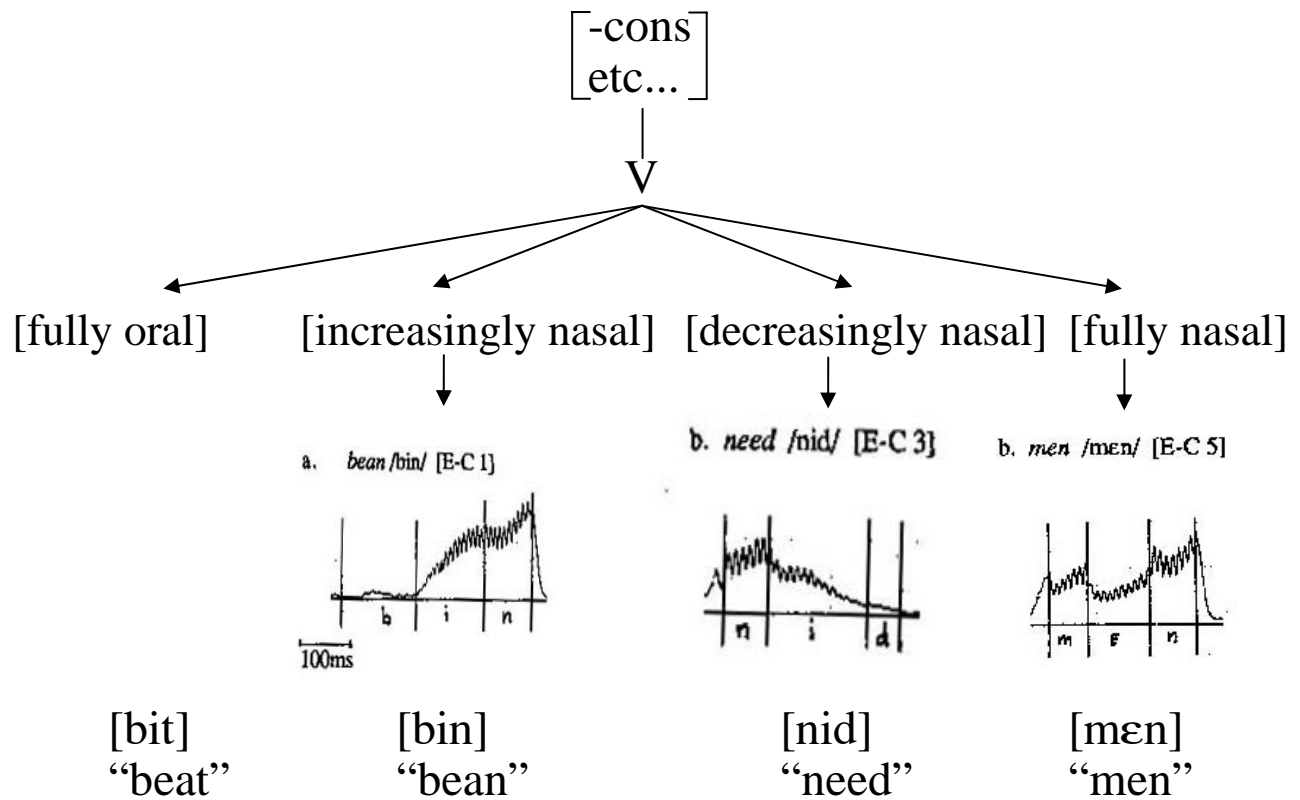
## BACKGROUND

- The categorical nature of an autonomous generative phonology derives from lexical combinations of distinctive features **(a)**.
- Allophonic variation, too, is viewed as a consequence of adding, subtracting or changing the value(s) of category-defining distinctive features from lexical representations **(b)**.
- “Low level,” gradient, interpolative, and coarticulatory effects are viewed as the product of (rules of) “phonetic implementation” (Chomsky and Halle 1968), the rationale being that gradient realizations constitute interpolations between (categorical) phonological targets. These are thus not regarded as part of the categorical phonology; they are often regarded as part of a “generative phonetics” (Pierrehumbert 1980, Cohn 1990, Keating 1990, Zsiga 1993, Kingston and Diehl 1994, etc.) **(c)**.

