

WORKING PAPERS IN LINGUISTICS # 25

PAPERS IN PHONOLOGY

Edited by Arnold M. Zwicky

Papers By

Donald G. Churma, Roderick D. Goman

and Lawrence Schourup

Department of Linguistics

The Ohio State University

1841 Millikin Road

Columbus, Ohio 43210

January, 1981

OSU Working Papers in Linguistics # 25

Papers in Phonology

The five papers in this volume deal with various aspects of phonological theory, and all show the influence of Stampe's Natural Phonology. Schourup's contribution (originally a seminar paper) proposes to revive the notion 'basis of articulation' within this framework. Churma's contributions (the first an extended extract from his 1979 Ph.D. dissertation) deal critically with argumentation in various versions of generative phonology. Finally, Goman proposes (in a revised extract from his 1979 Ph.D. dissertation) a Natural Phonological treatment of consonant processes, paralleling Donegan's vowel study in WPL 23.

AMZ

Table of Contents

	Page
List of Working Papers in Linguistics	iv
Lawrence Schourup, "The Basis of Articulation"	1
Donald G. Churma, "Diachronic Evidence for Synchronic Analyses in Phonology"	14
Donald G. Churma, "A Further Remark on the 'Halleian Syllogism'"	59
Donald G. Churma, "Some Further Problems for Upside- Down Phonology"	67
Roderick D. Goman, <u>On the Natural Phonology of Consonants</u>	107

List of Working Papers in Linguistics

- No. 1, December 1967 out of print
 Articles by Dale Elliott, Charles Fillmore, James Heringer, Terence Langendoen, Gregory Lee, Ilse Lehiste, and Sandra Thompson.
- No. 2, November 1968 (OSU-CISRC-TR-68-3) PB-182 596
 Articles by Charles Fillmore and Ilse Lehiste.
- No. 3, June 1969 (OSU-CISRC-TR-69-4) PB-185 855
 Articles by Dale Elliott, Shuan-fan Huang, Terence Langendoen, Gregory Lee, and Ilse Lehiste.
- No. 4, May 1970 (OSU CISRC-TR-70-26) PB-192 163
 Articles by Gaberell Drachman, Mary Edwards, Charles Fillmore, Gregory Lee, Patricia Lee, Ilse Lehiste, and Arnold Zwicky.
- No. 5, June 1969 out of print
Twana Phonology, by Gaberell Drachman.
- No. 6, September 1970 (OSU-CISRC-TR-70-12) PB-194 829
 Articles by Charles Fillmore, Ilse Lehiste, David Meltzer, Sandra Thompson, and Marcel Tatham.
- No. 7, February 1971 (OSU-CISRC-TR-71-7) PB-198 278
 Articles by Alexander Grosu and Gregory Lee.
- No. 8, June 1971 (OSU-CISRC-TR-71-7) PB-202 724
 Articles by Dale Elliott, Michael Geis, Alexander Grosu, Barry Nobel, Ann Zwicky, and Arnold Zwicky.
- No. 9, July 1971 (OSU-CISRC-TR-71-8) PB-204 002
 Articles by Zinny Bond, Richard Gregorski, Andrew Kerek, Ilse Lehiste, Linda Shockey, and Mary Wendell.
- No. 10, August 1971 Edited by Charles Fillmore out of print
 Articles on Case Grammar.
- No. 11, February 1972 ED 062 850
 Articles by James Heringer, Patricia Miller, Lawrence Schourup, and Richard Wojcik.
- No. 12, June 1972 (OSU-CISRC-TR-72-6) PB-210 781
 Articles by Richard Gregorski, David Meltzer, Ilse Lehiste, and Linda Shockey.

- No. 13, December 1972 ED 077 268
 Alexander Grosu, The Strategic Content of Island Constraints.
- No. 14, April 1973 ED
Studies in Phonology and Methodology. Articles by James Hutcheson, Ronald Neeld, Mieko Ohso, Lawrence Schourup, Holly Semiloff, Clare Silva, and Arnold Zwicky.
- No. 15, April 1973 ED 082 566
 Articles by Angeliki Malikouti-Drachman, Gaberell Drachman, Mary Edwards, Jonnie Geis, and Lawrence Schourup.
- No. 16, December 1973 ED
Mostly Syntax and Semantics. Articles by Michael Geis, Fred Goodman, Patricia Lee, Ronald Neeld, Jerrold Sadock, Clare Silva, and Arnold Zwicky.
- No. 17, May 1974 \$2.00
 Articles by Sara Garnes, Ilse Lehiste, Patricia Miller, Linda Shockey, and Arnold Zwicky.
- No. 18, June 1975 \$2.00
 Articles by Michael Geis, Sheila Geoghegan, Jeanette Gundel, Ronald Neeld, Geoffrey Pullum, and Arnold M. Zwicky
- No. 19, September 1975 Edited by Robert Herbert \$2.50
Patterns in Language, Culture, and Society: Sub-Saharan Africa. Proceedings of the Symposium on African Language, Culture and Society.
- No. 20, September 1975 Edited by Robert Herbert \$2.50
Proceedings of the Sixth Conference on African Linguistics.
- No. 21, May 1976 Edited by Arnold Zwicky \$2.50
Papers on Nonphonology. Articles by Steven Boër, Marion Johnson, Robert Kantor, Patricia Lee, William Lycan, and Jay Pollack.
- No. 22, February 1977 Edited by Olga Garnica \$3.50
Papers in Psycholinguistics and Sociolinguistics. Articles by Mary Louise Edwards, Sara Garnes, Olga K. Garnica, Roy C. Major, and John Perkins.

No. 23, December 1978 \$3.50
Patricia Donegan, On the Natural Phonology of Vowels.

No. 24, March 1980 \$3.50
Edited by Arnold M. Zwicky
Clitics and Ellipsis. Articles by Robert J. Jeffers,
Nancy S. Levin, Gregory T. Stump, and Arnold M. Zwicky.

Orders for those issues still available (indicated by price) should be sent to:

OSU WPL
Department of Linguistics
The Ohio State University
18rl Millikin Road
Columbus, OH 43210

Issues with PB numbers are available through:

The National Technical Information Center
The U.S. Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22151

Issues with ED numbers--11, 13, parts of 14 (articles by Holly Semiloff-Zelasko, "Glide Metatheses" [ED 104 140] and Clare M. Silva, "Metathesis of Obstruent Clusters" [ED 105 735], 15, and parts of 16 (articles by Clare M. Silva, "Adverbial -ing" [ED 095 720], Clare M. Silva and Arnold M. Zwicky, "Discord" [ED 095 719], and Arnold M. Zwicky, "Bibliography I, Voivs" [ED 096 826] and "Bibliography III, Forestress and Afterstress, Compounds and Phrases" [ED 095 721]) are available from ERDS (ordering information is in the back of each issue of Resources in Education):

ERDS
ERIC Clearinghouse on Language and Linguistics
Center for Applied Linguistics
3520 Prospect St., N.W.
Washington, D.C. 20007

The Basis of Articulation*

Lawrence Schourup

O. Introduction.

Most of those who have written about the basis of articulation have done so with mixed feelings. Bloomfield considered the basis worth discussing but characterized observations regarding it as necessarily "vague...hazy and inaccurate" (1933:127-8). Sweet, though according the topic prominent mention (1906:74-5; 1911:4), cautioned that "no language carries out the tendencies of its basis with perfect consistency" (1906:75). Malmberg, likewise granting the existence of an "articulatory basis", refers to the term as "a convenient, but not strictly scientific label" (1963:71). The reaction of Vildomec appears to typify that of many writers: claiming not to know what the basis is exactly, he assures us that it is nonetheless "of primary importance" (1963:218). Definitions have been attempted, such as this disconcertingly vast one by Honikman¹

the gross oral posture and mechanics, both external and internal, requisite as a framework for the comfortable, economic, and fluent merging of and integrating of the isolated sounds into that harmonious, cognizable whole which constitutes the established pronunciation of a language (1964:73)

but despite her attempt to refine the notion and revive interest in it, and the more recent attempt to do so by Drachman (1970), present-day linguists have all but abandoned this traditional concept.

There are at least two reasons for the current neglect of the basis of articulation. The first is practical: as Table 1 indicates, almost every aspect of phonology has on occasion been consigned to the basis. Clearly, by swallowing up all these considerations the basis has made itself unapproachable. But there have also been theoretical reasons to ignore the basis: Chomsky and Halle specifically exclude it from consideration on grounds that its effects are "not locatable in particular segments but rather extend over entire utterances" (1968:295). As such, the basis, though acknowledged to exist, is seen as essentially irrelevant to both underlying and phonetic representation, hence to phonology.

Table I
Some Aspects of Phonology Included in the Basis of
Articulation According to Various Writers

1. Favored position of the tongue (DHfHMST)
2. Degree of lip activity (DHMS)
3. 'Gravitation' of all articulatory muscles toward a particular locus or axis (H)
4. Syllable division (BD)
5. Degree of tension of the articulators (HaH MV)
6. Shape of lips (H)

7. Characteristics of timing, stress, and pitch (TV)
8. Time consumed by articulatory gestures (HT)
9. Segment inventories (BDM)
10. Retraction of the jaw (B)
11. Voice onset time (D)
12. Features of the articulators determined by race (T)
13. Precision of articulation (BM)
14. Location of resonance centers (D)
15. Extent of articulatory gestures (B)
16. Spread of nasalization (D)
17. Degree of nasalization (M)
18. Psychological dominance of vowels over consonants (D)
19. Point of articulation (D)
20. Diphthongization (M)

- B - Bloomfield 1933
- D - Delattre 1966
- H - Honikman 1964
- Ha - Haden 1938
- Hf - Heffner 1950
- M - Malmberg 1963
- S - Sweet 1906
- T - Thalbitzer 1904
- V - Vildomec 1963

1. The basis and Natural Phonology.

It is within the context of a natural theory of phonology that the possible significance of global properties of utterances reemerges. While at present the basis of articulation as such plays no important role in any phonological theory, the theory of Natural Phonology (Stampe 1969, 1973; Donegan and Stampe 1979) invites reassessment of the basis in a way that the standard generative model does not: in the generative framework a theory is explanatory if it provides a description of the set of possible grammars and a procedure for selecting the correct grammar for given data (Chomsky 1965:34). Natural Phonology identifies explanation instead with determining how phonology is "governed by forces implicit in human vocalization and perception" (Donegan and Stampe 1979:126), thus inviting the question of whether the way the tract is set up for speaking affects the nature and interaction of these phonological forces ('processes').

A difficult hurdle stands before anyone who would resuscitate the basis of articulation, however. Even if it can be shown that, say, the French tongue 'prefers' a particular position, how can we know that this position does not simply reflect the rule system of French; that is, how do we know that the favored tongue position is not a secondary effect deducible from the rule system of the language by somehow plotting what would be the most convenient 'homing' position for the articulations of French? If the basis of a language is of any great phonological importance, it must to some extent play a determining role.

It is at first hard to see what kinds of evidence might be brought to bear on the issue. If language L homes to tongue position T and employs process P, application of which is facilitated by T, we certainly can't claim that T is responsible for the existence of P in L; neither can we indisputably claim the reverse--that T is determined by the processes

of L--since these two possibilities are superficially identical. A reasonable response to this difficulty is to ignore it and view the basis as a pleasant mystery. But if the basis of a language does in fact determine aspects of its phonological structure, continuing to ignore the basis would hobble our understanding of phonology. In the remainder of this paper I want to explore evidence for the existence of bases of articulation and for their role in determining phonological structure. To make the discussion manageable, I will restrict attention to positional settings of the tongue and lips. This is not to imply that the basis is less ornate than Table I would suggest; it seems necessary to start, though, with something observable and fairly simple.

If bases of articulation exist, their reason for doing so is surely to accommodate the articulations of particular languages. What does not seem to have been properly appreciated is that the influence will also pass in the opposite direction. The point here is that there is a direct relationship between phonetic difficulty and actual physical properties of the tract. Lenition processes respond to specific difficulties involved in achieving the successive articulatory positions required for speech, but different things will be difficult for a 'normal' speaker and one whose tongue, for example, has been partially excised for medical reasons, or whose tongue is very large, or who suffers from a severe overbite. If it makes sense to speak of naturalness at all, lenitions must respond to the actual tract. Strictly speaking, there is of course one genetic tract per person, but if the muscles can be set up differently, the physical reality upon which the articulations of a language are imposed will also differ. If such settings exist, the tract set up by a native speaker of English and the one set up by a native speaker of French are physically distinct, and what is easy for the one will to an extent differ from what is easy for the other in ways that cannot be understood by looking at a universal set of processes defined with respect to the genetic tract alone.

Drachman (1970) suggests that the basis "is required in order to trigger the presently (plausible and) productive rules of the language... Thus, [for English] Palatalization but not Spirantization, vowel nasalization and flapping of dentals but not the English Vowel Shift" (475). If this were the whole story, the basis could be viewed as determined by the rules. But Drachman adds that in acquisition the basis "may... exceptionally, dominate the rule system, with the result that rules are... modified, reordered or even suppressed" (475). This bilateral influence on each other of the plausible, productive rules and the tract Drachman terms a "conspiracy". While he adduces no specific examples of children hitting on an incorrect basis, such effects probably do occur.² However, I would question the notion that tract dominance is a matter of modifying, limiting, or suppressing phonological processes, and I will argue in section 2.5 below that tract dominance needn't be exceptional.

