MOT 2003
University of Toronto
February 7-9 2003

How to explain natural classes without universal distinctive features

Jeff Mielke[1]
The Ohio State University
mielke@ling.ohio-state.edu

1. Introduction

(1) This paper claims that universal distinctive features are insufficient and unnecessary to adequately describe and explain the crosslinguistic similarity of phonological systems (natural classes, alternations, etc.).

- By "universal distinctive features", I mean innate cognitive entities which are part of Universal Grammar and provide explanation for phonological patterns.
- I am *not* arguing against cognitive categories. Phonological features can be thought of as "tags" which identify classes of sounds that are inductively found to be similar (be it with respect to an acoustic or articulatory property or phonological patterning)

(2) There are two reasons for the claim that universal distinctive features are insufficient and unnecessary:

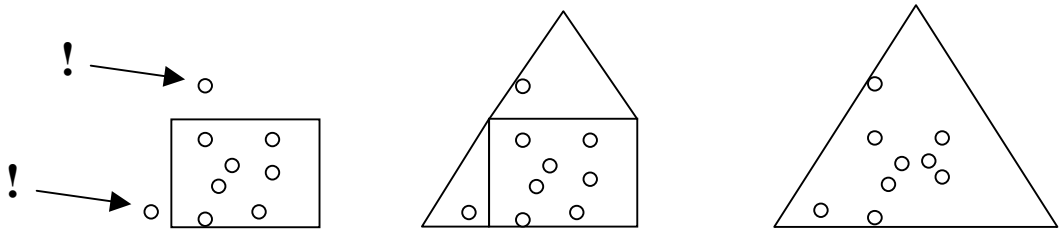
- Universal distinctive features cannot account for certain phonological phenomena. This paper considers the case of so-called "unnatural" classes.
- The same mechanisms which explain these problematic cases can also explain more common phonological processes.

(3) For example, as illustrated in (4), if your favorite theory of universal distinctive features (a box) fails to predict some of your data (4a), you can attribute them to a historical explanation, physiology, etc. (small triangles), as in (4b), or make historical explanation, physiology, etc. your theory of phonology[2] (4c).

[1] This research has benefitted greatly from the attention of my generals exam committee (Beth Hume, Keith Johnson, Brian Joseph, and Dave Odden) and also from discussions with Mike Armstrong, Steve Conley, Robin Dautricourt, Robin Dodsworth, Rich Janda, Grant McGuire, Scott Myers, Anton Rytting, Andrea Sims, Giorgos Tserdanelis, Steve Winters, and members of Phonies.

[2] The use of the term "phonology" refers to patterns of sounds observed in human language, and this is not to be confused with "Cognitive Phonology" (Hale and Reiss 2000).

(4) Theories of phonology



(a) (b) (c)

2. Why universal distinctive features?

(5) Features are widely assumed to be the explanation for phonological processes. E.g.

> "[S]ince features are universal, feature theory explains the fact that all languages draw on a similar, small set of speech properties in constructing their phonological systems… Feature theory… has provided strong confirmation for the view that languages do not vary without limit, but reflect a single general pattern which is rooted in the physical and cognitive capacities of the human species" (Clements and Hume 1995:245).

(6) **Are universal distinctive features *necessary* to explain crosslinguistic similarities?** Much of the similar patterning of speech sounds (in spoken languages) can be explained by uncontroversial facts such as those in (7).

(7) Some uncontroversial facts about language.

a. All spoken languages are produceable by the human vocal tract.
b. All spoken languages are perceivable by the human perceptual/auditory system.
c. All human languages are used for communication.
d. All human languages are learnable by human children.
e. All human languages are subject to human cognitive processes (in learning, memory, etc.)
f. All human languages convey social meaning.
g. All human languages change, often in ways related to (a-f) (and the remnants of these changes can be observed synchronically).

(8) Given these facts, it might not be surprising to find similar patterning among speech sounds in language after language, with or without universal distinctive features. But the claim of cognitively-based universality was accepted with fairly little debate.

(9) **QUESTION: Are distinctive features universal?**

(10) **ANSWER (1930s and 1940s): No.** Following, e.g., Baudouin de Courtenay (see 1972), emphasis is placed on examining the functioning of sounds in each language as an independent system.

- "[I]t is evident that the determination of the phonemic content of a phoneme presupposes its prior classification in the system of distinctive oppositions existing in the language in question" (Trubetzkoy 1939:67).
- "The description of a system of values and the classification of its elements can be made only from that system's own perspective, that is, from the perspective of the tasks that the system fulfils" (Jakobson 1942:241).

(11) **ANSWER (1960s-present): Yes.** Chomsky and Halle (1968:164) emphasize the importance of universal grammar: "[distinctive features] must be determined absolutely, within general linguistic theory, and independently of the grammar of any particular language."

(12) Distinctive features have not always been assumed to be the source of explanation in phonology. One assumption that has remained constant in feature theory is that there is a small (but not necessarily universal) set of distinctive features. The reasons for this claim have varied over time.

(13) **QUESTION: Why is there a small set of distinctive features?**

(14) **ANSWER (1940s): Economy.** Jakobson proposes distinctive features as a means of reducing the number of "unmotivated" contrasts, which strain perception and memory, and even notes (1942:233) that dogs and fish can be trained to distinguish very similar tones if a meaning is associated with each tone, just as a human listener can perceive minute differences between speech sounds if they are contrastive in the his/her language.