(c)

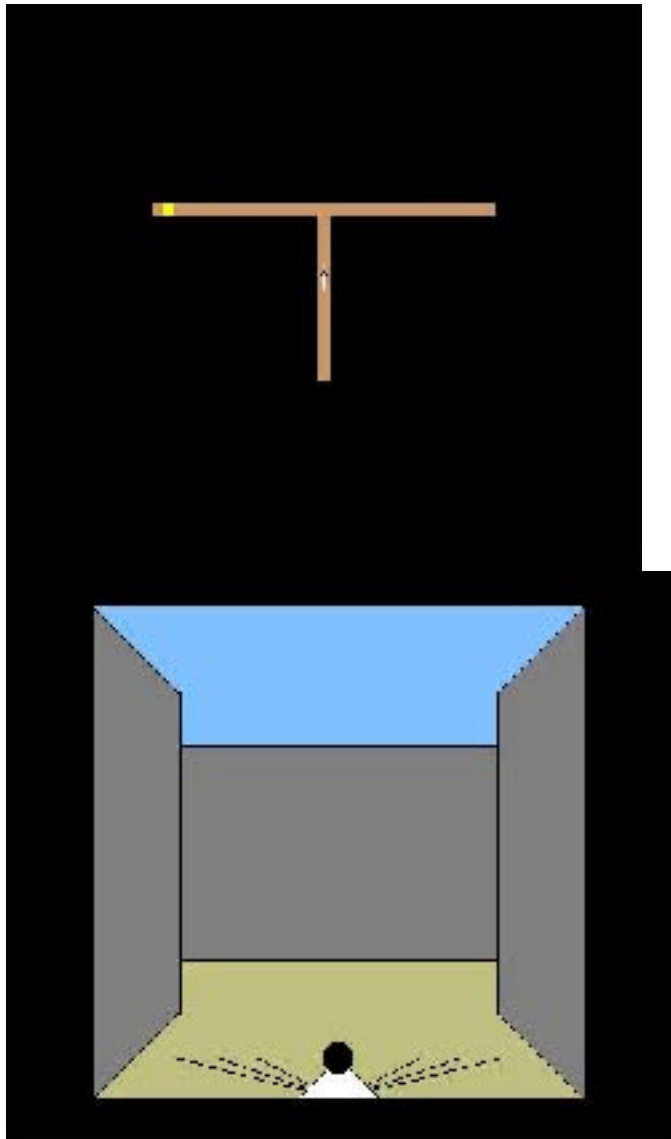


- Example: nasality on English vowels (Cohn 1990):



(nasal traces from Cohn 1990)

- An alternative conception:
  - (a) The categorical nature of phonology derives not from the stability and uniqueness of hypothesized URs, but from the stability of morpheme meaning across allomorphic contexts.
  - (b) Gradient phonetic exponence is not a diagnostic for a phonetics-phonology distinction, but is as much a part of the phonological system as is the segment (or gesture).
- **Probability matching** in lower animals (Gallistel 1990, Labov 1994)



- Gallistel (1990:352ff.): Rats in a T-Maze were rewarded with food 75% of the time at one end, 25% of the time at the other. When provided with feedback, rats matched the probability of reward—running to the one end 75% of the time, the other end 25% of the time—despite the fact that they would receive more rewards if they ran to the one end 100% of the time.
- The “irrationality” of such behavior is only apparent; from a broader evolutionary point of view, in the context of natural, populated settings, the observed behavior is actually beneficial.

- Humans engage in similarly “irrational” behavior in terms of speech production: learners come to largely reproduce the nuances of variation they perceive their elders engaging in, despite the fact that certain of these variants are more successful at keeping contrastive elements distinct.
- For example, young English-learning children initially produce their stops unimodally, with short-lag VOT regardless of category. Through three years of age, a bimodal distribution begins to develop, but still, voicing lead is extremely infrequent, though less so for labials. Even up to six years of age, the lenis category involves fewer tokens with voicing lead than adults’. Finally learners come to match the nuanced variability of their elders (Preston and Yeni-Komshian 1967, Preston, Yeni-Komshian, and Stark 1967, Zlatin and Koenigsknecht 1976), even though fully voiced variants are more distinct from voiceless aspirates than are the more commonly produced devoiced variants.
- Labov (1994:583ff.), “It is not a hypothesis that children do probability matching [during language learning-D.S.]. It is simply a description of the observed facts...”
- Nonetheless, sound changes may slowly progress due to phonetic and/or functional factors, which influence the perception of the speech signal, consequently affecting the variability over which probabilities are matched.
- The facts of probability matching would seem to offer support for this alternative approach to phonological categorization, as learners betray a nuanced mastery of the variability they perceive their elders to engage in, which is undeniably part of their phonological knowledge.