To the extent that processes dominate the basis, it does make sense to say that the basis "triggers" them. This is simply a precise restatement of the traditional claim that the basis facilitates certain characteristic language-particular features of pronunciation. But it is peculiar to state that in the opposite case--when the basis dominates--the same processes are suppressed, modified, or limited by the basis, for this deprives the processes of their phonetic motivation. Suppose that a child mistakenly adopts a basis that fails to facilitate palatalization in a language in

which adults do palatalize. The child has not thereby suppressed palatalization, but has only adopted a tract configuration for which palatalization is irrelevant.³ It is not that the process is muzzled, but that the dog-house is empty; for that child with that tract set as it is, palatalization is not a natural, phonetically motivated process. If we wish to think of natural processes as representing the system of implicit restrictions on the speech capacity, we cannot simultaneously hold that processes which do not apply in a language due to its basis are suppressed or limited, since this ascribes to them some metaphysical existence apart from the basis in phonetic reality which chiefly recommends them to us.

The point is that processes conspire with the basis more fundamentally than Drachman suggested: the processes of a language exist with respect to particular bases which provide their phonetic motivation. I expect that if this fact is taken properly into account, it will initially complicate the study of phonology, but will eventually lead to better understanding of the differences between languages.

2. Reality of the basis.

To this point in the discussion it is an open question whether the basis is a mythic beast or a psycho-physical fact. Indeed, most of the evidence in favor of bases of articulation is of a questionable sort, namely, impressions gained listening to speakers of a language or attempting to approximate their speech. Honikman's (1964) and Sweet's (1906) comparisons of French, German, and English, for example, are of this sort. The best evidence would be cineradiography or some less carcinogenic technique for observing the articulators during speech. There seem to be no published cineradiographic studies that specifically address this question, which is not surprising, since the task is formidable. It would require close comparisons of series of measurements from the running speech of numbers of informants using very high speed film. Such studies are definitely needed to validate claims about the basis and would, I feel sure, repay the effort.

2.1. Hesitation vowels.

Lass suggests that hesitation vowels "might...be a source of information about a truly linguistic 'neutral position'" (1976:44); a connection between hesitation vowels and the basis is also implied by Hinds (1973:259). The idea that the hesitation vowel of a language and its basis are related has great appeal.⁴ The hesitation vowel in French is mid to low, somewhat front, and often somewhat rounded (discussion below). The articulatory position that would produce this vowel is precisely that which has been imputed repeatedly to the French basis (e.g. by Sweet, Delattre, Honikman, Bloomfield; a parallel claim can be made for German, where the basis is traditionally deemed similar to that of French, though not quite so front).

2.2. Incidence of English uh.

There appear to be no careful studies of hesitation vowels in particular languages. In an informal study of my own I found that there is not as much variation in these vowels in English as the literature suggests. Key (1977:94) mentions various hesitation 'noises' for English. Coupling her list and mine, a small set emerges:

[ə:] [ə̃m] [əb] [ɛ:] [æ:] [ɛ:] [a:].⁵

What this list fails to express is the overwhelming predominance of [ə:]. I recorded the incidence of hesitation vowels for 42 speakers while listening to a radio call-in show. Of 42 callers, 41 used hesitation vowels. Of these 41, 39 used [ə:], three also used [əm] and one used [əb]. Evidently, [ə:] is the hesitation vowel par excellence in English (cf. Maclay and Osgood 1959:24). Although er is a frequently cited possibility and occurs in comics and transcriptions of conversations, I have heard it in speech only a few times, and invariably it marked occurrence of a speech error, as in "I know it from the show...er...the movie." [ə:] in this case may be a reduced form of the word or. Note, though, that er is the British spelling of [ə:] or [ɜ:]; so some American English [ɜ:]'s might arise as spelling pronunciations.

2.3. Function of uh.

There is evidence that uh functions in discourse. It occurs more in dialogues than monologues (see Rochester 1973), no doubt due in large part to its use to hold the turn. Filled pauses occur less frequently in stories to passive audiences (Levin and Silverman 1965), where the turn-taking dynamics are relaxed. Consistent with this, Davy and Quirk (1969:114) comment that 'voiced pause' is not used when speaking to oneself. Studies of the placement of uh and other hesitations (Boomer 1975; Goldman-Eisler 1958) indicate that uh appears primarily in three locations: (a) at grammatical juncture, (b) at other constituent boundaries, and (c) before the first content word within a constituent. Such studies have not been sensitive to the discourse functions of uh. Without recourse to discourse structure, placement of uh before the first content word of a constituent cannot be fully explained. It is clear that to hold a turn, one could begin a constituent before completely planning it, since beginning at all reserves the turn, while not doing so invites an interruption (cf. Sacks, et al 1974:718-20). I have observed four distinct uses of uh in conversation: (a) to reserve the turn during, for example, word search, (b) to indicate desire to take the turn, (c) to indicate disagreement with what another has said without intention to take the turn, and (d) to indicate presence in conversational settings where this might be questioned, for example, on the telephone.

To say that uh has complex functions in discourse is not, however, to disqualify it as a neutral vowel with properties attributable to the basis. Three considerations bear on this: (a) If uh is learned as an arbitrary word, there is no way to account for the identity of uh's vowel with that of the other two most common filled pause alternatives, [əm] and [əb]; (b) Filled pauses in various languages⁶ are never fully high, back, front, or rounded, though they can approximate these qualities to a degree (see my comments on l'e muet below). If filled pauses are learned, we might expect to find [o], [u], [ʊ], [ü], or even a diphthong turning up as the regular hesitation vowel in some languages, but we don't; (c) One would expect the hesitation vowel of a language to involve the minimal vocal gesture that will hold a place in speech. Simply initiating voicing is the easiest way to accomplish this. The quality of the vowel would then directly reflect the tongue position of the basis of articulation.

It is almost certainly incorrect to regard uh as a speech error (e.g. Clark and Clark 1977:262), both because it is clearly functional, and because it appears to interact with a linguistic rule. Jefferson (1975:183-4) notes that the definite article in English regularly takes its prevocalic form [ðɪ] before uh, rather than its preconsonantal form [ðə] (unless a speaker wishes to convey the impression of correcting an error, in which case the wrong variant may be chosen to display this intention). Jefferson concludes that uh is "at least a projectable syntactic unit, and is perhaps characterizable as having the status of a word in the English language." The point I wish to make is that while uh is certainly a linguistically functional unit, and may even have the status of a word, its phonetic quality is not arbitrary; in fact, one could argue that this vowel is non-phonemic.

2.4. French schwa.

As mentioned above, the hesitation vowel in French is somewhat front and often rounded. The same is true of l'e muet, the French schwa in le, je, que, etc. A parallel statement can be made regarding the English unstressed schwa in sofa, which seems indistinguishable from the English hesitation vowel. The quality of the French schwa has been hotly disputed among those who care. It is now fashionable to regard l'e muet as central, and in one way this seems correct. There is excellent evidence that l'e muet doesn't significantly overlap with the French front rounded vowels and belongs further back than them on an acoustic vowel diagram. Schane (1968:30) is correct to assign l'e muet the feature [-front] based on Pleasants' general conclusion that

il a son point d'articulation en arrière de celui de eu fermé
et même de eu ouvert...Les caractéristiques de l'articulation
de [ə]...sembleraient indiquer que [ə] est une voyelle centrale.

(1956:247)

However, the phonology and the phonetics of this vowel are distinct matters. It is clear from other remarks by Pleasants, including the following, that French schwas are in fact somewhat fronted:

L'[ə] de nos expériences se place juste à la limite qui sépare
les voyelles antérieures des voyelles centrales. (1956:58).

For further evidence that l'e muet is phonetically fronted, we may look to its stressed variant. It is often claimed that in imperatives like dites-le, where normally unstressed schwa received stress, /ə/ is replaced by /ɸ/. The rationale for this is that speakers recognize that /ə/ cannot bear stress and so substitute another vowel for it (see Price 1971:78). If this analysis is correct, the phonetic fronting of /ə/ would explain why /ɸ/ in particular is chosen as the substitute for /ə/. But a careful look at the phonetic quality of stressed schwa makes the substitution of /ɸ/ look implausible. Pleasants argues extensively (1956:38-43; 59; 72-3; 253-68) that stressed /ə/ is phonetically distinct from both stressed /ɸ/ and /æ/. It appears, therefore, that stress simply enhances the frontness of /ə/.

What I am suggesting is that if a language uses a neutral vowel in positions of reduction, the quality of the vowel will be identical to that of the hesitation vowel of the language, and both will reflect the basis. In some languages the neutral vowel will not be transparently equivalent to the hesitation vowel because of coarticulation with adjacent segments, but one might expect the basis coloring to come through, for example, in open syllables after [h]. These remarks are of course speculative, but the identity of the hesitation and schwa vowels in French (and, apparently, English) makes it at least plausible that schwas reflect the basis.⁷

2.5. Epenthesis.

Daly and Martin (1972) suggest a minor connection between the basis of articulation and the phonological rules. They claim that the phonetic properties of epenthetic vowels introduced for syllabification "are at least partially determined by the base of articulation of the particular language" (1972:608). They propose the following three restrictions (1972: 610):

1. A language may have an epenthetic /i/ if and only if that language has a palatal or a palatalized series of consonants;
2. A language may have an epenthetic /a/ if and only if it has a pharyngeal or pharyngealized (perhaps a glottal or glottalized) series of consonants;
3. A language may have an epenthetic /ɨ/ if and only if it has a labio-velar or (labio)velarized series of consonants.

Unfortunately, they cite no data in support of these restrictions,⁸ nor of their general claim.

Responding to Daly and Martin, Hinds (1973) argues against these claims using examples of marked epenthetic vowels in languages which have no marked consonants. He argues exclusively from loan words in Japanese and Korean, citing as precedent Daly and Martin's use of Turkish loan words to illustrate a claim about harmony. The use of loan phonology invalidates Hinds' objections, however, since vowel insertion in borrowing is fundamentally different from proper epenthesis. Hinds' evidence against two of Daly and Martin's claims is from Japanese, which 'epenthesizes' [i], [e], [a], [o], and [u] in loan words. Ohso (1973), however, makes it clear that the insertion of these vowels involves interpreting the foreign target with respect to the segment inventory and processes of Japanese. In fact, from the point of view of a speaker of Japanese, such vowels are not inserted at all--they are seen as vowels that were incorrectly deleted by the foreign speaker. They are insertions only from the point of view of the language borrowed from.

Daly and Martin are probably right to say that the basis has a hand in determining the quality of epenthetic vowels, but it will be difficult to sort out the basis-influenced vowels, because some languages epenthesize phonemes unrelated to their basis; Egyptian Arabic, for example, epenthesizes [i] in all epenthetic environments (Broselow 1976), though it shares with other forms of Arabic a rather pharyngeal, rather low, rather back hesitation vowel. That other languages do epenthesize basis-colored vowels is indicated by the epenthesis of l'e muet in French (Pleasant 1956:155-60; Schane 1968:31-2) or of schwa in English (as in athlete, [æθəlit]).

2.6. French [R].

Delattre (1966:10-11) discusses 'le Mode Antérieur' of French, a collection of characteristics which together give French pronunciation a fronted quality. In his words, "parler sur le Mode Antérieur veut dire porter les lieux d'articulation, les centres des cavités de résonance, le plus possible vers l'avant de la cavité orale" (10). He includes the following as characteristic of this mode: (a) the convex and bulged-forward shape of the tongue body; (b) by comparison with English, a further forward point of articulation for [t], [d], [n], and [l]; (c) pronunciation of [s], [z], [ʃ], and [ʒ] with the apex of the tongue curved downward so that the fricative aperture is strictly laminal; (d) preponderance of front vowels over back ([ieyøœ] vs. [uoc]); (e) preponderance of rounded vowels over unrounded ([yøœuoc] vs. [ieE]); (f) greater coarticulatory rounding (especially of labial consonants: cf. Fr. pour and Eng. poor) with following rounded vowels than in English.

Interestingly, Delattre includes French uvular [R] as a further symptom of this anteriority. On the historical replacement of apical trilled [r] by [R], Delattre says, "c'est grâce à cet r dorsal que la langue peut conserver sans interruption la position bombée convexe qui favorise la résonance antérieure générale" (1966:11). A Frenchman whose tongue tip suddenly adhered to his lower teeth would be a gastronomical cripple, but could enunciate his order perfectly.⁹

It would be wrong to say that the French basis of articulation determines that French /r/ should be uvular; this is simply not the case. Some dialects of French preserve the apical trill, and there was a long period in the history of French when the apical /r/ lingered despite the presence of other aspects of anteriority.¹⁰

The issue that all of this raises is an historical one. There is too much anteriority in French to be due to mere accident. The change to uvular [R] in particular is striking. It appears that the basis of articulation of French is holding sway in the court of phonological change. The effect seems at odds with that of push and drag chains, which aim at distinctiveness. The general tenor of developments in French has been, as much as possible, to move the focus of articulation to the front part of the mouth. There are limits on how far this kind of thing can go, but French seems to be doing its best to reach them.