(15) **ANSWER (1950s): Because there is no evidence for more features.** In *Preliminaries to Speech Analysis*, Jakobson, Fant and Halle (1952) argue for twelve distinctive features.

- At the time, they did not "detect" any more than twelve features that were necessary to express contrasts of all the worlds languages.
- Information Theory research (e.g. Shannon & Weaver 1949) was influential in the formation of feature theory in the 1950s (particularly the binarity of features).

(16) **ANSWER (1960s and 1970s): Because it's part of Universal Grammar.** Chomsky's theory of syntax (e.g. Chomsky 1965) assumes formal and substantive universals, some of which can be extended to phonology, and notes that "Jakobson's theory of distinctive features can be interpreted as making an assertion about substantive universals with respect to the phonological component of a generative grammar. (*Aspects*, p. 28)

This approach to phonology continues in *SPE* (Chomsky and Halle 1968)

(17) **ANSWER (as early as 1968 but especially in the 1980s and 1990s): Because representations should be explanatory.** Assuming a structured, universal set of distinctive features allowed researchers to place the burden of explanation in these universal representations (e.g., Clements 1985, Sagey 1986, Avery and Rice 1989, McCarthy 1991, Clements and Hume 1995, and many more)

(18) **But are universal distinctive features *able* to explain sound patterns?** Some problems have been noticed over the years.

- **Sign languages** appear to use entirely different phonological features and organization (e.g., Brentari 1998, Eccarius 2002, Miller 2002, Sandler 1989). How might the two sets of innate features have evolved separately?
- Ladefoged (1984) observes that many facts of **language-specific phonetics** are consistent within a given speech community but not explainable from universal principles of phonology or phonetics.
- Port (1996) notes that **incomplete neutralizations** are a problem for a theory of distinctive features which assumes that two phones are either identical or distinct.
- **"Unnatural" classes** (those that cannot be described by any widely-accepted theory of distinctive features) should not exist if universal distinctive features are the source of regularity in phonology.

3. Explaining natural and unnatural classes without universal distinctive features

This section deals with classes of sounds whose patterning together in phonological rules is problematic for a theory based on universal distinctive features. I will start by summarizing what I am proposing in order to explain the data I am about to present.

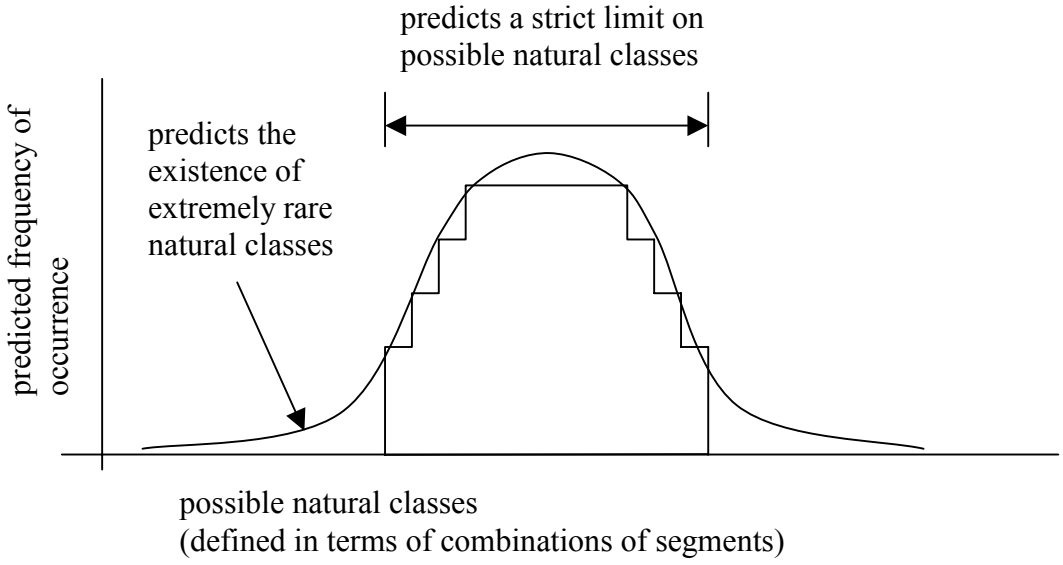
The Proposal:

1. Phonetically natural phonological processes are the direct result of phonetically-conditioned sound change. (e.g. articulatorily-defined natural classes, processes like intervocalic voicing and voiced-obstruent-conditioned tone lowering)
2. Phonetically unnatural phonological processes come from these phonetically natural processes (possibly in many steps), via overgeneralization, analogy, telescoping, restructuring, etc.[3]

(19) This proposal does not require any new assumptions. Historical explanation is already widely used for problematic cases that fall "outside the box" (4).

- **The proposal** predicts that some classes are more likely than others to arise through language change, but none are explicitly ruled out (20, curve).
- **The universal distinctive features hypothesis** predicts that certain classes of sounds (those which are inexpressible with the features of a given theory) are impossible (20, boxy shape).