- Given the facts of probability matching, the target/interpolation diagnostic for phonological categorization would seem rather arbitrary; categories change when *meanings* change.

### EXEMPLIFICATION

- Labiality in Trique (Hollenbach 1977). High labial vocalism is always realized on either side of a velar; it does not similarly flank alveolars (labial consonants are extremely rare).

[ <u>nuk</u> wah]	strong	[ <u>run</u> e]	large black beans
[ <u>dug</u> wah]	to twist	[ <u>ut</u> ah]	to anoint
[ <u>ʒug</u> wa]	to be twisted	[ <u>ut</u> fē]	to get wet
[ <u>dug</u> wane]	to bathe (someone)	[ <u>ut</u> i]	to nurse
[ <u>rug</u> wah]	hearth stones	[ <u>ut</u> a]	to gather
[ <u>duk</u> wa]	possessed house	[ <u>dun</u> a]	to leave something
[ <u>zug</u> wi]	(name)	[ <u>gun</u> ah]	to run
[ <u>dug</u> we]	to weep	[ <u>rud</u> aʔa]	stone rolling pin
[ <u>rug</u> wi]	peach	[ <u>ʒut</u> fē]	hens, domestic fowl
[ <u>dug</u> wi]	together with, companion	[ <u>gun</u> i]	to hear

- Longacre (1957):

Proto-Trique:	[uga]	(*[ugwa])	[uda]	(*[udwa])
	[ugi]	(*[ugwi])	[udi]	(*[udwi])
	[uge]	(*[ugwe])	[ude]	(*[udwe])

Modern Trique:	[ugwa]	(*[uga])	[uda]	(*[udwa])
	[ugwi]	(*[ugi])	[udi]	(*[udwi])
	[ugwe]	(*[uge])	[uda]	(*[udwe])

- Trique trans-velar harmony may be historically rooted in the greater likelihood of coarticulation here, since such coarticulation serves to *enhance* the acoustic distinction between the velar and alveolar places of articulation; trans-alveolar harmony cannot be similarly motivated, since superimposed labiality would serve to *diminish* the velar-alveolar acoustic distinction.

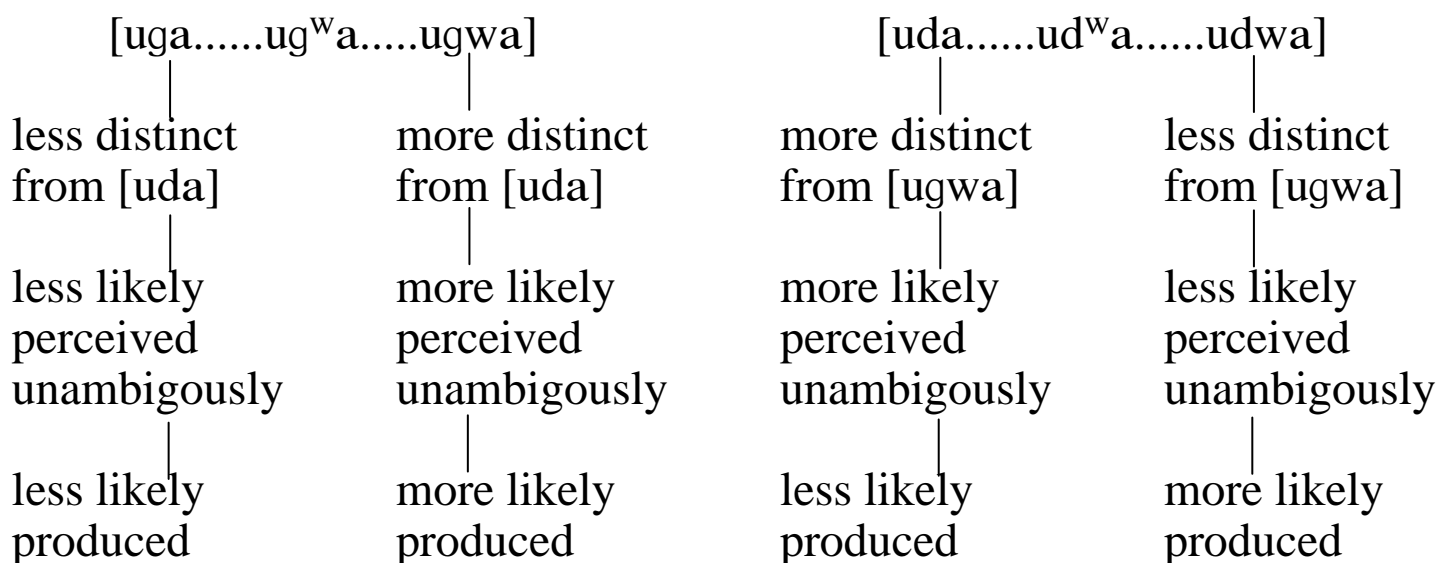
- Articulatorily, labiality spreading is equally likely through both alveolars and velars; but by considering the functional benefit of spreading through velarity, and the counter-functionality of spreading through alveolarity, we might motivate the sound change.