It is precisely in a case like this that the basis, nonexceptionally, influences the processes of a language. The change from [r] to [R] accommodates the basis, rather than the other way around. The picture that emerges is this: the primary function of the basis is to accommodate the articulations, but in so doing the basis becomes itself a thing, with its own habits and sluggishness, and so influences the kinds of variation likely in the language, and thus also its diachronic development. It seems very likely that this state of affairs can explain why it is that some languages retain quite marked series of consonants--say, pharyngeal ones--over long periods of their development, without tending to give them up in favor of less marked articulations. Just as processes must be defined with respect to the basis, so must the markedness of segments. The claim that pharyngeals are marked should come as a surprise to the speaker of Arabic, whose basis is low and back.

3. The basis and 'neutral position'.

Jakobson et al. (1951) proposed the existence of a 'neutral position' of the tract, assumed to be universal and to approximate the position for a very open [æ]. The neutral position is one taken during speech, since it is claimed to be important "for predicting the effects on formant positions of variations in the overall length of the vocal cavity of different individuals," and "also serves as a reference point for the tenseness feature" (18).

Chomsky and Halle (1968:300) adopted the notion of a speech neutral position, though they specified the position of the tongue body as roughly that for the vowel in English bed, but with the blade of the tongue at rest (compare 1968:300 and 304). Lass makes a revealing comment in this connection:

It is interesting that the neutral position in SPE is much closer than that given by Jakobson et al...there is no discussion of why it has shifted so far up, which tends to make one suspicious that it is an analytical convenience rather than a fact about languages. Actually Chomsky and Halle need an essentially 'front' and 'mid' neutral position, because the features [high, low, back] are defined in terms of the deviation from just such a position. (1976:44).

Chomsky and Halle distinguish the neutral position from that associated with the basis of articulation (1968: compare 295 with 300), which allows them to claim that the neutral position is universal--and use it as a reference point for the distinctive features--while claiming that the basis is not. This gambit, however, leaves them in the awkward position of claiming that the deviations from neutral involved in producing, for example, a [+high] segment are deviations from the--at that point--abstract speech neutral position, not the actual in-speech homing position determined by the language-specific basis.

I know of no evidence whatsoever for the separate existence of a universal neutral or speech-ready position, nor for that matter, any evidence that the positions for unfilled and filled pause are distinct, or that any of the positions just mentioned are in fact distinct from that for the basis of articulation.

4. Secondary articulations.

Daly and Martin (1972:612) observe that while the presence of velarization in a language tends to correlate with the quality of the basis of articulation, the presence of a series of velars does not; in general the basis of a language is more associated with its secondary than its primary articulations. This can be explained by referring to Perkell's observation that in cineradiographic studies there can be observed two separate articulatory systems at work, a slow and gross vowel-producing system ('extrinsic') and a quick, precise consonant-producing system ('intrinsic'):

...the production of a consonant can be thought of as being a gesture superimposed on the continuously varying vowel producing system...coarticulation effects of vowels are, for the most part, manifested by influencing the position of the consonant-articulating organs rather than by altering the manner of articulation...

the positioning element of consonant production is performed by the slow extrinsic system and it is strongly influenced by coarticulatory effects. This positioning aspect presumably also operates to produce secondary features of consonant articulation such as palatalization, labialization, and pharyngealization.

Deformation of the articulatory organs is superimposed on the positioning element, and the deformation is performed by the action of fast, precise intrinsic musculature. (1969:65-66)

If we consider the positioning elements of the basis to operate on the gross, extrinsic system, we can neatly account for its correlation with secondary rather than primary articulation.

5. Conclusion.

As I have been using the term 'basis of articulation', it designates the language specific homing positions of the articulators in running speech. Such positions appear to directly determine the quality of the predominant hesitation vowel of a language and may also influence the quality of vowels used in positions of reduction and epenthesis. In a larger discussion of the basis it would be necessary to include many of the other global phonological properties of utterances listed in Table I, and others.

In general it seems necessary to distinguish setting phenomena from 'local' ones, like processes, and to acknowledge the connection between the two. I have suggested that this connection is one of interdependence: the basis accommodates the articulations of a language, but the direction of influence may also be reversed so that the basis is itself accommodated. Finally, I have claimed that to characterize such notions as 'phonetic motivation' and 'markedness' without reference to bases of articulation is to buy universality of description at the expense of phonetic reality.

Research needed to clarify this area of phonology includes the following: (a) careful instrumental studies of the homing positions of the articulators in running speech for various languages; (b) close measurements of the phonetic properties of naturally occurring hesitation vowels for various languages; (c) comparison of the data from (a) and (b); (d) cross-language studies of the historical persistence of bases of articulation; (e) studies of the acquisition of bases of articulation and their influence on the sound substitutions children use.

Footnotes

*I am grateful to Arnold M. Zwicky for many helpful comments. Thanks also to Jonas Nartey for suggesting readings on cineradiography, and to Paul Gallagher and Mohammed Sawai'e for interesting examples.

¹Not all of the writers I will mention use the term 'basis of articulation' but they all seem to be getting at the same thing, whether they use 'organic basis' (Sweet), 'tract setting' (Honikman), or one of a half dozen other terms.

²While I am aware of no published remarks to this effect, it is a commonplace among students of child language that some children begin with, for example, a very palatal(ized) inventory, or lots of rounding.

³I am not talking here about a child who is trying not to palatalize for some reason and uses the basis to accomplish this goal (if such a thing ever happens) but the simpler case of a child who just plain gets the basis wrong.

⁴It may not be that the hesitation vowel of a language is invariably determined by its basis of articulation. It is conceivable that in some languages this vowel would be learned independently, putting it on a par with 'hesitation words' like the English interjection well, but I know of no such languages.

⁵Except with [ɛ̃m] and [əb], some nasalization is optional, probably a result of not controlling the velum rather than intending to produce a nasalized vowel.

⁶I hesitate to cite the hesitation vowels for languages other than English, French, and German, because I have gotten conflicting opinions from native speakers of other languages. Often an approved literary hesitation vowel is offered, but it differs from what the speakers really do. I prefer to send this question to the phonetics lab--it is, after all, a simple enough matter to make spectrograms of hesitation vowels occurring in ordinary conversation.

⁷Once again, it is best left to the laboratory to determine if a language's schwa and its hesitation vowel are alike.

⁸I doubt that these restrictions are right in lumping together primary and secondary articulations. See my comments in 4.0.

⁹It is amusing and instructive to try to read a passage of French aloud with the tongue loosely held in the position just described, and then a passage of English. The English is quite distorted, while the French sounds only faintly, if at all, unnatural.

¹⁰The change to [R] cannot be dated precisely, but it is known to have occurred in the late seventeenth or early eighteenth century. A thorough discussion of this change is found in Nyrop (1914:42-8).

References

- Bloomfield, Leonard. (1933). *Language*. New York: Holt, Rinehart and Winston.
- Boomer, Donald S. (1965). Hesitation and grammatical encoding. *Language and Speech* 8.148-9.
- Broselow, Ellen I. (1976). *The Phonology of Egyptian Arabic*. Ph.D. Dissertation, University of Massachusetts.
- Chomsky, Noam. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam, and Morris Halle. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Clark, Herbert H. and Eve V. Clark. (1977). *Psychology and Language*. New York: Harcourt Brace Jovanovich.

- Daly, D. M. and L. W. Martin. (1972). Epenthetic processes. *Papers in Linguistics* 5.4.
- Davy, D. and R. Quirk. (1969). An acceptability experiment with spoken output. *Journal of Linguistics* 5.109-20.
- Delattre, Pierre. (1966). *Studies in French and comparative phonetics.* *Janua Linguarum* 18. The Hague: Mouton.
- Donegan, Patricia J. and David Stampe. (1979). The study of natural phonology. in *Current Approaches to Phonological Theory*, ed. by Daniel A. Dinnsen. Bloomington, IN: Indiana University Press.
- Drachman, Gaberell. (1970). Rules in the speech tract. in *Proceedings of the Sixth Annual Meeting of the Chicago Linguistic Society.*
- Goldman-Eisler, Frieda. (1958). Speech production and the predictability of words in context. *Quarterly Journal of Experimental Psychology* 9-10: 96.
- Haden, Ernest F. (1938). The physiology of French consonant changes. Language dissertation No. 26. Supplement to *Language* 14.4.
- Heffner, R-M. S. (1950). *General Phonetics.* Madison: University of Wisconsin Press.
- Hinds, John. (1973). On 'epenthetic vowels'. *Papers in Linguistics* 6.2.
- Honikman, Beatrice. (1964). Articulatory settings. in *In Honor of Daniel Jones*, ed. by D. Abercrombie, et al. London: Longmans.
- Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle. (1952). *Preliminaries to Speech Analysis.* Cambridge, MA: MIT Press.
- Jefferson, Gail. (1975). Error correction as an interactional resource. *Language in Society* 3.181-199.
- Jespersen, Otto. (1904). *How to Teach a Foreign Language.* (trans. by Bertelsen). London: George Allen and Co. Ltd.
- Key, Mary Ritchie. (1977). *Nonverbal Communication: a Research Guide and Bibliography.* Metuchen, NJ: Scarecrow Press.
- Lass, Roger. (1976). *English Phonology and Phonological Theory: Synchronic and Diachronic Studies.* Cambridge: Cambridge Univ. Press.
- Levin, H. and I. Silverman. (1965). Hesitation phenomena in children's speech. *Language and Speech* 8.67-85.
- Maclay, Howard, and Charles E. Osgood. (1959). Hesitation phenomena in spontaneous English speech. *Word* 15.19-44.
- Malmberg, Bertil. (1963). *Phonetics.* New York: Dover.
- Nyrop, Kr. (1914). *Manuel phonétique du français parlé.* 3rd edition. Copenhagen: Copenhagen et Kristiania.
- Ohso, Mieko. (1973). A phonological study of some English loan words in Japanese. *OSU WPL* 15:1-26.
- Perkell, Joseph. (1969). *The Physiology of Speech Production.* Cambridge, MA: MIT Press.
- Pleasants, Jeanne Varney. (1956). *Études sur l'e muet.* Paris: Librairie C. Klincksieck.
- Pope, M. K. (1934). *From Latin to Modern French with Especial Consideration of Anglo-Norman.* Manchester, England: Manchester University Press.
- Price, Glanville. (1971). *The French language present and past.* London: Edward Arnold.
- Rochester, S. R. (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research* 2.51-81.
- Sacks, Harvey, Emanuel Schegloff, and Gail Jefferson. (1974). A simplest systematics for the organization of turn taking for conversation. *Language* 50.696-735.

- Schane, Sanford. (1968). French phonology and morphology. Cambridge, MA: MIT Press.
- Stampe, David. (1969). The acquisition of phonetic representation. in Proceedings of the Fifth Annual Meeting of the Chicago Linguistics Society.
- Stampe, David. (1973). A Dissertation on Natural Phonology. Ph.D. Dissertation, University of Chicago.
- Sweet, Henry. (1906). A Primer of Phonetics. 3rd ed. Oxford: Clarendon Press.
- Sweet, Henry. (1911). A Primer of Spoken English. Oxford: Clarendon Press.
- Thalbitzer, William. (1904). A Phonetical Study of the Eskimo Language. Copenhagen: B. Luno.
- Vildomec, Verobj. (1963). Multilingualism. The Netherlands: A. W. Sythoff-Leyden.

Moreover, it seems to me, it is absolutely crucial that this be the main concern of phonologists, unless one is willing to claim that the only concern of phonology is the specification of the actual pronunciation of utterances. The achievement of even the latter end is, of course, no mean accomplishment, but by itself it is of little or no value to ultimately more important and interesting questions having to do with determining the nature of textual representations (also referred to as underlying representations or underlying forms) and of phonological rules in descriptively adequate grammars of natural languages. Such a concern is, perhaps contrary to popular belief, not peculiar to generative phonology and those theories which have been influenced by and/or have reacted to generative phonology (e.g., the "natural phonology" of Stampe 1973, Hooper and Stampe 1979, and the "Natural Generative Phonology" of Hooper 1975 and the references cited there). Rather, as pointed out by Chomsky (1978:304-5), most American structuralists, for example, implicitly accepted such a concern as legitimate in linguistic research, despite the probability that the great majority of them would have rejected a proposal of this nature if it was explicitly put to them. This must be so because the sequence of papers which appeared shows that a set of procedures for phonetic analysis, say, would be proposed, another investigator would point out that using these procedures would lead to an absurd result in certain cases and perhaps propose a revision of the original principles, etc. Unless these investigators were implicitly accepting a point of view very much like that which concerns itself with descriptive adequacy, there would appear to be no basis for claiming that the results in question were in fact absurd. Descriptive adequacy has thus been an implicit concern in most of those phonological theories in which a concern of this nature has not been explicit, as it has been in the case of generative phonology and the other current theories mentioned above.

