MCWOP7
University of Iowa
10AM Sunday, October 10, 2001

On the categorical nature of coarticulation and other interpolative gestures

Daniel Silverman
University of Illinois
dan@uiuc.edu

INTRODUCTORY PROPOSALS

1. Coarticulation is fully relevant to phonology.
2. In connection with variability of production, coarticulation can influence the direction of sound changes.
3. The target of language acquisition is the variability itself.
4. Psychoacoustic experimental findings may be viewed as supporting these hypotheses.

BACKGROUND

5. The categorical nature of an autonomous generative phonology derives from lexical combinations of distinctive features.
6. Allophonic variation, too, is viewed as a consequence of adding, subtracting or changing the value(s) of category-defining distinctive features from lexical representations.
7. "Low level," gradient, interpolative, and coarticulatory effects are viewed as the product of (rules of) "phonetic transcription" (Chomsky and Halle 1968), the rationale being that gradient realizations constitute interpolations between (categorical) phonological targets. These are 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.).

8. An alternative conception:
   - 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. (This is a *non-autonomous* phonology.)
   - 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 theoretical construct "segment" (or "gesture").

9. Probability matching in lower animals (Gallistel 1990, Labov 1994): 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.

10. 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.

11. 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.

12. 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.

13. 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…"

14. 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.

15. 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.

16. Given the facts of probability matching, the target/interpolation diagnostic for phonological categorization would seem rather arbitrary: **phonological categories change when semantic categories change**.

EXEMPLIFICATION

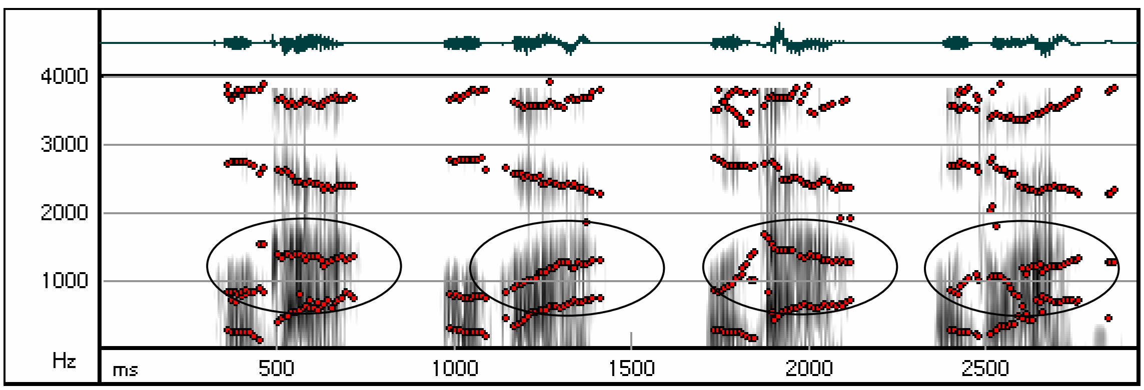
17. Labiality in Trique (Hollenbach 1977). When high labial vocalism precedes a velar, it always follows the velar as well, in the form of labialization; it does not similarly flank alveolars (labial consonants are extremely rare).

[nukwah]	strong	[rune]	large black beans
[duɡwah]	to twist	[utah]	to anoint
[ʒuɡwa]	to be twisted	[utʃe]	to get wet
[duɡwane]	to bathe (someone)	[utʃi]	to nurse
[ruɡwah]	hearth stones	[uta]	to gather
[dukwa]	possessed house	[duna]	to leave something
[zuɡwi]	(name)	[ɡunah]	to run
[duɡwe]	to weep	[rudaʔa]	stone rolling pin
[ruɡwi]	peach	[ʒutʃe]	hens, domestic fowl
[duɡwi]	together with, companion	[ɡunɨ]	to hear

18. 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])

19. 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.



[uga] [ugwa] [uda] [udwa]

20. Note in particular F2 loci/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.

21. 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

22. There is inherent variability in speech production, thus [uga...ug^{w}a...ugwa], [uda...ud^{w}a...udwa] are among the possible variants.

23. 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.

24. 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.

25. Ambiguous tokens will sometimes be impossible to categorize, and hence will not be added to the pool of tokens over which probabilities are calculated.

26. 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.