(20) Predicted natural classes

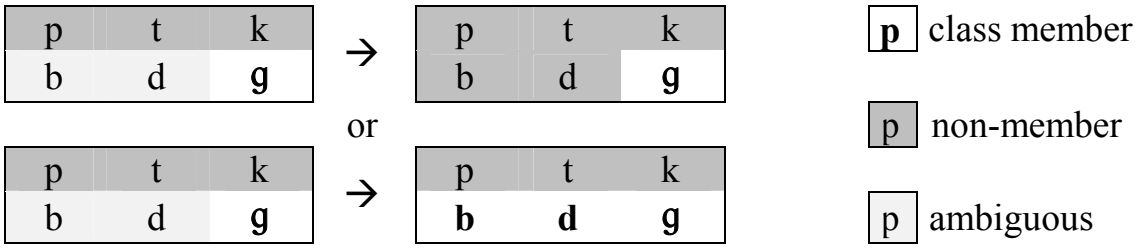


(21) "Natural" processes are the ones which are transparently related to their phonetic roots, and are generally (and perhaps circularly) predicted to be common. "Unnatural" processes are predicted by universal distinctive feature theory not to occur, but are predicted to be rare by the proposal (because they are less likely to arise from common diachronic processes).

[3] For diachronic explanations of "natural" and "unnatural" phenomena, see also, e.g. Anderson (1981), Dolbey and Hansson (1999), Buckley (2000), Hale and Reiss (2000), Hyman (2001), Myers (2002), and Mielke (2002).

(22) **Generalization in linguistics:**

- Generalization is a general cognitive process which is not specific to language, but occurs when a language learner infers a class from available positive evidence.
- For example, given evidence that [g] undergoes a phonological process (say spirantization), that voiceless stops do not, and no evidence either way about [b] or [d], a language learner may infer that this process applies only to [g], or that it applies to all sounds produced with closure voicing.



- Generalization yields the most interesting results when a language learner arrives at the "wrong" generalization (by undergeneralizing or overgeneralizing the prevailing pattern).
- Undergeneralization and overgeneralization are commonly observed in language-learning children (e.g. Pinker 1994 and references cited).
- The "wrong" generalization becomes *right* if it catches on and spreads (social factors become important here).
- Generalization is seen in cases such as the way different speech communities have generalized the use of the English verbal inflectional suffix *–s*:

I digitize.	We digitize.
You digitize.	You digitize.
She digitizes.	They digitize.

Standard American English

I digitizes.	We digitize.
You digitizes.	You digitize.
She digitizes.	They digitize.

Some dialects of Scots English (Tom Stewart, p.c.)

I digitize.	We digitize.
You digitize.	Y'all digitize.
She digitize.	They digitize.

Some dialects of AAVE (Green 1998)

3.1. Sanskrit "ruki" rule

(23) One of the most well-known "unnatural" classes is the set of segments that condition retroflexion of a following /s/ in Sanskrit: /r u k i/ (Whitney 1960, Renou 1961, Zwicky 1970).

(24) Sanskrit nouns with the locative plural ending *-su*

a.	jāsu	jā-	'progeny'
	marutsu	marut-	'wind'
	apsu	ap-	'water'
b.	svasṛṣu	svasṛ-	'sister'
	śatruṣu	śatru-	'enemy'
	vākṣu	vāc-	'voice'
	agniṣu	agni-	'fire'

(25) Sanskrit verbal roots

a.	vas-	'clothe'
	bhās-	'shine'
	tsar-	'creep up on'
	psā-	'devour'
	bharts-	'revile'
b.	dhṛṣ-	'dare'
	bhūṣ	'adorn'
	akṣ	'attain'
	dviṣ	'hate'
	kṣudh-	'crush'

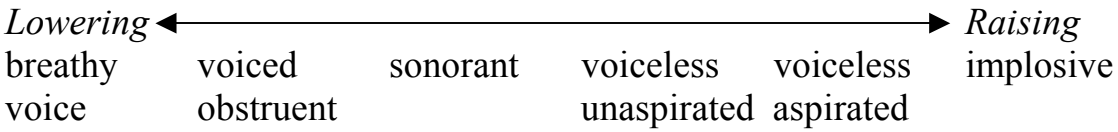
(26) The "ruki" rule does not necessarily require generalization, but it is probably the result of the merging of two rules, one conditioned by /k/ (which has no lexical exceptions) and the other conditioned by /r u i/ (with many exceptions) (Zwicky 1970). Nevertheless, because of this historical change, learning Sanskrit meant learning this rule, and this class of sounds.

3.2. Zina Kotoko

(27) Zina Kotoko exhibits consonant-tone interactions (Odden 2002).

- The most typical phonological consonant-tone interaction is when consonants act as depressor consonants, lowering the tone of adjacant vowels, often from H to L (see e.g., Bradshaw 1999).
- All known cases include voiced obstruents among the class of depressor consonants (Bradshaw 1999). This is consistent with the phonetic observation that voiced obstruents lower the F0 of a following vowel (Hyman and Schuh (1974), in (27))

(28) Hyman and Schuh's (1974) hierarchy of phonetic F0 lowering:



(29) Consonant-tone interactions arise when this phonetic lowering is interpreted as phonological tone (see Bradshaw (1999) for an analysis involving a single feature for voice and low tone).