Diachronic Evidence for Synchronic Analyses
in Phonology*

Donald G. Churma

1. Introduction. There is currently little agreement among phonologists as to the form which phonological rules and phonological representations in 'correct' grammars of natural languages should take. The term 'correct' is presumably understood by most investigators as being at least roughly equivalent to what Chomsky (1964, 1965) calls "descriptively adequate." This latter concept ("descriptive adequacy") is, according to Chomsky (1964:105), to be equated with what Sapir (1933) terms "psychological reality" (cf. also Chomsky 1976). That is, modern phonologists, like (among others) Sapir and Chomsky, are concerned with determining the form in which speakers of a language store in long term memory the pronunciations of the lexical items of their language, and the form of the rules (if any) which convert the stored pronunciation into the actual pronunciation. To put it in slightly more technical terms, phonologists today are concerned with determining the nature of lexical representations (also referred to as underlying representations or underlying forms) and of phonological rules in descriptively adequate grammars of natural languages.

Such a concern is, perhaps contrary to popular belief, not peculiar to generative phonology and those theories which have been influenced by and/or have reacted to generative phonology (e.g., the "natural phonology" of Stampe 1973, Donegan and Stampe 1979, and the "Natural Generative Phonology" of Hooper 1976 and the references cited there). Rather, as pointed out by Chomsky (1978:304-5), most American structuralists, for example, implicitly accepted such a concern as legitimate in linguistic research, despite the probability that the great majority of them would have rejected a proposal of this nature if it was explicitly put to them. This must be so because the sequence of papers which appeared shows that a set of procedures for phonemic analysis, say, would be proposed, another investigator would point out that using these procedures would lead to an absurd result in certain cases and perhaps propose a revision of the original principles, etc. Unless these investigators were implicitly accepting a point of view very much like that which concerns itself with descriptive adequacy, there would appear to be no basis for claiming that the results in question were in fact absurd. Descriptive adequacy has thus been an implicit concern in most of those phonological theories in which a concern of this nature has not been explicit, as it has been in the case of generative phonology and the other current theories mentioned above.

Moreover, it seems to me, it is absolutely crucial that this be the main concern of phonologists, unless one is willing to claim that the only concern of phonology is the specification of the actual pronunciation of utterances. The achievement of even the latter end is, of course, no mean accomplishment, but by itself it is of little or no relevance to ultimately more important and interesting questions having to

do with the nature of the human mind. That is, the ultimate concern of generative phonologists, in my view (and, at least implicitly, that of most other phonologists), is the light which the study of the phonological part of human language can shed on the nature of the mind. Such light simply cannot be shed unless we go beyond the mere specification of the pronunciation of utterances to concern ourselves with the way in which these pronunciations come about when speakers of natural languages produce them.

Despite the importance of the issues addressed by modern phonology, there is, as noted above, little agreement concerning the resolution of these issues. Probably the most important reason for this lack of agreement is that phonologists until quite recently (with some notable exceptions such as Sapir), have taken as their data almost exclusively the facts of pronunciation in various natural languages. For, while no one can deny that such facts must be considered when attempting to answer the questions discussed above, neither, it seems to me, can it be denied that there are other data which, if they are available, cannot be neglected if we are to find satisfactory answers to our important questions. It has been well known at least since Chao (1934) (and undoubtedly well before then) that pronunciation data alone cannot in general supply a unique answer to phonological questions, even given rather specific assumptions about the form which such answers may take (e.g., the assumptions of classical phonemics).

It is this lack of uniqueness for solutions of phonological problems, of course, which has led to the current widespread disagreement among phonologists as to the correct answers to their questions. Recently (Sapir can again be considered 'recent' in this respect), there have been several attempts to resolve these difficulties by appealing to data other than the pronunciations of utterances themselves. In my opinion, it is evidence of this type (which has been termed "external evidence" by, e.g., Botha 1973, and "substantive evidence" by such researchers as Skousen 1975) which is required to settle the disputes which have arisen concerning whether or not proposed grammars (or parts of them) are descriptively adequate.¹

Moreover, it appears to me, it is issues of this kind (i.e., at the level of descriptive adequacy) which can most profitably be pursued, given the current state of our understanding of the nature of language. It makes little sense to concern ourselves with the admittedly ultimately more philosophically profound issue (cf. the above discussion of "the nature of the human mind") of "explanatory adequacy" (cf., for example, Chomsky 1964:63) of theories of language which provide "a general basis for selecting a grammar which achieves" the still unclear level of descriptive adequacy. The actual practice of contemporary phonologists bears out this assessment: most current theoretical research is concerned with the descriptive adequacy of various proposed grammars, and not with constructing explanatorily adequate theories. Such a concern is entirely appropriate in my view: determination of the nature of descriptively adequate grammars seems clearly to be logically prior to work on explanatorily adequate theories.

Despite the apparently clear desirability of using "external" or "substantive" evidence (hereafter, simply 'external evidence') to decide phonological issues, there is little consensus among phonologists concerning the success of arguments employing such evidence. One major reason

(and perhaps the most important) is that the premises of these arguments have been left, in many cases, implicit. When the premises of an argument are not made explicit, it is difficult to evaluate it, since the truth of the premises is a crucial concern.² It is therefore of considerable importance that these premises be made explicit, insofar as possible, so that they may be evaluated with respect to their truth or falsity (or, probably more appropriately, the likelihood of these premises). Once the premises of an argument have been made explicit, of course, it can be determined whether or not a logically valid form of inference is being employed, as well.

In this paper, I will examine one kind of such evidence, data from language change,³ and attempt to evaluate its relevance for determining the descriptive adequacy (i.e., psychological reality) of synchronic analyses in phonology. This kind of external evidence is particularly important, it seems to me, since, first of all, it has been used probably more frequently than any other kind, but also because, as noted by Sommerstein (1977), it is perhaps the only kind of external evidence which is generally accepted as being relevant to synchronic analysis. The approach taken here is a methodological examination of four arguments which have been put forth in favor of certain phonological analyses, together with replies to them and, in one case, a methodological critique. In so doing, I will attempt to make explicit the premises on which these arguments are based, and to evaluate these premises as potential 'universals' of language change which could serve as a firm basis for using diachronic data in assessing synchronic analyses.⁴ Such an investigation, I feel, is long overdue, given the importance of the issues and the type of evidence involved, as well as the essential lack of any comparable work.⁵ Before proceeding with the investigation, however, I will first give a brief sketch (section 2) of my background assumptions about the nature of phonological systems, and (section 3) of a framework for discussing and evaluating scientific arguments.

2. Background assumptions. In a work of a methodological character, it is well to make as few assumptions as possible, since the entire undertaking could be compromised if one of the assumptions accepted should prove false. However, it is extremely difficult, if not impossible, to do work (even when it is methodological work) in a theoretical vacuum, and I will take this opportunity to specify what will be assumed here regarding the nature of the phonological component of grammars of natural languages.

The assumptions made are given in (1) below. The last assumption (1c) actually follows from (1a) and (1b), but I include it here for the sake of clarity.

- (1) a. the physical speech signal is, at least partially, segmentable into discrete sequential units⁶ (i.e., there is a level which corresponds roughly to the "systematic phonetic" level of Chomsky and Halle 1968)
- b. there may be a level more 'abstract' than the level just specified in that representations at the former level need not correspond one-to-one to those at the latter (the former level will be referred to here as the level of 'lexical representation,' a level which

may or may not be the same--perhaps depending on the lexical item--as the systematic phonetic level)

c. due to the potential discrepancy between the level in (a) and that in (b), there may exist 'phonological rules' which convert representations at the latter level into representations at the former level.

Many of the arguments in the literature which make use of external evidence are concerned with ascertaining the nature of the level of lexical representation, and this work will correspondingly be primarily concerned with this issue.⁷ This issue is quite an important one, since resolving it often results automatically in the resolution of other issues, such as whether or not we are dealing with a case of "rule inversion" (cf. Venemann 1972 and section 4.3 below), but it will not be my exclusive concern.

3. On the nature of scientific argumentation. In Churma (1979, Ch. II), I argued that many scientific arguments can be understood as instances of what has come to be known as Bayesian inference (cf., for Example, Salmon 1967). The logical basis of this form of inference is in probability theory, and it can easily be shown (see Churma 1979, Appendix) that a formula corresponding to this form of inference follows straightforwardly from the axioms of the probability calculus. It can, of course, be quantified, but for present purposes, it will suffice to characterize Bayesian inference as in (2).

- (2) Unless hypothesis H is true, it would be unlikely that piece of evidence E would occur.
E does occur.
Therefore, it is likely that H is true.

Of crucial importance here is how the "unlikely" nature of E is to be determined. I have argued (in some detail, since this particular view does not seem to have been seriously considered in the logico-philosophical literature) in Churma (1979, Ch.II) that such determination is ultimately subjective in nature. Among other things, this view permits an explanation of how it is possible for rational investigators to disagree about the force of a given argument: they disagree because of incompatible beliefs (or degrees of belief) about the truth of the major premise in (2). The following section of this paper can also be taken as an argument in favor of this view.

Another straightforward consequence of the axioms of the probability calculus (cf. Churma 1979, Ch. II and Appendix) is what I have called "'almost deductive' inference". This kind of inference resembles standard deductive inference (in particular, modus ponens and modus tollens), but here the truth of the premises need not be known more than probabilistically. Thus, in the case of 'almost modus ponens', we have the situation illustrated in (3).

- (3) If A, then B is likely.
A is likely.
Therefore, B is (somewhat) likely.⁸

'Almost modus tollens' takes an analogous form. Let us turn now to an examination of the use of data from diachrony in the justification of synchronic analyses.

4. Appeals to diachronic evidence. In this section I will examine four sets of arguments which make use of evidence from historical change, and replies which have been made to each of them. After reconstructing in more explicit fashion the arguments and counterarguments, I will give some discussion of the validity of the argument forms used and (more often) of the truth/falsity or likelihood of the premises on which these arguments are based. The arguments considered are found in Kiparsky (1968), Hooper (1976), Schuh (1972) and Skousen (1972, 1975); the rejoinders are from, respectively, Stampe (1973), Harris (1978), Leben (1974) and Kiparsky (1973a, b). A "remark" on Leben's reply by Schuh (1974) is also given some attention, as is Botha's (1973) methodological critique of Kiparsky's argument. After each presentation of the relevant arguments, I will give some discussion of the cogency of the arguments in question. In the final section, I will discuss the important issue of possibly taking seriously some of these premises (or generalizations of them) as universal principles of historical change which could serve as a legitimate basis for other arguments which attempt to justify synchronic analyses on the basis of data from diachrony.

4.1. Kiparsky-Stampe-Botha.

4.1.1. Kiparsky on the brace notation. I will consider first of all the argument presented in Kiparsky (1968) (and reiterated in Chomsky and Halle 1968) in favor of the psychological reality of the brace notation. This argument is noteworthy in that it has received extensive, and essentially negative, discussion by Botha (1973). It is also a particularly transparent application of the Bayesian schema (1). Furthermore, Stampe's (1973) reply (discussed in the next subsection) explicitly acknowledges "the beauty of Kiparsky's argument" (Stampe 1973:48), although Stampe argues against Kiparsky's conclusion. It thus appears worthwhile to consider this argument in some detail; I reproduce below verbatim in (4) Kiparsky's argument (the numbering in the quotation is Kiparsky's).

(4) In English, underlying long vowels, which are otherwise realized as diphthongs, are shortened in two main phonological environments: before two or more consonants (for example, keep:kept) and in the third syllable from the end of the word (for example, vain:vanity, severe:severity). The rules which bring these shortenings about are the following:

- 5'. V → [-long]/__CC
5''. V → [-long]/__C...V...V

The theory of generative grammar requires that 5' and 5'' be collapsed into a single rule as follows:

5. V → [-long]/__C^C{...V...V

It asserts that of the two descriptively equivalent grammars, one of which contains the two rules (5' and 5'') as separate processes, and the other as a single

process combined into 5 by factoring out their common part and enclosing the remainder in braces, it is the latter which is the psychologically correct one.

Rule 5 arose in Early Middle English as a generalization of a much more restricted process of shortening. In Old English, vowels were shortened before three or more consonants (for example, gōdspell > godspell, bræmblas > bræmblas) and in the third syllable from the end provided they were followed by two consonants (for example, blēdsian > bledsian). The corresponding rules were:

6'. V → [-long]/ CCC

6''. V → [-long]/ CC...V...V

Again, these rules must be collapsed as before:

6. V → [-long]/ CC^C...V...V

On comparing the Old English rule in 6 and the Early Middle English (and indeed Modern English) rule in 5 we see that the only difference between them is that the later rule (5) has lost one of the required consonants in its environment. It represents a simpler, more general form of the Old English vowel-shortening process. It will apply in all cases where 6 applied but also in cases where 6 would not have applied. Evidently the change from 6 to 5 is an instance of simplification, which we have seen to be one of the basic mechanisms of linguistic change. But in a linguistic theory in which the brace notation plays no role, the relation between the Old English and Early Middle English shortening process is a different one. If the brace notation were not part of linguistic theory we would have two separate changes--namely, 6' > 5' and 6'' > 5''--on our hands and we would be faced with the very peculiar fact that two separate, unrelated rules have undergone an identical modification at the same point in the history of English.