### **FORCES AT WORK**

- There is inherent variability in speech production, thus [uga...ug<sup>w</sup>a...ugwa], [uda...ud<sup>w</sup>a...udwa] are among the possible variants.
- However, [ugwa] variants render the u-velar-V sequences *more distinct* from their u-alveolar-V counterparts; [udwa] variants render u-alveolar-V variants *less distinct* from their u-velar-V counterparts.
- Consequently, listeners are more likely to *perceive* [ugwa] and [uda] unambiguously; hence they are more likely to *produce* [ugwa] and [uda] in their own speech, as a consequence of probability matching.
- That is, the variability engaged in by elders will be largely matched by learners. But nonetheless, due to the greater likelihood of unambiguous perception of certain variants over others, learners's calculated probabilities may differ slightly from their elders's, in that the variants which contrast more sharply with oppositions will more often be perceived correctly, hence, in turn, be more likely produced.



- Conceivable diachronic scenario:



∴ gradual move towards [ugwa]

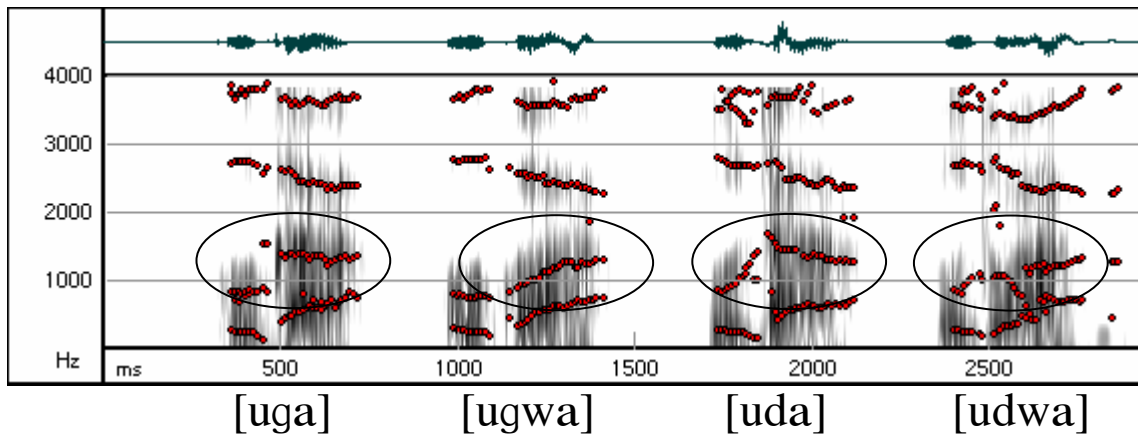
∴ stability of [uda]

## EXPERIMENT

- A laboratory condition may serve to recapitulate elements of the hypothesized historical scenario in “sped up” form, with the introduction of varying amplitudes of white noise into the speech signal, and having listeners report on their perception.