- Sometimes sonorants also function as depressor consonants (in Nupe, Ngizim, Ewe, and Kanazawa Japanese) (Odden 2002). This is not surprising because of where sonorants are located in Hyman and Schuh's phonetic hierarchy.
- In Zina Kotoko, depressor consonants can include sonorants, glottal stop, and most interestingly, implosives (Odden 2002, to appear).

(30) In the recent past, underlying M (on the initial syllable) is realized as M after [h] and voiceless obstruents.

a.	her-ə́m	'bite'	hərc-ə́m	'slice'
	hwat-ə́m	'inflate'	həl-ə́m	'steal'
b.	skal-ə́m	'pay back'	sap-ə́m	'chase'
	pay-ə́m	'bury'	kah-ə́m	'take a handful'
	ka'd-ə́m	'cross'	sək-ə́m	'send'
	tam-ə́m	'touch'	cənh-ə́m	'be sated'

(31) Underlying M is lowered to L after voiced obstruents, sonorants, glottal stop, and implosives.

a.	ghàg-ə́m	'close'	gàh-ə́m	'pour'
	zə̀gl-ə́m	'carry'	bghwàr-ə́m	'jump pl.'
	gə̀'b-ə́m	'answer'	gə̀'d-ə́m	'open'
	gùlm-ə́m	'twist'	vàlf-ə́m	'give back'
	dùnk-ə́m	'throw'	zàk-ə́m	'beat'
	vìt-ə́m	'blow a fire'	jìk-ə́m	'begin'
b.	yèy-ə́m	'call'	wèh-ə́m	'be tired'
	làb-ə́m	'tell'	rà'd-ə́m	'pull'
	màr-ə́m	'die'	làkf-ə́m	'bring'
c.	'ə̀kf-ə́m	'approach'	'ə̀k-ə́m	'snatch'
d.	'də̀v-ə́m	'put'	'də̀h-ə́m	'write'
	'bàl-ə́m	'dance'	'dàm-ə́m	'eat'

(32) It is surprising that implosives, which have the phonetic effect of raising F0, act as phonological depressors in Zina Kotoko. This is surprising, but clearly not impossible. If "natural" is defined in terms of phonetic transparency or frequency of occurrence, then this is "unnatural", but it clearly occurs in nature.

(33) **Hypothesis:** Originally only voiced obstruents in Zina Kotoko were depressor consonants, and speakers *overgeneralized* this category to include other phonetically voiced sounds such as sonorants and implosives.

p	t	tʃ	k	ʔ
f	s	ʃ	x	h
b	**d**	**dʒ**	**ɡ**	
v	**z**		**ɣ**	
ɓ	ɗ			
m	n			
w	r, l	y		

the phonetic basis

→

p	t	tʃ	k	ʔ
f	s	ʃ	x	h
b	**d**	**dʒ**	**ɡ**	
v	**z**		**ɣ**	
ɓ	**ɗ**			
m	**n**			
w	**r, l**	**y**		

the overgeneralization

(34) What the theory of universal distinctive features explains:

- It explains tone-lowering straightforwardly if it assumes the feature L/Voice (Bradshaw 1999), but cannot explain how this process could also be conditioned by implosives.
- It cannot explain the inclusion of glottal stop.

(35) What the proposal in this paper explains:

- It explains the occurrence of voiced obstruents in all known classes of depressor consonants: They are the phonetic basis because they lower F0 in adjacent vowels.
- It explains how sonorant consonants and implosives can also function as depressor consonants: they share a phonetic property with voiced obstruents (voicing), even thought this is not the property which formed the phonetic basis for the first step.
- It does not really explain the inclusion of glottal stop either, but (correctly) predicts that in most cases it will not be inductively grouped with voiced segments, and crucially does not rule out its inclusion.