The last sentence of this passage bears a strikingly close relationship to the major premise in the Bayesian schema (2). I rephrase this sentence slightly here to emphasize the parallel, as in (5) below.

- (5) Unless the brace notation were a part of linguistic theory, it would have been unlikely for the phonological system of Old English to have changed into that of Early Middle English in the way that it did.

It is clear that Kiparsky regards such a change as quite unlikely from his characterization of it as a "very peculiar" one. When we add as a minor premise a statement that such a change did in fact occur, of course, it follows from (1) that the brace notation is (strongly) supported (the more unlikely the change in question is considered, the stronger the support offered by the fact that it did occur), as long as we agree with these premises.⁹

4.1.2. Stampe on Kiparsky's argument. Stampe (1973:48) begins his critique of Kiparsky's argument by, as noted above, "granting the beauty of this argument..." I take this as implying that Stampe, at least, feels that the argument form employed by Kiparsky is a legitimate one (as well as a "beautiful" one). This, despite the fact that Stampe does not accept Kiparsky's conclusion, is an indication of the strong intuitive appeal of the Bayesian argument schema.

What apparently motivated Stampe to question the soundness of Kiparsky's argument (though, again, not the validity of the argument form) seems to be that it "seems odd that phonetically motivated changes of this sort, changes which at least in their inception were imposed on the language by its speaker, and not vice versa, should be subject to the sort of cognitive analysis implied by the brackets" (Stampe 1973:48). In terms of the Bayesian schema (1), even if the brace notation were psychologically real, changes of this type would be unlikely.¹⁰

Stampe goes on to present an alternative analysis of the facts of Old English and Early Middle English, one which would not make a change of the type that occurred appear unlikely, thus disputing Kiparsky's (implicit) claim about the unlikelihood of this change without the psychological reality of braces. In so doing, however, he not only disputes the value of the probability that E is true unless H is true; he also calls into question whether or not E is actually true. Thus, instead of Kiparsky's rules, Stampe (pp. 48-9) has (6) and (7) for Old English and Early Middle English, respectively.

- (6) V → -long/___CC. (where . is a syllable boundary)
(7) V → -long/___C.

That is, in Old English vowels were shortened in a syllable which was closed by two consonants, while at the later stage only one consonant was required for the shortening to take place. In favor of this formulation of the rules, Stampe (p. 49) adduces counterexamples to Kiparsky's analysis which are in accord with his own (OE hiehsta 'highest', EME rēspondent, where the st and sp are syllabified with the following syllable). Thus the evidence on which Kiparsky bases his argument is not only unlikely, but does not occur. That is, the psychologically real grammar of Old English contained neither the collapsed rule 5 nor the two separate rules 5' and 5'', and similarly for that of Early Middle English.

It is worth pointing out that the evidence in this case is not unambiguous observational data,¹¹ but linguistic analyses which may or may not be correct. It should not be surprising, then, if one investigator finds that he disagrees with another concerning whether or not the 'evidence' cited in an argument actually obtains. What needs to be done in a case like this is to find independent evidence for preferring one of the analyses over the other, as Stampe has done in this case by offering the counterexamples noted above. Something analogous is true, I would maintain, in the case of many major premises.

4.1.3. Botha on Kiparsky's argument. Botha (1973:94-111; 136-66) has also subjected Kiparsky's argument to some criticism, but on grounds quite different from those offered by Stampe. Using the framework developed in

Toulmin (1964), Botha "reconstructs" this argument (pp. 100-1) as in (8).

(8) If the brace convention...is incorporated in the general-linguistic theory, then it follows that the linguistic changes C_1 and C_2 could have occurred as a single unitary change C .
The linguistic changes C_1 and C_2 occurred as a single unitary linguistic change.

∴ The brace convention...should be incorporated in the general-linguistic theory and should be assigned psychological reality.

Botha is unhappy with this argument because (p. 101) "there is a distinct qualitative difference in content between the evidential statement, which as the minor premise refers to anhistorical and diachronic state of affairs and the second half of the conclusion, in which a claim is made about a nonhistorical, nondiachronic mental state of affairs". But, as noted in section 4.1.1, the first and second halves of the conclusion appear to be considered logically equivalent by Kiparsky, and therefore the second half must not be qualitatively different unless the first one is.

There are further problems with Botha's reconstruction, and I would like to give a brief discussion of this issue, in the course of which it should become clear that the entire argument as reconstructed is incoherent unless Kiparsky's views on psychological reality as it relates to what is part of the "general-linguistic theory" are accepted, and, in fact, it seems that any argument from external evidence (in particular, historical change) implies acceptance of quite similar views.

It is somewhat difficult to interpret this reconstruction, since it seems paradoxical to me to speak of more than one linguistic change (here " C_1 and C_2 ") having "occurred as a single unitary change." What Botha may mean by this is that we can use the brace notation to formally express as a "single unitary linguistic change" what is actually two separate ones. But if this is the case, then it is not clear that even the first conjunct of the conclusion is qualitatively relevant to the premises: why should linguistic theory incorporate a device that enables us to write as one change what is in reality two changes?

Let us suppose, then, that what Botha means here is something like 'what otherwise would appear to be separate changes can be expressed as the "single unitary linguistic change" that it in fact is,' since this seems to be the only other alternative for resolving the paradox. This appears to be what both Kiparsky and Stampe are arguing--that there is a single change involved, not two.

But if we accept this, there is a problem concerning the major premise: why should we expect the incorporation of a purely descriptive theoretical device into the "general-linguistic theory" to have any inductive consequences concerning possible linguistic changes? That is, if the brace convention is a purely descriptive device (i.e., if it does not have psychological reality), then the premises are not relevant at all to actual linguistic changes, but only to how we can describe them. Thus, in order to get a reasonable major premise, we must attribute psychological reality to the brace convention--we must be willing to claim that it is more than just a descriptive device. Since Botha's reconstruction is considerably changed by now, I give below in (9) a re-reconstruction on the basis of the preceding discussion.

(9) If the brace convention is psychologically real, then it follows that what otherwise would appear to be separate linguistic changes are in fact a single unitary linguistic change.

What otherwise would appear to be separate linguistic changes are in fact a single unitary linguistic change.

∴ The brace convention is psychologically real.

This, I feel, is somewhat closer to what Kiparsky intended, and it can be seen that Kiparsky's conclusion is the result of a perfectly legitimate (inductive) argument. (Of course, Stampe's counterarguments would still hold, since under his analysis the single change would no longer otherwise appear to be separate ones.) But even given this version, there is no provision in the reconstruction for how strongly the evidence supports the conclusion (recall that the argument is a nondemonstrative one, and therefore does not establish the conclusion), and both Kiparsky and Stampe apparently feel (cf. the latter's granting of the "beauty" of Kiparsky's argument) that, if the premises are granted, the conclusion would be strongly supported. What is missing here is a reference to how likely it would be for the apparently different linguistic changes to in fact be one even if the brace notation did not have psychological reality, a reference which can be readily seen in the Bayesian schema (2), but which would seem to be difficult to incorporate into any reconstruction of the type that Botha gives. That is, Kiparsky's argument is reconstructible much more appropriately in terms of a Bayesian schema than in Botha's terms. Thus, the form of Kiparsky's argument is quite legitimate, although Stampe has given reason to doubt the premises involved.

4.1.4. Discussion. What can be concluded from the above considerations?

It seems to me that Kiparsky is quite correct in his contention that a certain aspect of linguistic theory is supported insofar as it renders an otherwise unlikely phenomenon comprehensible. This is just an instantiation of Bayesian inference, and as such has the same formal characteristics as many other (non-linguistic) scientific inferences; the argument form is a valid one. In this particular case, unfortunately, it seems likely that at least one of the premises involved is not true, and so, of course, the conclusion is not supported by the premises. It is worth pointing out here that, although in this case the determination of the required likelihoods seems relatively straightforward, such determination will in many cases be pretty clearly subjective; indeed, even in this case, a convinced opponent of the 'reality of the syllable' would undoubtedly find the issue much less straightforward than I have implied it to be.

4.2. Hooper-Harris.

4.2.1. Hooper on vocalic alternations in Spanish. Hooper is concerned with certain vowel alternations found in Spanish verb forms, and how these alternations should be treated in a synchronic description of Spanish. The alternations in question are: (i) mid vowel/diphthong (cf. t[e]ndémos/ t[yé]nden 'we/they tend', c[o]cémos/c[wé]cen 'we/they cook'); (ii) high vowel/mid vowel (pedímos/píden 'we/they ask for'); and (iii) high vowel/mid vowel/diphthong (mintiéran 'that we lie' / mentímos/m[yé]nten 'we/they lie'). Following Hudson (1974, 1975), Hooper proposes to treat these alternations as "suppletive" in nature, so that the verb stem for tend,

for example, would be represented in the lexicon as $t\{\overset{e}{y_e}\}nd$. A rule is then formulated which specifies the environments in which each of the alternants is found. The rule which Hooper eventually settles on is given in (10) (Hooper's 29 p. 159).

$$(10) \left\{ \begin{array}{l} ye/we \\ e/o \\ <i/u> \end{array} \right\} \rightarrow \left\{ \begin{array}{l} ye/we / [+stress] \\ e/o </__C_o i> \\ <i/u> \end{array} \right\}$$

What this means is that if a lexical entry contains a $\{\overset{y_e}{e}\}$, then ye appears under stress, while e appears elsewhere (the o/we alternation is precisely analogous in this respect); if there is a $\{\overset{e}{i}\}$ in the lexical entry, then e appears if the following vowel is i and i appears elsewhere; and if

we have a $\left\{ \begin{array}{l} ye \\ e \\ i \end{array} \right\}$, ye is found under stress, e if unstressed and an i

follows, and i otherwise. Hooper's historical argument is an attempt to support this analysis, and thereby Hudson's "suppletive" theory of lexical representation on which it is based. Let us now examine this argument.

These alternations have undergone analogical leveling in some dialects of Spanish, so that only the diphthong appears where formerly there was a mid vowel/diphthong alternation, Hooper states. In addition, the high/mid vowel alternation has been leveled in favor of the high variant, and the mid vowel has been leveled out of the three way alternations (cf. Hooper 1976: 167). Given rule (10), Hooper claims, "the leveling is accounted for by the mere loss of the mid vowel case in each alternation. Subsequent to this loss, verbs such as contar are underlying /kwent-/, verbs such as pedir are underlying /pid-/, and verbs such as mentir are underlying

/m $\{\overset{y_e}{i}\}nt-/"$ (Hooper 1976:168). The rule given in (11) (Hooper's 44) applies to this latter class of verbs.

$$(11) \left\{ \begin{array}{l} ye/we \\ i/u \end{array} \right\} \rightarrow \left\{ \begin{array}{l} ye/we \\ i/u \end{array} \right\} / [+stress]$$

Thus, she concludes, "the analysis involving rule (29) [here, rule (1)], based on Hudson's model, gives a uniform account of all three alternation types."

Hooper goes on to argue that "it is impossible to account for this leveling" if a diacritic analysis such as that of Harris (1969, 1974) is adopted. (In such an analysis, lexical representations contain a high vowel in cases where a high vowel appears as an alternant, and a mid vowel in the other case, along with a diacritic indicating that the vowel is subject to lowering, to diphthongization, or to both, and there are rules which lower high vowels to mid vowels and convert vowels under stress to the corresponding diphthongs if these vowels are marked with an appropriate diacritic.) She continues: "the diacritic representation...implies that through historical simplification, the diacritic will be lost, and the underlying form will replace all other allomorphs." While the developments

in some cases are in accord with this implication, this is not the case with respect to the former mid vowel/diphthong alternations (where there are now only diphthongs), since here the underlying form contained the mid vowel, and not the diphthong in Harris's analysis. Hooper argues further (p. 169) that no other diacritic analysis can account for this leveling, but I will not go into detail about this part of her argument.

Let us now attempt to reconstruct this argument in more explicit fashion. Given only the discussion in the next-to-last paragraph above, one might be tempted to reconstruct this argument as in (12).

(12) If rule (10) (and the associated lexical representations) are correct, then a uniform account of the development of these alternation types can be given.

A uniform account should be given (i.e., the development was uniform).

Therefore rule (1) (etc.) are correct.

It should be noted that the minor premise in this reconstructed argument was never explicitly stated by Hooper. However, it is clearly needed to arrive at even an 'inductive' argument form (cf. the discussion in section 4.1.3 above), and Hooper apparently regarded it as being true.

If we examine the paragraph which follows this one, however, it becomes clear that the overall argument form which Hooper intends is the stonger Bayesian one. A revised reconstruction which takes this into account is given in (12').

(12') Unless rule (10) (etc.) were correct, it would be unlikely that a uniform account of the development of these alternation types could be given.

A uniform account should be given.

Therefore rule (10) (etc.) are correct.

That is, no diacritic analysis can give a uniform account, and such an analysis is the only readily imaginable alternative. This, then, appears to be the form of Hooper's argument; and again the argument form appears to be a legitimate one; an assessment of the premises will be given in section 4.2.3.