## SUBJECTS/DESIGN

- 10 UIUC graduate linguistics students; all native English speakers.
- Sound files consisting of four relevant phonetic sequences were digitally recorded in the UIUC phonetics lab at a sampling rate of 22,050 Hz: [uga], [ugwa], [uda], [udwa]. Closure durations: [uga]: 40 msec, [ugwa]: 54 msec, [uda]: 50 msec, [udwa]: 51 msec.

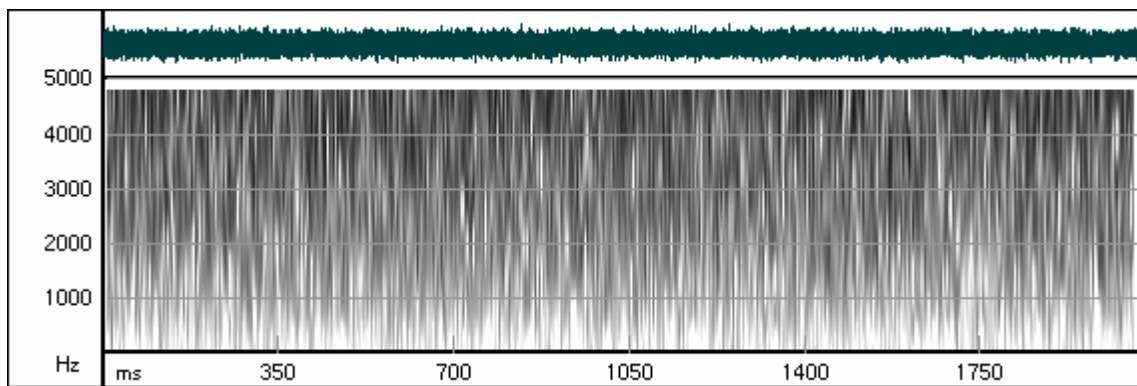


- Note in particular F2 transitions at stop release:

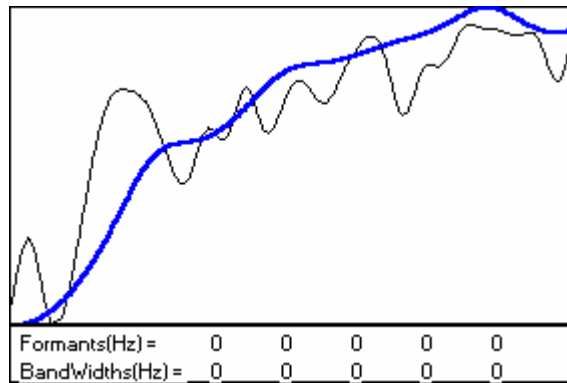
[uga]:	1500 Hz
[ugwa]:	900 Hz
[uda]:	1800 Hz
[udwa]:	1000 Hz

[ugwa] and [uda] are maximally distinct; the presence or absence of labiality renders forms maximally similar.

- Each file was overlaid with four levels of white noise.



Waveform and spectrogram of first level of noise



FFT/LPC at randomly chosen window

- Using PsyScope, subjects listened with headphones to 1000 trials—50 of each of the 20 sound files—in randomly generated blocks of 100, with a 2 second inter-trial interval, and untimed rests between blocks. Using the keyboard, Ss reported which sound sequence they heard ([uga], [ugwa], [uda], or [udwa]). Ss were encouraged to guess if they were undecided.

pre-randomized:

uga	50
uga <sup>1</sup>	50
uga <sup>2</sup>	50
uga <sup>4</sup>	50
uga <sup>8</sup>	50
ugwa	50
ugwa <sup>1</sup>	50
ugwa <sup>2</sup>	50
ugwa <sup>4</sup>	50
ugwa <sup>8</sup>	50
uda	50
uda <sup>1</sup>	50
uda <sup>2</sup>	50
uda <sup>4</sup>	50
uda <sup>8</sup>	50
udwa	50
udwa <sup>1</sup>	50
udwa <sup>2</sup>	50
udwa <sup>4</sup>	50
udwa <sup>8</sup>	50

## RESULTS

perceived → presented ↓	uda	udwa	uga	ugwa	total presented:
uda (1500 total)	1117	40	144	15	1316
udwa (1500 total)	126	813	71	291	1301
uga (1500 total)	355	46	964	43	1408
ugwa (1500 total)	15	501	45	848	1409
total perceived:	1613	1400	1224	1197	5434

raw tally, pooled across middle three noise levels;  
non-responses excluded

- Lowest raw total of errors are shaded:  
[ugwa] → [uda]  
[uda] → [ugwa]
- Highest raw total of errors:  
[ugwa] → [udwa]