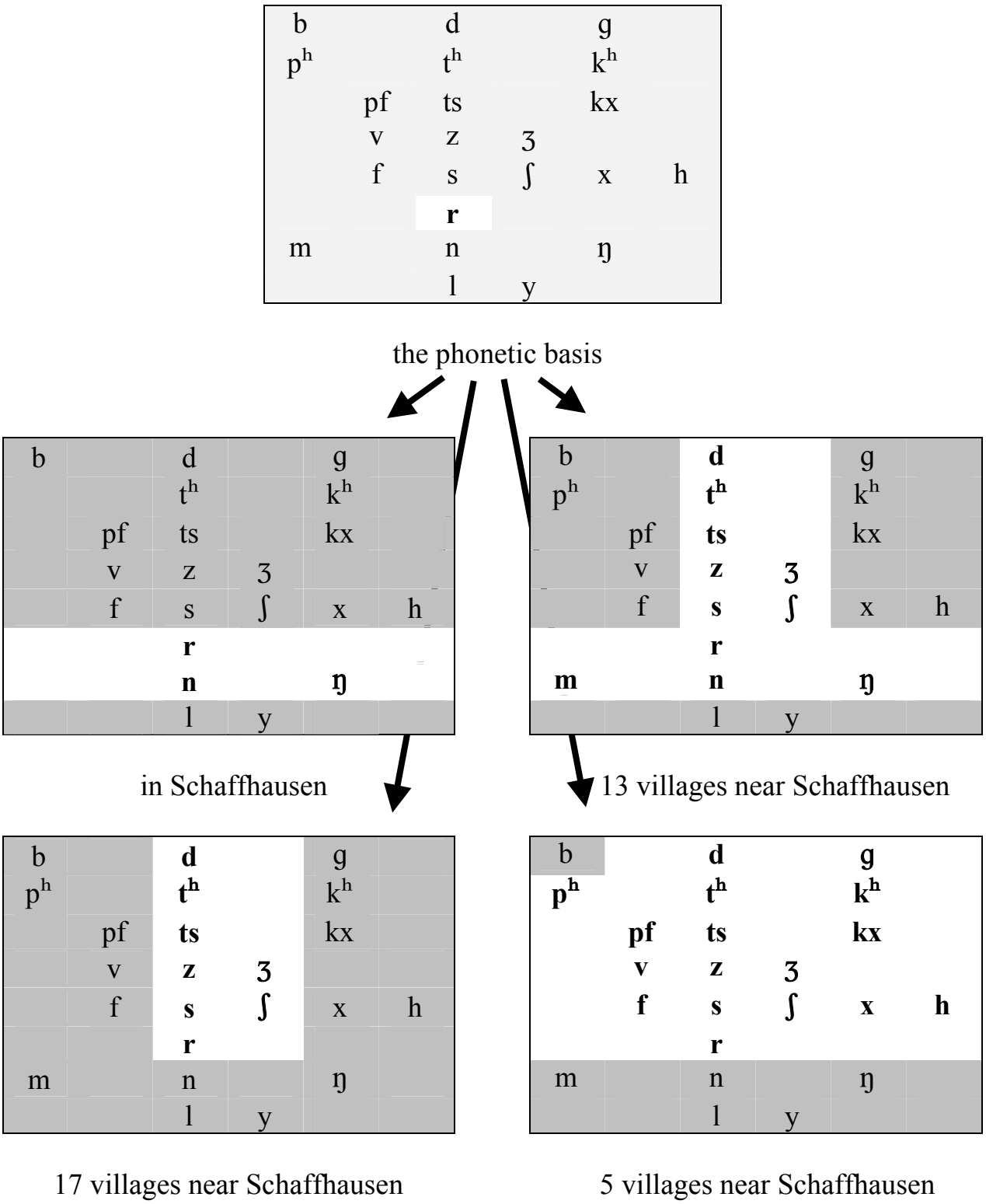
3.3. Swiss German

(36) In northeastern varieties of Swiss German, /o/ is lowered to [ɔ] before certain consonants (Keel 1982, Wanner 1941, Janda and Joseph 2001). In and around the city of Schaffhausen, four different versions of /o/-lowering are observed, in which different classes of sounds condition lowering:

(37) Schaffhausen [o]-lowering generalizations

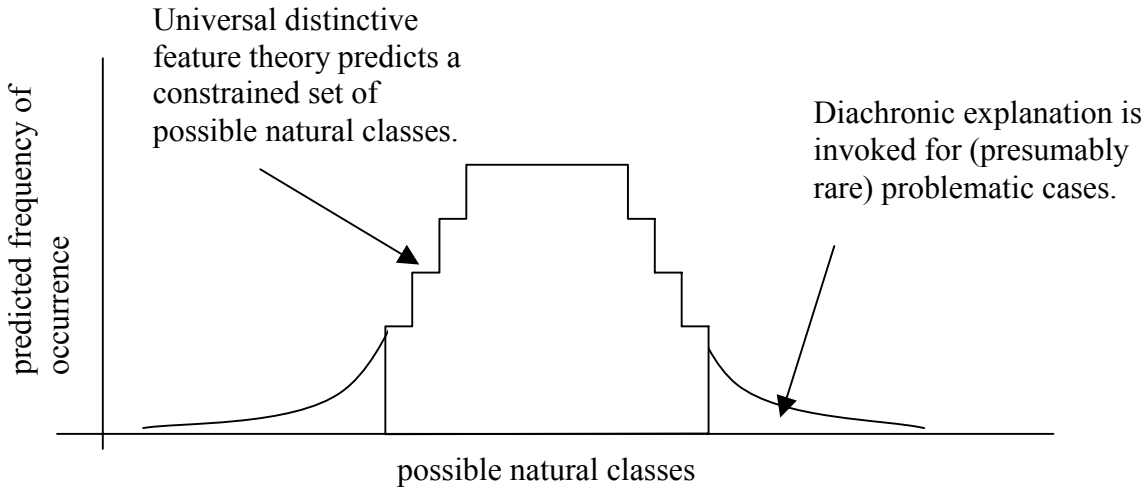
a.	(in Schaffhausen)	/o/ → [ɔ] / __ [r] and nasals
b.	(in 13 villages)	/o/ → [ɔ] / __ [r], nasals, and coronal obstruents
c.	(in 17 villages)	/o/ → [ɔ] / __ [r] and coronal obstruents
d.	(in 5 villages)	/o/ → [ɔ] / __ [r] and all obstruents except [b]

(38) According to Janda and Joseph (2001), /o/ lowering appears to have originally occurred only before /r/. It has been overgeneralized differently in different communities, apparently for sociolinguistic purposes (Janda and Joseph 2001):