4.2.2. Harris's reply. Before proceeding to Harris's reply, let me first sketch briefly his (latest) analysis, since he contrasts it with that of Hooper with respect to predictions about historical change. He maintains the basic diacritic approach outlined in the previous section: some lexical items (those subject to diphthongization) are marked with a diacritic [D], while others (and some which are also marked [D]) are marked with the diacritic [HM] (for 'high-mid'--these lexical items are subject to the high/mid vowel alternations). He has two separate rules which correspond to Hooper's rule (10); these are given in (13) and (14) (Harris's 4 and 10).

(13)
$$\left[\begin{array}{c} +\text{stress} \\ D \end{array} \right] \rightarrow [-\text{syllabic}] \left[\begin{array}{c} -\text{back} \\ -\text{high} \end{array} \right]$$

$$(14) \left[\begin{array}{c} \text{V} \\ -\text{low} \end{array} \right] \rightarrow \left\{ \begin{array}{l} [-\text{high}]/[\overline{\text{HM}}]C_1 \left[\begin{array}{c} +\text{syll} \\ +\text{high} \end{array} \right] \\ [+high] \end{array} \right\} / [\text{X}[\overline{3 \text{ CONJ}}]\text{Y}]_{\text{verb}} \quad \begin{array}{l} \text{(a)} \\ \text{(b)} \end{array}$$

That is, stressed e and o are converted to ye and we, respectively, in lexical items which are marked [D]; and in third-conjugation verbs which are marked [HM], a mid vowel appears when an i follows and a high vowel appears otherwise.

Let us now turn to Harris's counter-arguments. His basic approach is to first question Hooper's facts (cf. Stampe's criticism of Kiparsky in section 4.1.2). He agrees that the mid vowel has been leveled out of those alternations where a high vowel is one of the alternants (i.e., types (ii) and (iii) described in section 4.2.1). But, he goes on (p. 54), "in a number of dialects--overlapping, but not coextensive with those in which the high-mid alternation is lost--the diphthongization alternation has been partially lost. Here the loss is sporadic, affecting particular lexical items at random." The alternations have been leveled here in favor of the monophthong (contrary to Hooper's account), so that, e.g., "standard...qu[yé/e]r- 'want' has become qu[e]r." The leveling of the mid vowel/diphthong alternation in favor of the diphthong mentioned by Hooper does in fact occur in some Chicano dialects, but it is somewhat more complicated than Hooper's account suggests: "this particular leveling has occurred only in first-conjugation verbs with the [wé] ~ [o] alternation... First conjugation verbs with the [yé] ~ [e] alternation, and verbs of the other conjugation with either alternation are not affected."

This requires that the account of the historical changes, assuming Hooper's analysis, be somewhat different. In particular, Harris claims, "the loss of the high-mid alternation consists of three seemingly unrelated changes: (a) loss of mid vowels in individual lexical disjunctions, (b) loss of the environment of rule 17 [here, rule (10)--DGC], and (c) loss of the second case of rule 19."¹² That is, contrary to Hooper's claims, her analysis does not give "a uniform account of" even just the high/mid part of the alternations. Furthermore, Harris argues, "nothing in this account predicts loss of the mid vowel in all forms, rather than the high vowel in some and the mid in others, or the high vowel in all forms."

Concerning the sporadic leveling of the diphthongization alternation mentioned above, Harris has little to say--he apparently feels that it is unproblematic for either his or Hooper's analysis (cf. pp. 54, 56). He does, however, note (p. 55) that "nothing in Hooper's account reflects the fact that diphthongization is lost in one particular morphological subclass, as described above," in the case of the leveling which affected only the o/we alternation in first-conjugation verbs.

Harris also argues that on his analysis, "loss of this alternation implies loss of all and only the machinery associated with just this alternation, namely rules 14¹³ and 10a [here (14a)] and, obviously, the triggering feature [HM]. Nothing else changes. In particular, rule 10b [i.e., (14b)] remains in effect, guaranteeing that the only possible result of the loss of the alternation is precisely the survival of high vowels" (pp. 53-54). In the case of the sporadic leveling of diphthongization, "all that changes is that individual lexical items lose their exceptional diphthongization-marking property...This behavior in exceptional forms

is unremarkable" (p. 54). As for the levelings in favor of the diphthongs, Harris claims (p. 54) that "this means that, in addition to the general process of diphthongization (which remains unaltered), the grammar of these innovative dialects must contain a special statement referring to the absence of first-conjugation stems with the 'back' branch of diphthongization" (he refers the reader in a footnote to Harris (1974) for a formulation of such a "special statement").

Harris goes on to present some data not discussed by Hooper. Some speakers of the Chicano dialects just referred to, presumably due to influence from the mass media, schooling, and recent immigrants who speak more standard varieties of Spanish, have apparently reestablished the high/mid vowel alternation in a systematically characterizable way: "all diphthongizing third-conjugation verbs have high and mid stem-vowel variants; no non-diphthongizing verb does" (p. 56). That is, assuming that these dialects are more innovative than the other Chicano dialects which have leveled out the mid vowel variant completely, the only verbs which have reestablished the high/mid alternation are those third-conjugation verbs which in Harris's analysis were previously (and still are) marked with the diacritic [D]. Such facts, Harris maintains (p. 56), "are beyond the reach of Hooper's analysis, which treats all of the innovations in question as collections of accidents, rather than systematic changes."

Let us now summarize Harris's counter-arguments, and cast them in a more explicit reconstructed form. First of all, he claims that Hooper's analysis does not offer a uniform account of the historical changes with respect to the high-mid alternations, (i.e., that the major premises in (12) and (12') are false) and that (assuming that "all and only the machinery associated with just this alternation"--see above--is meant to be roughly equivalent to "uniform") his analysis does. This appears straightforward, and I will not attempt to explicate it further. His second point is that Hooper's analysis does not predict loss of the mid vowel in the high/mid alternations, while his does. This argument, as it stands, is of the form given in (15).¹⁴

- (15) Analysis A_1 makes prediction P.
Analysis A_2 does not make prediction P.
P is found to be true.
Therefore, A_1 is to be preferred to A_2 .

It is of some interest to note that this argument, if strengthened somewhat, takes on a Bayesian appearance. Such a strengthened version is given in (15').

- (15') Unless Harris's analysis (A_1) were correct, the facts P would be unlikely.
P is true.
Therefore, A_1 is likely to be correct.

Harris apparently does not want to make this stronger claim, however, presumably preferring simply to establish the superiority of his analysis over Hooper's. It should be noted that in so doing Harris is making no claims as to the ultimate correctness of his analysis, but only that it is to be preferred to that of Hooper. His final argument appears to be that his analysis allows one to specify which verbs reintroduce the high/

mid alternation, while Hooper's does not. This argument form is sufficiently similar to that in (15) that I see no need to reconstruct it here (but note that "prediction" will have to be replaced by something like "generalization").¹⁵

4.3.2. Discussion. As noted above, Harris has called into question the accuracy of the data on which Hooper bases her argument. That is, the "development of these alternation types" referred to in (12) and (12') above does not correspond to their actual historical development. If the facts are indeed as Harris and his sources (which are the same, for the most part, as those of Hooper) describe them, then Hooper's argument is incoherent, since an implicit part of the major premise is not true. Even if the facts were exactly as Hooper has described them, however, there appears to be a problem for her, since Hudson's suppletion theory "proposes that the direction of leveling is always toward the form designated as the elsewhere case [i.e., the one without a conditioning environment--DGC] of the distribution rule..." (Hooper (1976:129)). This is not, on her account, what has happened in the diphthongization case--the levelin has been toward the diphthong, which is not the elsewhere case in any expansion of the "distribution rule" (10). Thus, either this case (and the similar case of the leveling out of the mid vowel in the three-way alternations) is a counterexample to this proposal and it must be modified, or the suppletion theory of lexical representations has been falsified. Note also that attempting to revise the suppletion analysis somewhat so as to make the diphthongs the elsewhere case would result in incorrect predictions concerning the sporadic leveling of the diphthongization alternations, where the diphthongs are leveled out and mid vowels remain. In any event, the fact that the diphthongization alternation levels out in different dialects in different directions indicates that the suppletion theory (or any other theory, undoubtedly) must give up its claims for predictive power concerning the direction of leveling.¹⁶

Let us turn now to Harris's arguments. First of all, does Hooper's analysis in fact make the loss of the high/mid alternation appear to consist of "three seemingly unrelated changes"? It seems to me that it does not; once the mid vowel in each "individual lexical disjunction" has been lost in the high/mid cases, the other changes follow automatically. That is, the fact that the environment of the second part of rule (10) is missing is an automatic consequence of the lack of lexical items containing both a mid and a high vowel in the braces of a suppletive lexical representation; there could never be an opportunity for this environment to be met, and so it is not needed. Similarly, the loss of the second part of Harris's rule 19 (see note 12) is the automatic consequence of there no longer being any mid vowels in the lexical representations of third-conjugation verbs. These changes do not appear to me to be at all unrelated. Moreover, Harris's analysis seems to be saying pretty much the same thing as Hooper's: once the "triggering feature [HM]" is lost, there is nothing for rule (5a) (his 10a) to apply to. Harris is not quite correct that his rule 14 has been lost, however. In actuality, it must have changed to something like (16).

$$(16) \text{ if } \begin{bmatrix} 3 \text{ conj} \\ D \end{bmatrix}, \text{ then } \#X \begin{bmatrix} +\text{syll} \\ +\text{high} \end{bmatrix} C_o \#$$

That is, all diphthongizing third-conjugation verbs have a high vowel. Concomitantly, Harris's rule 13 (p. 46), given below in (17), must be modified to something like (18).

(17) If [D], then $\begin{bmatrix} +\text{syll} \\ -\text{high} \\ -\text{low} \end{bmatrix}$

(18) If $\begin{bmatrix} +1 \text{ conj} \\ D \end{bmatrix}$, then $\begin{bmatrix} +\text{syll} \\ -\text{high} \\ -\text{low} \end{bmatrix}$

That is, it is no longer true that all diphthongizing verbs have mid vowels--now, only first-conjugation verbs do. These changes again appear to be necessary consequences of the across-the-board loss of the diacritic [HM]. I can see little difference between the two analyses in this respect. It should be pointed out, however, that Hooper's analysis contains no analog of rule (17), and, it appears to me, it would not be at all easy to formulate a redundancy statement of the form which Hooper employs to express this fact. That is, it is apparently very difficult, if not impossible, to express in the notation used by Hooper the fact that whenever there is an alternation involving a diphthong in standard Spanish, there is also a mid vowel involved. (This is related only somewhat marginally to the arguments from historical change, but since rule (17) must be replaced by rule (18) for Harris, it cannot be completely ignored.) Thus, in sum, Harris's claim that Hooper's analysis does not provide a uniform account of the high/mid vowel alternations appears to be false.

The second issue to be discussed in this context is whether or not Hooper's analysis predicts the direction of leveling in the high/mid alternations. If we accept the principle that leveling favors the elsewhere variant (this principle has of course been shown above to be very difficult to maintain, however), then it seems clear that it does, since the high vowel is the elsewhere variant and it is the one that 'wins out' in the leveling. The fact that the principle which makes the prediction is probably incorrect should not be regarded too highly, by the way, since the corresponding principle for the diacritic analysis (that the underlying form shows up in cases of leveling) also appears to be incorrect in view of the data under discussion, since there are dialects which level out the diphthongization alternation in favor of the diphthong instead of the underlying monophthong. Thus, the second premise in (15) appears not to be true.

As for Harris's third argument, it is not at all clear to me that it is in fact the case that there is no way to specify which verbs reintroduce the high/mid alternation. In particular, just as Harris can state that precisely those third-conjugation verbs which are marked [D] reintroduce the alternation, it seems that Hooper can just as easily state that all third-conjugation verbs which contain a stem 'vowel' roughly of the form $\left. \begin{array}{l} \text{diphthong} \\ \text{high vowel} \end{array} \right\}$ reintroduce this alternation. Again, the premise that Hooper's analysis cannot state this generalization appears to be false.

To sum up, when the data have been straightened out and the premises of the arguments are examined with these data in mind, the historical changes which have affected Spanish vowel alternations appear to shed little light concerning a choice between a Hooper-type suppletion analysis and a Harrisian diacritic analysis. Some general implications of these data in particular will be discussed below (see section 5.)

4.3. Schuh-Leben

4.3.1. Schuh on 'rule inversion in Chadic'.¹⁷ Schuh first presents data which indicate that in the Chadic language Kanakuru earlier stops "weakened to corresponding sonorants in phonologically specifiable environments," so that *t, *d, *ɗ became r, *k and *g became y, and *p and *b and *ɓ became w. These weakenings resulted in many cases in synchronic alternations between stops and sonorants (e.g., yilik 'tongue' vs. yiliy-no 'my tongue'). But the synchronic rule, Schuh maintains, does not mirror the diachronic change; rather, "the rules producing the alternations...are 'hardening' rules and the sonorant variants are underlying" (p. 384).