27. Conceivable diachronic scenario:

<table>
<tr><td colspan="2">[uga…...…ug^{w}a…….ugwa]</td><td></td><td colspan="2">[uda……….....ud^{w}a……......udwa]</td></tr>
<tr><td colspan="2">↓ ↓</td><td></td><td colspan="2">↓ ↓</td></tr>
<tr><td>less distinct from [uda]</td><td>more distinct from [uda]</td><td></td><td>more distinct from [ugwa]</td><td>less distinct from [ugwa]</td></tr>
<tr><td>↓</td><td>↓</td><td></td><td>↓</td><td>↓</td></tr>
<tr><td>less likely perceived unambigously</td><td>more likely perceived unambigously</td><td></td><td>more likely perceived unambigously</td><td>less likely perceived unambigously</td></tr>
<tr><td>↓</td><td>↓</td><td></td><td>↓</td><td>↓</td></tr>
<tr><td>less likely prodcued</td><td>more likely produced</td><td></td><td>more likely produced</td><td>less likely prodcued</td></tr>
<tr><td colspan="2">∴ gradual move towards [ugwa]</td><td></td><td colspan="2">∴ stability of [uda]</td></tr>
</table>

EXPERIMENT

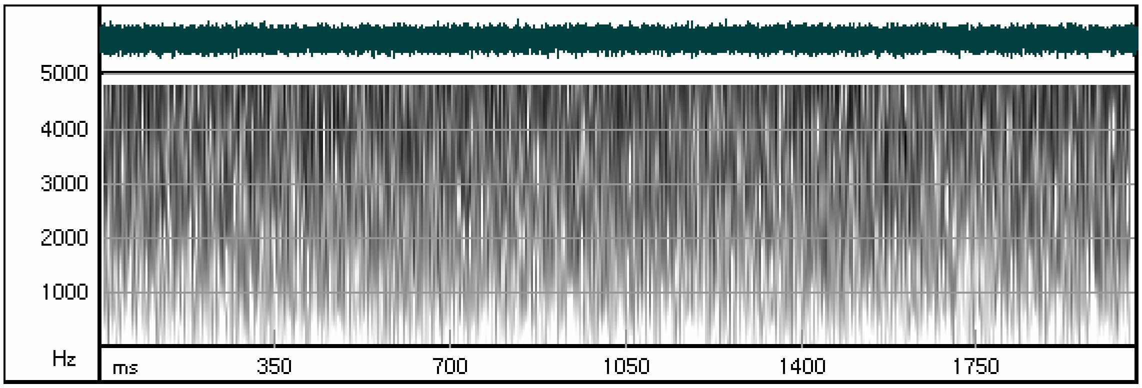
28. A laboratory condition may serve to recapitulate elements of the hypothesized historical scenario in "sped up" form, with the introduction of varying amplitudes of white noise into the speech signal, and having listeners report on their perception.

SUBJECTS/DESIGN

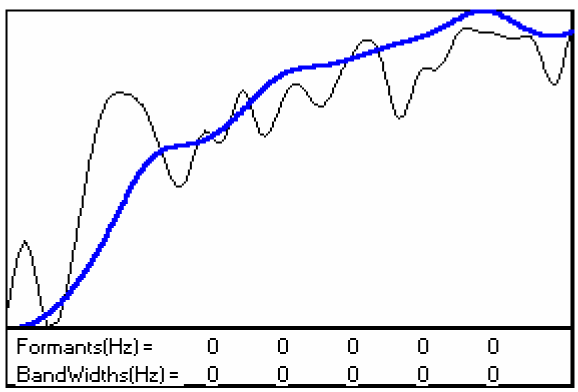
29. 10 University of Illinois graduate linguistics students; all native English speakers; Double-blind.

30. Sound files consisting of four relevant phonetic sequences were digitally recorded in the Department of Linguistics' 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.

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



Waveform and spectrogram of first level of noise



FFT/LPC at randomly chosen window during noise

32. 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.

33. pre-randomized (superscripts = relative noise levels):

uga	50
uga^1	50
uga^2	50
uga^4	50
uga^8	50
ugwa	50
$ugwa^1$	50
$ugwa^2$	50
$ugwa^4$	50
$ugwa^8$	50
uda	50
uda^1	50
uda^2	50
uda^4	50
uda^8	50
udwa	50
$udwa^1$	50
$udwa^2$	50
$udwa^4$	50
$udwa^8$	50

RESULTS

34.

perceived → presented↓	uda	udwa	uga	ugwa	total perceived:
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 (statistics not yet completed)

35. Lowest raw total of errors are shaded:

[ugwa]	→	[uda]	(15)
[uda]	→	[ugwa]	(15)

36. Highest raw total of errors:

[ugwa]	→	[udwa]	(501)
[uga]	→	[uga]	(355)
[udwa]	→	[ugwa]	(291)
[uda]	→	[uga]	(144)

DISCUSSION

37. 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.

38. 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.

39. 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.

40. 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 (virtually never) or gradient (virtually always), phonetic realizations seem to be categorized together as long as meaning remains stable, regardless of phonetic variability.

41. 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.

42. 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.

43. The exquisite articulatory control that speakers display in their productions is best evidenced by the fact that they are able to largely match the variability present in the ambient pattern. The probability matching itself, no doubt, betrays an extremely sophisticated statistical

analytic ability on the part of learners. Moreover, their actual productions betray evidence that they are able to implement their calculated probabilities in their own speech.

44. On this view, learners's articulatory talents are harnessed in service to *copying*, not *modifying* (improving upon or otherwise) the ambient speech pattern.

45. 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.