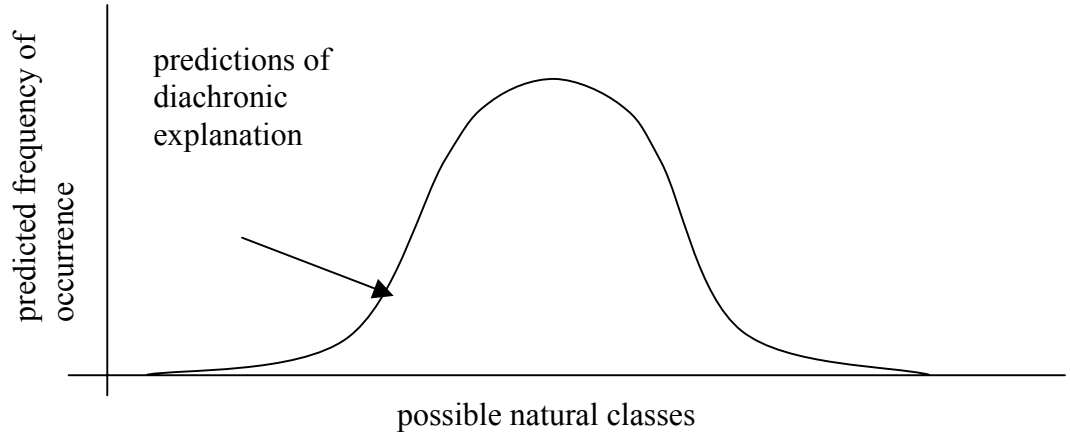


4. Toward an explanatory model

(39) Diachronic explanation is already used to explain problematic cases:



(40) There is no reason to assume that only "unnatural" processes have diachronic roots. "Natural" diachronic processes would likely look a lot like the processes that traditionally don't need to have diachronic explanation invoked for them.

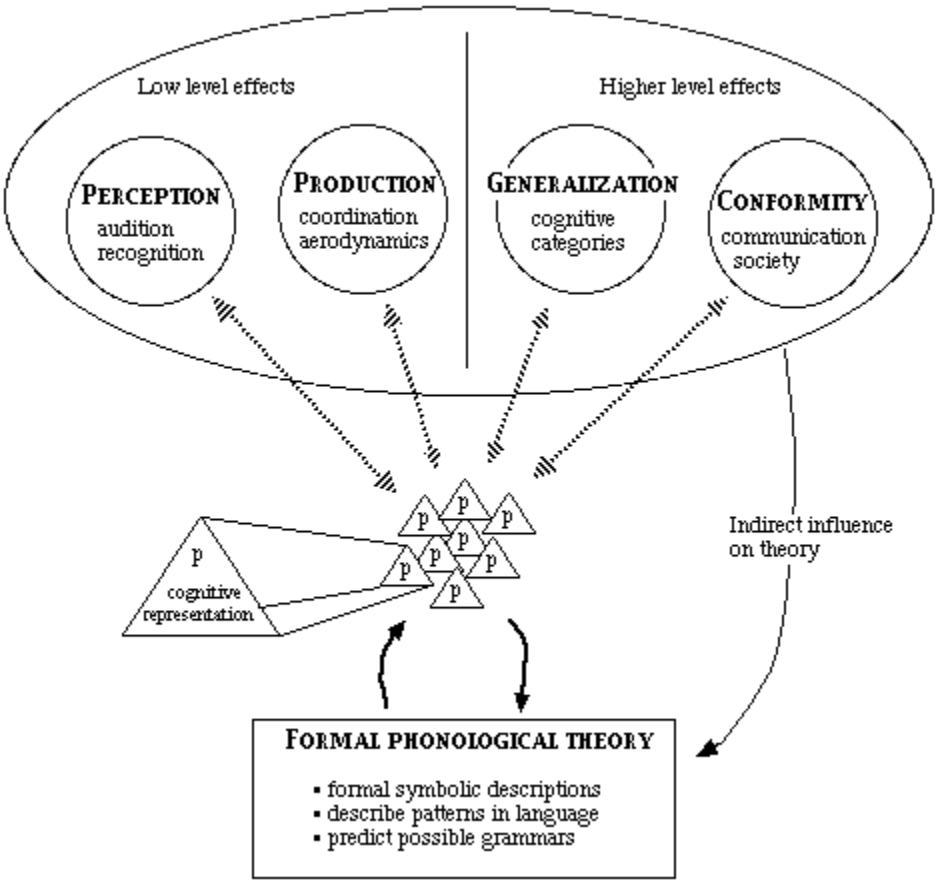


Some research programs are consistent with this proposal. To name a few:

(41) **The "Big Bang" Theory** (Janda and Joseph 2001:2-3):

> "[S]ound change originates in a very 'small', highly localized context over a relatively short temporal span… purely phonetic conditions govern an innovation at this necessarily somewhat brief and limited point of origin… the purely phonetic conditions of [the innovation] are rapidly supplanted during spread… via speakers' imposition of phonological and sociolinguistic conditions, thereby deflecting the future course of the process."