In support of his contention about the rule having been inverted, Schuh offers three arguments: (a) in synchronic alternations, the sonorants currently alternate only with voiceless stops, regardless of their etymological source, and, furthermore, the sonorant/voiceless stop alternation has been extended even to etymological sonorants; (b) it would otherwise be impossible to distinguish stop-sonorant sequences which are subject to a rule of schwa-epenthesis from those which are not--only underlying stops (in Schuh's analysis) trigger epenthesis; and (c) plurals of certain nouns and verbs, which formerly contained various (etymological) stops, have been regularized so that they contain only p, ɗ and k (the singulars contain the corresponding sonorants); also, some singulars have alternate forms for the plural, one showing "hardening" and the other the sonorant found in the singular (cf. Schuh 1972:286-9).

Let us now attempt to reconstruct these arguments in more explicit fashion. Schuh's actual wording of argument (a) (p. 386) is repeated below in (19). The argument, in somewhat more explicit terms, can thus apparently be taken to be representable as in (20).

(19) ...regarding the contemporary Kanakuru consonant alternations as an inverse rule where sonorant → stop rather than the historical process stop → sonorant explains the regularization of the alternations giving only voiceless stops as alternates of sonorants. It also explains why etymological sonorants now alternate with stops.

(20) If the alternations are regarded as being due to an inverse rule, then the regularizations are explained. Therefore, the alternations are due to an inverse rule.

As stated, there appears to be something missing in this argument. For one thing, it seems clear that merely regarding the alternations as being due to an inverse rule cannot possibly play a part in an explanation of the regularizations (or anything else, for that matter)--these alternations must in fact be due to an inverse rule. It therefore seems desirable to amend the premise in (20) so as to delete the "regarded as being" part. I am reasonably confident that Schuh would have no objection to such an amendment; he was, after all, arguing that the rule inversion analysis is correct. The second thing which appears to be missing from this argument is a minor premise. Since it is easy to add a premise which would probably seem obvious to all linguists, and which would make the argument have roughly

the form of the inductive arguments discussed by Botha, it appears reasonable to add such a premise; the revised form of (20), including the revision in the major premise, is therefore given as (20').

- (20') If the alternations are due to an inverse rule, then the regularizations are explained.
The regularizations should (must) be explained.
Therefore, the alternations are due to an inverse rule.

Since the argument as given in (20') is inductive in form, it is subject to all the weaknesses of such argument types. In particular, it is quite possible that something other than the rule's being inverted could explain the regularizations. It is therefore possible that Schuh intended his argument to have a stronger Bayesian form, although it is difficult to tell since there are no connectives in Schuh's actual argument (I supplied the if in the reconstructions). To allow for this possibility, I give a Bayesian version of (20') in (21).

- (21) Unless the alternations were due to an inverse rule, it is unlikely that the regularizations could be explained.
The regularizations should (must) be explained.
Therefore, it is likely that the alternations are due to an inverse rule.

Let us turn now to Schuh's second argument. He presents (p. 386) the data given in (22).

- (22) a. a wupə-ro 'he sold (it) to her' [cf. wupe 'to sell']
b. a gup-ro diyii 'he forged a hoe for her' [cf. guwi 'to forge']
c. ʃi kuke-mai 'he is learning it' [cf. kuke 'to learn']
d. si duŋ-ŋai 'he is beating it' [cf. duyi 'to beat']
[cf. also a duk-ro 'he beat (it) for her']

Schuh argues (pp. 386-7) that "if we were to take stops as underlying in all cases and derive the sonorants from them, there would be no way to distinguish the medial consonant in the verb root in [22a] from that in [22b] and the medial consonant in the verb root in [22c] from that in [22d] for the purposes of epenthetic ə insertion...Likewise, by not distinguishing k and ɣ underlyingly, we would have no way to predict which words have velars which assimilate to a following nasal, as in [22d]." What Schuh apparently means by this is that the underlying form of the verb stem for the above forms is the same as the surface form of the infinitive; after a rule which deletes final -i in verbs everywhere except pre-pausally has applied, the epenthesis rule mentioned earlier applies, breaking up stop-sonorant clusters,¹⁸ and then a rule which assimilates velars to a following nasal and the inverted version of weakening (which applies in the complement of the former weakening environments) apply to produce the stops in the left column of (22). Sample derivations for (22a, b) are given in (23). This argument thus seems to be reconstructible in the modus tollens form given in (24).

(23)	/wupe-ro/	/guwi-ro/
<u>i</u> -deletion		guw -ro
epenthesis	<u>wupe-ro</u> ¹⁸	
(assimilation)		<u>gup -ro</u>
strengthening		

- (24) If stops are underlying in all cases, there is no way to distinguish the required consonants with respect to epenthesis.
 These consonants should (must) be distinguished.
 Therefore, stops cannot be underlying in all cases.

The minor premise here is again only implicit in Schuh's actual words, but is so innocuous looking that it can reasonably be supplied here.

Let us now proceed to Schuh's third argument. As noted earlier, it has two parts: "first, the stops found in the plural [which have not undergone weakening] do not always reflect their etymologies....In fact, the alternations are always w/p, r/d and y/k" (pp. 388-9), with the exception of the verb 'to die', where r alternates with t. Secondly, "plural hardening ...involves an alternation which is both phonetically unmotivated and requires arbitrary marking of those lexical items which undergo it," which leads it to be replaced by "more regular processes which do the same semantic or syntactic work" (p. 389). In this case, the sonorants show up in the plural. Examples given by Schuh to support these claims include those in (25) (cf. pp. 387, 389--ngin is the most common and productive plural suffix).

- (25) 'hen' yaawe (sg.) yaapiyen/yaawingin (pl.) (cf. Tangale yabe)
 'gazelle' sere (sg.) sediyen/serengin (pl.) (cf. Dira kite)

It is not clear exactly what form these arguments should take here, since all Schuh does is mention the facts. It seems, however, that he intends something like (26), in which the arguments, of course, take a Bayesian form.

- (26) a. Unless the rule has been inverted, it would be unlikely that the (non-weakened) stops in the plurals would not reflect their etymologies and rather have a single stop for any sonorant in the singulars.
 They do not reflect their etymologies, etc.
 Therefore, it is likely that the rule has been inverted.

- b. Unless the rule has been inverted, it would be unlikely that the plural rules would be replaced by a more regular process in which the singular sonorant shows up in the plural.
 These rules have been so replaced.
 Therefore, it is likely that the rule has been inverted.

There is some evidence in Schuh's comment on Leben's reply that he did in fact intend a Bayesian argument at least for (26a), in that he argues (p. 289) that having an underlying stop and a rule of weakening in singulars "totally ducks the issue of why the stops in the plural have all been neutralized, and moreover why even historical sonorants now alternate with

stops..." That is, unless there had been an inversion, such changes would have been unlikely.

Schuh gives one final argument concerning historical change, this time in Hausa, another Chadic language. He first argues for a series of changes often referred to as 'Klingenheben's Law' (since these changes were first systematically described in Klingenheben 1928) on the basis of "synchronic alternates, dialect variants, and comparative evidence" (p. 391), whereby original velar stops became w, alveolar obstruents became ř, and labial stops, including etymological m, became w, all syllable-finally; there is again little question about whether this is an accurate description of the history of Hausa.

Schuh then proceeds to argue that this process has taken an inverted form from a synchronic perspective. One bit of evidence for inversion, Schuh argues, has to do with the formation of plurals. "The language is losing those plural forms where obstruents alternate with sonorants" (p. 393). The innovative plurals contain the sonorant which is found in the singular. Thus we find forms such as those in (27),¹⁹ where the singular w is the result of Klingenheben's Law, and the original plural is formed by infixing -aa- and adding a suffixal -ee or -aa to the noun stem (t becomes c ([č]) and z becomes j ([ǰ]) before front vowels by a general process of the language).

(27)	<u>Singular</u>	<u>Plural</u>
'buffalo'	Ḅawnaa	Ḅakaanee/Ḅawnaayee
'heart'	zuwciyaa	zukataa/zuwciyooyii
'Tuareg'	buwzuu	bugajee/buwzaayee

Moreover, Schuh continues (p. 394), there is further evidence "in the word gwauroo 'bachelor' which has the plural alternates gwauraayee and gwagwaaree. This second alternate has to be an analogical reformation resulting from the neutralization of *P and *K in syllable final position. The -u- in gwauroo comes from *P, not *g, as can be seen in the dialect variants gwabroo or gwamroo". Schuh again is not very specific about the actual form of his argument, presumably feeling that it should be clear from the form of his previous arguments. However, given the considerable lack of clarity concerning the precise form of his previous arguments, it is not at all clear whether this one should be reconstructed in an inductive, Bayesian, or modus tollens form, although it seems safe to assume that it is meant to be of one of these types. I give a Bayesian version in (28), mainly because, as will be seen in the next section, Leben appears to be construing it in this way.

- (28) Unless Klingenheben's Law has been inverted, it would be unlikely that the regularization of plurals and the analogical reformation of the plural of gwauroo would occur. They do occur.
Therefore, it is likely that Klingenheben's Law has been inverted.

4.3.2. Leben's reply. Leben argues (p. 265) that "Schuh's evidence does not lead to the intended conclusion." He goes on (pp. 265-6) to claim

that "the positing of a synchronic stage with...inverse rules constitutes a middleman which it would be advantageous to eliminate in principle from the realm of possible phonological systems". In order to show that it is possible to eliminate inverse rules, of course, Leben must counter each of Schuh's arguments, and he proceeds to attempt to do so.

Concerning Schuh's first argument, Leben argues (p. 267) that "if we do not assume that sonorants became basic, it is still possible to explain the historical developments." Recall that what is to be explained is the fact that sonorants now alternate only with voiceless stops. Leben proposes an alternative explanation for this: "In the examples given by Schuh, the voiceless stops resulting from this regularization appeared in a typical devoicing environment, immediately preceding a voiceless stop.... If, in addition, etymological d, b, etc., ceased to surface phonetically as voiced stops, then future generations would be presented with no synchronic evidence for setting up underlying voiced stops in these words." Thus, the voiceless stops could become underlying for this reason in the case of etymological stops. As for the historical sonorants which now alternate with stops, Leben notes (p. 268) that "the only instances he cites of the extension of the alternation to historical sonorants occur in word final position, and Schuh himself notes (p. 386) that 'word final is a position of neutralization where stops and sonorants cannot contrast either phonetically or underlying [sic].'" The sonorants have been eliminated by this neutralization rule in favor of voiceless stops, which "will naturally be subject to the same alternations as any other instance" of voiceless stops. What Leben has done, then, is to argue that there is another possible explanation of the regularizations, i.e., that the if-clause in (20') can be replaced by something else (which would make this inductive argument unconvincing) or the major premise in (21) is false.

Similarly, Leben argues (p. 270) that it is possible to account for the varying susceptibility of stop-sonorant clusters to epenthesis without an inverse rule. In particular, he makes use of the same rules as those mentioned in connection with Schuh's argument, except that a rule of weakening, which mirrors the historical process, replaces Schuh's strengthening. This process, it should be noted, did not affect stops which were preceded by a short vowel and followed by e, or stops which were followed by e. These rules result in derivations such as those in (29).

(29)	/wupe-ro/	/gupi-ro/
<u>i</u> -deletion and epenthesis	wupə-ro	gup-ro
(assimilation)	-----	-----
weakening	-----	-----

The isolation form of /gupi/ weakens to guwi, but that of wupe (as in all other forms with stops between a short vowel and e) does not. This completes Leben's argument that the major premise in (24) is false.

Leben does not have much to say about Schuh's third argument, apparently feeling that inverted rules which apply only "in a small subset of nouns and in a small class of verbs that form plural stems" (p. 270) need not be eliminated from the class of possible rules. He does note, however, that he sees "no good reason for assuming that [the process at issue] did involve Weakening in singulars," and argues that "even if it were

shown that this morphological rule had become inverted, the case for Schuh's other inverse rules would not become any more plausible. For one thing, this morphological rule converts r into d, and thus it does not reinforce Schuh's earlier proposal of a rule to convert r into t." It thus appears that Leben is willing to concede that this could be a case of rule inversion, but argues that since the inversion has happened to a morphological rule, it is not of the same type as Schuh's earlier examples, which presumably involved phonological rules, and therefore is irrelevant to establishing whether or not it is possible for the latter type of rule to become inverted.

Concerning the Hausa example, Leben first argues (p. 274) that Schuh's "proposed solution seems unnecessary, since the regularized plurals are functioning to reduce allomorphy in the singular-plural paradigm." He goes on to point out that "there are perhaps over a dozen different ways of forming plurals in Hausa; a noun may take a number of different plurals, all with the same meaning. A given noun or adjective must be marked for which way or ways its plural is formed". Then he argues that "looking at the correspondence between the singular and the old plural and comparing this to the correspondence established with the regularized plural, it is hardly surprising that the regularized plurals should be gaining ground, to the detriment of the older forms." That is, he is apparently questioning the major premise in (28).