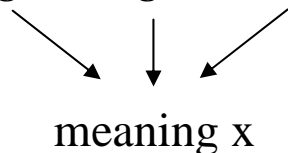
## DISCUSSION

- The ability to psychoacoustically discriminate signals has often been shown to be reflected in the sorts of contrasts that exist, or that are rare or absent, from linguistic sound systems.
- The operative assumption here is that noise introduced into the speech signal might serve to induce a “sped up” rate of misperception in certain contexts, and thus serve to reflect one origin of real-world sound change.
- The variability inherent in speech production may be the fodder for these sorts of sounds changes: the more distinct the variant from an

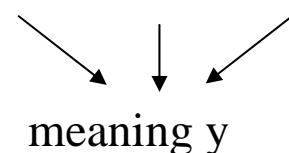
acoustically similar contrastive value, the more likely the system will wend towards this variant.

- Questions arising regarding an autonomous generative phonology? The present scenario is consistent with the hypothesis that *stability of morpheme meaning* rather than supposed *phonological targets* are the decisive factor in determining phonological categories. Whether stable or gradient, phonetic realizations seem to be categorized together as long as meaning remains stable, regardless of phonetic variability.

[uga.....ug<sup>w</sup>a.....ugwa]



[uda.....ud<sup>w</sup>a.....udwa]



- The facts of probability matching may be viewed as supporting this approach, as learners betray a nuanced mastery of the variability they perceive their elders to engage in, which is undeniably part of their phonological knowledge. Indeed, the facts of probability matching are consistent with the hypothesis that categorical phonological (phonetic) targets do not exist. Rather, the target of phonological acquisition seems to be variability itself.
- These experimental findings do not bear directly on the issues of phonological categorization or probability matching (no meanings were associated with the sound sequences, etc.). Nonetheless the findings may be seen as consistent with the sorts of diachronic scenarios that are likely, given the facts of probability matching.

**Thanks to Margaret Russell, Kyle Chambers, and Molly Mack, for their technical, methodological and statistical assistance. Thanks also to “my loyal subjects,” whose dirty looks I must now endure, since this was such an unpleasant experiment to participate in.**

- Chomsky, N, and M. Halle (1968) *The sound pattern of English*. New York, Harper and Row.
- Cohen J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 196:37-46.
- Cohn, A. (1990) *Phonetic and phonological rules of nasalization*. PhD dissertation, UCLA.
- Gallistel, R. (1990) *The organization of learning*. Cambridge MA, MIT Press.
- Hollenbach, B.E. (1977) Phonetic versus phonemic correspondence in two Trique dialects. In W.R. Merrifield, ed., *Studies in Otomanguean phonology*. SIL.
- Kingston, J., and R.L. Diehl, (1994) “Phonetic knowledge,” *Language*, 70:419-454.
- Labov W. (1994) *Principles of linguistic change, internal factors*. Oxford, Blackwell.
- Longacre, R.E. (1957) Proto-Mixtecan. *IJAL* 23.4.III.
- Pierrehumbert, J.B. (1980) *The phonology and phonetics of English intonation*. PhD Dissertation, MIT.
- Preston, M.S., and G. Yeni-Komshian (1967) “Studies on the development of stop consonants produced during the second year of life,” *Annual Report, Neurocommunications Laboratory, Johns Hopkins University School of Medicine* 3:211-222.
- Preston, M.S., G. Yeni-Komshian, and R.E. Stark (1967) “Voicing in initial stop consonants produced by children in the prelinguistic period from different language communities,” *Annual Report, Neurocommunications Laboratory, Johns Hopkins University School of Medicine* 2:305-323.
- Zlatin, M.A., and R.A. Koenigsknecht (1976) “Development of the voicing contrast: a comparison of voice onset time in stop perception and production,” *Journal of Speech and Hearing Research* 18:93-111.
- Zsiga, E.C. (1993) *Features, gestures, and the temporal aspects of phonological organization*. PhD dissertation, Yale University.