(42) Hume and Johnson (2001) propose **a general model of the interplay of external forces and phonology**, where phonological systems are subject to various "filters" on diachronic sound change which give rise to phonological patterns that are described by formal phonological theory.



(43) Dresher (2001) proposes the **Successive Division Algorithm**, which involves adding features as contrasts are discovered.

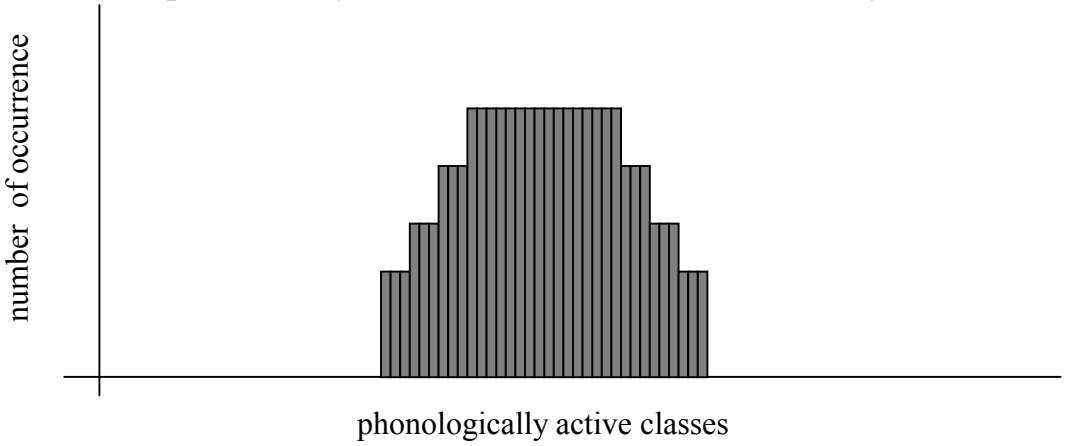
- The language learner starts with the hypothesis that all segments are variants of a single phoneme, and proceeds to cut this phoneme into smaller units as contrasts (such as consonant vs. vowel) are discovered.
- The model does not explicitly assume universality.
- If cuts are inferred from observed phonological observations and phonetic similarity (as above) and the language learner does not have a perfect data set to learn from, then this learning algorithm allows overgeneralization, and combined with the above models, predicts a distribution of natural classes (and alternations) similar to (40).

5. Conclusions

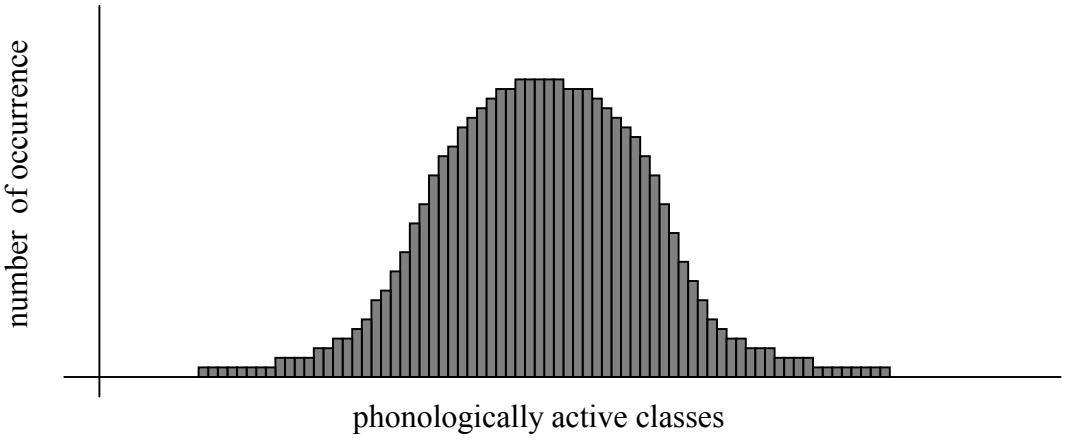
This paper has shown how the existence of "natural" and "unnatural" classes can be explained by one mechanism (without the use of universal distinctive features).

The proposal predicts the existence of *phonologically active classes* which are not predicted by the theory of universal distinctive features. As data, the forthcoming dissertation (Mielke 2004 or so) will include a survey of classes of sounds participating in phonological patterns, drawn from descriptive grammars of a wide variety of languages. The predictions of the two hypothesis are illustrated in (44) and (45).

(44) Classes predicted by universal distinctive feature theory



(45) Classes predicted by the proposal



References

Anderson, S. (1981) Why Phonology Isn't "Natural". *Linguistic Inquiry 12(4)*. 493-539.

Baudouin de Courtenay, Jan Niecislaw (1972) *A Baudouin de Courtenay anthology; the beginnings of structural linguistics*. Edward Stankiewicz (ed. & trans.) Bloomington: Indiana University Press.

Bradshaw, Mary (1999) *A Crosslinguistic Study of Consonant-Tone Interaction.* OSU dissertation.

Brentari, Diane (1998) *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.

Buckley, Eugene (2000) What should phonology explain? Handout from SUNY Buffalo linguistics colloquium.

Chomsky, Noam (1965) *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

Chomsky, Noam, and Morris Halle (1968) *The Sound Pattern of English*. Cambridge, Mass.: MIT Press.

Clements, G.N. (1985) The geometry of phonological features. *Phonology Yearbook 2*. pp. 225-252.

Clements, G.N. and Elizabeth V. Hume (1995) The Internal Organization of Speech Sounds. in Goldsmith, John (ed.) *The Handbook of Phonological Theory*. Cambridge Mass.: Blackwell. pp. 245-306.

Dolbey, A.E. and G.Ó. Hansson (1999) The source of naturalness in synchronic phonology. *CLS 35, Vol. 1*. Chicago: CLS. pp. 59-69.

Dresher, Elan (2001) Contrast and Asymmetries in Inventories. Handout of paper presented at the Asymmetry Conference, UQAM.

Eccarius, Petra (2002) Finding common ground: A theoretical representation of handshape across multiple sign languages. Paper presented at MCWOP 8, Indiana University.

Hale, Mark, and Charles Reiss (2000) Phonology as Cognition. In N. Burton-Roberts, P. Carr, and G. Docherty (eds.) *Phonological Knowledge: Conceptual and Empirical Issues.* Oxford University Press. pp.161-184.

Harris, John (1994) English Sound Structure. Cambridge, Mass.: Blackwell.

Hume, Elizabeth V. and Keith Johnson (2001) A model of the interplay of speech perception and phonology. In Hume, Elizabeth V. and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. New York: Academic Press.

Hume, Elizabeth and David Odden (1996) Reconsidering [consonantal]. *Phonology 13.*

Hyman, Larry M. (2001) The Limits of Phonetic Determinism in Phonology: *NC Revisited. In Hume, Elizabeth V. and Keith Johnson (eds.) *The Role of Speech Perception in Phonology.* New York: Academic Press.

Hyman, L., and R. Schuh (1974) Universals of tone rules: evidence from West Africa. *Linguistic Inquiry 5*: 81-115.

Jakobson, Roman (1942) The Concept of Phoneme. reprinted in Waugh, Linda R., and Monique Moville-Burston (eds.) *On Language* (1990) Cambridge Mass.: Harvard University Press. pp. 218-241.

Jakobson, Roman, C. Gunnar M. Fant, and Moris Halle (1954) *Preliminaries to speech analysis: the distinctive features and their correlates*. Cambridge, Mass.: MIT Press.

Janda, R., and B. Joseph (2001). Reconsidering the Canons of Sound-Change: Towards a Big Bang Theory, ms., OSU.

Keel, William (1982) *Atomic Phonology and Phonological Variation*. Tübingen: Gunter Narr.

Ladefoged, Peter (1984) 'Out of Chaos Comes Order'; Physical, biological, and structural patterns in phonetics. in Van den Broecke, M.P.R. and A. Cohen (eds.) *Proceedings of the ICPhS 10.* Dordrecht: Foris Publications. pp. 83-95.

McCarthy, John. J. (1991) Semitic Gutturals and Distinctive Feature Theory. in Dunlap, Elaine, and Jaye Padgett (eds.) *UMOP 14*. pp. 29-50.

Mielke, Jeff (2002) A time and a place for phonetic explanation in phonology. OSU ms.

Mielke, Jeff (2004 or so) OSU PhD dissertation.

Miller, Christopher (2002) The anatomy of handshape phonology, paper presented at MOT 2002, McGill University.

Myers, Scott (2002) Gaps in factorial typology: The case of voicing in consonant clusters, ms., University of Texas at Austin.

Odden, David (2002) Consonant-Tone Interaction in Zina Kotoko. Paper Presented at MCWOP, Indiana University.

Odden, David (to appear) The verbal tone system of Zina Kotoko. in Schmidt, Odden & Homberg (eds.) *Aspects of Zina Kotoko grammar*, Lincom Europa.

Pinker, Steven (1994) *The Language Instinct*. New York, NY: W. Morrow and Co.

Port, Robert F. (1996) The discreteness of phonetic elements and formal linguistics: response to A. Manaster Ramer. *Journal of Phonetics 24*. pp. 491-511.

Renou, L. (1961) *Grammaire Sanscrite*. Adrein-Maisonneuve, Paris.

Rice, Keren, and Peter Avery (1995) Variability in a Deterministic Model of Language Acquisition: A Theory of Segmental Elaboration. in Archibald, John (ed.) *Phonological Acquisition and Phonological Theory*. Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 23-42.

Sagey, Elizabeth C. (1986) *The representation of features and relations in non-linear phonology.* MIT PhD dissertation.

Sandler, Wendy (1989) *Phonological representation of the sign: Linearity and nonlinearity in American Sign Language.* Dordrecht: Foris.

Shannon, C.E., and W. Weaver (1949) *The mathematical theory of communication.* Urbana: University of Illinois Press.

Wanner, Georg (1941) *Die Mundarten des Kantons Schaffhausen: Laut- und Flexionslehre* [Hans Wanner (ed.)]. Frauenfeld: Huber.

Whitney, W.D. (1960) *Sanskrit Grammar*, [9th issue of 2nd ed.] Harvard University Press, Cambridge, Mass.

Zwicky, Arnold M. (1970) Greek-letter Variables and the Sanskrit ruki Class. *Linguistic Inquiry 1*: 549-55.