He goes on (p. 275) to "examine the question of whether rule inversion was even possible in the cases proposed by Schuh." His position is that it was not, since having a stage in the history of Hausa with inverse rules "entails eliminating an otherwise valid constraint on plural formation" (that stems which end in a consonant cluster have -aa- inserted between these consonants, while stems with a final glide-consonant sequence have -aa- inserted after the consonant in the formation of plurals) at this stage, while (p. 276) once this stage "began to be overcome by the regularization of the plural forms..., Hausa went back to the old restriction on -aa- Insertion..." "This scenario," says Leben (p. 277), "...is totally unacceptable." Moreover (p. 276), "the putative relaxation did not have any effect on the derivation of plurals that had pre-existing stems ending in a glide-consonant sequence" (that is, they continue to form their plurals by inserting the -aa- after the stem final consonant). This argument thus appears to be of a modus tollens type, as indicated in (30).

- (30) If Kl ingenheben's Law has been inverted, then the restriction on -aa- insertion was eliminated and then re-introduced, and the elimination of this constraint had no effect on pre-existing stems endings in glide-consonant. This did not happen (i.e., "is totally unacceptable"). Therefore, Kl ingenheben's Law has not been inverted.

Leben goes on to propose his own analysis of the development of Hausa plurals,²⁰ one which does not require an inverted rule, but rather makes use of "competing underlying forms" (p. 277). This proposal entails, for example, that "the existence of the covariants ɓakaanee and ɓawnaayee in Hausa simply constitutes evidence for two competing underlying forms /ɓakn-/ and /ɓawn-/." Thus, there is an analysis other than Schuh's which is compatible with the regularizations observed to have occurred; i.e.,

Leben claims that the major premise of (28) is false. Finally, Leben suggests that his analysis, but not Schuh's, gives a possible explanation of why faŕkee 'trader' has only fataakee as a plural, and has not regularized. It has to do with the existence of "the derived form fatawcii 'trading' (derived by an unproductive process), where w comes from k in /fatk-/ by Klingenheber's Law. Therefore, restructuring of /fatk-/ as /faŕk-/, though it would succeed in reducing allomorphy in the singular-plural paradigm, would at the same time obscure the relationship of fatawcii to its root" (p. 278), while there is no corresponding form in the cases which are undergoing regularization. This appears to be a rather minor point in Leben's discussion, and so I will not attempt to make the argument form more explicit.

4.3.3. Discussion. Concerning the regularizations of the alternations in Kanakuru so that there is only a voiceless stop which, regardless of its etymological source, alternates with the sonorants which are the result of the historical weakening process (as well as etymological sonorants), there can be little question that Leben's account is superficially at least as plausible as that of Schuh. Schuh is apparently in agreement with this assessment, since he states in his comment on Leben's reply (p. 279) that "Leben has registered a number of valid criticisms of my analyses..." and does not explicitly comment on the issue in question. Thus, Leben has apparently succeeded in showing, at least to Schuh's satisfaction, that the if-clause in (20') is not the only possible means of providing an explanation for the regularizations, or, alternatively, that the major premise in (21) is false. It is possible, of course, that there could be data from Kanakuru which would be counterexamples to Leben's analysis-- there could be nonetymological voiceless stops which alternate with sonorants, and which are not in a devoicing environment, or there could be etymological sonorants which alternate with voiceless stops which are in other than word-final position. Schuh is apparently not aware of any evidence of the type just mentioned, since he does not bring it up in his comment, but this is clearly an empirical question. It should be noted in this regard that if there are in fact no data of this kind, Leben's account would appear to be supported, since the lack of stops occurring in other kinds of environments would appear a priori to be quite unlikely, unless perhaps some facts about the structure of Kanakuru preclude such data.

However, such data clearly do exist, even among the examples discussed by Leben, although he is correct in his statement that the examples cited by Schuh in his first argument contain no data of this type. Thus, we find guwi and a gup-ro diyii illustrating the w/p alternation (cf. Schuh, p. 385 and Leben, p. 268), where p is clearly neither in a devoicing environment nor in word-final position. Thus, since the alternating consonants come from etymological *ḡ (cf. Schuh, p. 385), Leben's analysis cannot be maintained. He thus has shown neither that there is an alternative explanation for Schuh's facts nor, alternatively, that the major premise of (21) is false (i.e., he has not explained the disappearance of other stops from the stop-sonorant alternations).

The question of epenthesis is likewise in principle an empirical one, since Schuh's and Leben's analyses, though both generate the forms discussed by Schuh, make different predictions about the behavior of other possible

forms. Leben's analysis essentially claims that all verbs with stem-final -e or with a final consonant exhibit 'epenthesis,' while Schuh's predicts that verbs with etymological sonorants (which now alternate with stops due to the analogical leveling mentioned by Schuh) will not. That is, it is possible that, in addition to verbs with non-alternating stops (presumably (22a) is an example of this type) due to the inhibiting effect of the preceding short vowel and the following -e, there are verbs with etymological sonorants which alternate with stops. Leben's analysis predicts 'epenthesis', since the e would be between a stop and a sonorant at the point in the derivation at which this rule is applicable, but Schuh's analysis predicts none since what is an underlying stop for Leben is an underlying sonorant for Schuh in such cases. The two analyses also make different predictions concerning the behavior of verb stems in final -e, but with a long vowel preceding the pre -e consonant. These considerations show that there are potential differences between the two analyses even at the level of observational adequacy (one of them must generate incorrect surface forms), although both offer an account of the data at hand, with respect to the epenthesis facts. Choosing between the two analyses thus depends solely (or at least primarily) on synchronic data, and not on historical evidence in this case; both appear to be somewhat satisfactory accounts of the historical data presented by Schuh in his first two arguments.

Concerning the question of which account of epenthesis is in accord with the synchronic facts of Kanakuru, Newman (1974) gives some discussion which indicates that the facts are closer to being the way Schuh has described them, rather than as in Leben's analysis. First of all, it should be noted, the schwa-epenthesis rule is actually somewhat different than in either Schuh's or Leben's formulation--sequences of two consonants are subject to epenthesis if the first is voiceless or prenasalized, and the second may be any consonant, not just a sonorant as stated by Schuh. The sequence dr is also subject to epenthesis, as are all triconsonantal clusters (cf. p. 3). The exact statement of the epenthesis rule is not crucial to the present concern, however, and I will not pursue this matter further here. What seems most relevant in this respect is that Newman specifically states (p. 4) that "the invariant voiceless stops [i.e., those never in a weakening environment]...are...subject to [epenthetic schwa insertion], while the still unspecified archiphonemes are not." (Archiphonemes are used by Newman to represent the alternating consonants.) Thus, etymological sonorants (archiphonemes for Newman) which alternate with voiceless stops will not exhibit epenthesis even in words with final -e, contrary to Leben's analysis, but in accord with that of Schuh. Moreover, data presented elsewhere by Newman indicate that this is in fact the case, e.g., adowe 'he tied it' (where the w is presumably an etymological w, since it is in an environment which prevented weakening), but a dop-t²¹eru 'he (went and) tied it' (p. 9). Thus, Leben's account apparently cannot be maintained in this case, either.

As for Schuh's third argument (that concerning "plural hardening"), of course, Leben offers little objection to Schuh's analysis, and there is therefore not much to be said about it. The little that he does have to say (p. 270)--that he sees "little reason for assuming that [plural hardening] did involve Weakening in the singulars"--seems to be clearly

off the mark, since as Schuh (1974:280) points out, he did supply "cognate items from other languages" which suggest that there was indeed weakening in the singulars. Of course, Leben is not terribly concerned with this issue, since he apparently feels that inverted morphological rules are not of the type "which it would be advantageous to eliminate in principle from the realm of possible" (pp. 265-6) grammars of natural languages. And Schuh states (correctly, it seems to me) that Leben "is correct in noting that the plural rules...do not make rule inversion more plausible for the cases discussed earlier in the article..." (p. 290).

There is considerably more which can be said about the Hausa argument. First of all, it does indeed appear to me, at least, that "it is hardly surprising that the regularized plurals should be gaining ground," given the previous state of affairs which Leben describes. Thus, Schuh's major premise in (28) seems to be false. (This argument would not be any more successful if it had been reconstructed in one of the other forms, of course.)

On the other hand, Leben's contention that there could not be any inversion in the Hausa case seems to be on much less firm ground, and, of course, merely showing that the major premise in (28) is not true does not suffice to show that Schuh's analysis is incorrect. It is therefore of some interest to pursue this issue further, especially since, judging from Leben (1979), he feels that he has successfully shown that Schuh's analysis is an impossible one. Recall that Leben's claim concerning this issue is essentially that Schuh's analysis entails the loss of the condition of -aa- insertion at one stage in the history of Hausa and its reintroduction at a later stage. The question now is whether this sequence of events is really as unbelievable as Leben claims it is. On first glance, it does indeed seem that the reintroduction of a constraint of precisely the same form as one which has recently been lost from the language would be an extremely unlikely event. But if we look more closely at Schuh's account, it can be seen that this account appears to give an automatic explanation for the sequence of events in question.

Let us assume, then, following Schuh, that Klingenberg's Law has been inverted and that underlying forms have been appropriately restructured. Thus, the underlying form of the stem for ḡawnaa, for example, is /ḡawn-/. At the first stage of rule inversion, where the plural is still ḡakaanee, there must be a relaxation of the constraint on -aa- insertion, since it now must break up the glide-consonant sequence w-n, whereas before only consonant-consonant sequences could be broken up in this way. The derivation of ḡakaanee would thus include something like ḡawn- → ḡaw-aa-n- → ḡak-aa-n-, where the last change is the result of the inverse rules, at this stage. Leben's first question (p. 276) is why "the putative relaxation of the condition...did not have any effect on the derivation of plurals that had pre-existing stems ending in a glide-consonant sequence." That is, such forms continue to show -aa- after the stem-final glide-consonant sequence. The answer to this question seems to be simply that allowing -aa- to break up the glide-consonant sequences in such cases would result in an incorrect plural form; if the language learner is to have 'correct' plural forms, he must learn which glide-consonant sequences are broken up by -aa- and which add it after the final consonant, and those which fall into the latter category are of course precisely those which "had pre-existing stems ending in a glide-consonant sequence." Moreover, if a language learner were to 'make a mistake' in the direction of allowing older glide-consonant sequences to be broken up, the result would be an increase in allomorphy, since the glide would thereby be put in an environment which

triggers the inverse rules. Since, as Leben would apparently agree (p. 274), we expect changes of this type to reduce allomorphy, it seems not at all surprising that the older forms should retain the allomorphy-minimizing manner of -aa- insertion.

Can there be an equally plausible explanation for the reintroduction of the constraint on -aa- insertion? (This constraint would now be operative only in the speech of those who had completely regularized the plural system.) It seems to me that this reintroduction, given Schuh's analysis, is the straightforward consequence of the complete regularization of plural formation (note that the constraint cannot be operative in the speech of those who retain the older plurals for some of the singulars which were affected by Klingenheben's Law), which, as noted above, has a clear motivation--the reduction of allomorphy. That is, the result of the complete elimination of this type of allomorphy is that there happen to be no longer any cases where -aa- breaks up a glide-consonant sequence. Thus, the "generalization about -aa- insertion" was not, as Leben puts it (p. 277), "rediscovered," but simply reintroduced by a perfectly natural historical process, the reduction of allomorphy in the singular-plural paradigm. These considerations constitute essentially an argument to the effect that the minor premise in (30) is false.²²

Since, as noted above, Leben has given an alternative analysis which does not make the regularizations in question appear unlikely,²³ it seems that in this case, as in the Spanish case, the data from historical change have little to say about the form of the synchronic grammar of the language in question before the changes. It is perhaps worth noting, however, that Leben and Schuh agree that the grammar of Hausa does not contain underlying forms which were exactly like those in the grammar before Klingenheben's Law took place; at least one positive conclusion about the form of the grammar of Hausa with respect to the underlying forms contained there, that not all of them are the same as before the changes occurred, can thus apparently be made.

4.4. Skousen-Kiparsky.

4.4.1. Skousen on Finnish. Skousen's basic point is that many of the phonological rules which have been posited by generative phonologists to account for morphological alternations in Finnish are not part of a descriptively adequate grammar of Finnish; many of his arguments in support of this position invoke data from historical change, and it is these arguments which will be examined here. I will focus on Skousen 1972, since Kiparsky's reply is directed solely toward this work (he apparently did not have access to Skousen's 1972 dissertation, on which Skousen 1975 is based, when he wrote his reply), although reference will also be made to Skousen 1975 when it can be of help in clarifying the issues involved.

Before giving the arguments against the validity of the rules mentioned above, Skousen gives (p. 569) "some substantive evidence for a phonetically-plausible rule that speakers do capture," and I feel that it is worthwhile to take a brief look at this discussion before proceeding to his arguments against phonetically-plausible rules. The rule in question, a fairly old one found in Savo dialects, geminates a consonant "when it is preceded by a short, stressed syllable and followed by a long vowel or diphthong." Skousen's formalization of this rule is given in (31).

(31) $C \rightarrow C: / \bar{V} _ _ VV$

Thus, for example, older téköö 'he does' has become tékköö in these dialects. Evidence that this rule (in this form) is still productive includes the fact that recent loan words undergo it, and that "more recent phonetic rules ... have set up surface exceptions to the rule of gemination," but "in every case, the rule of gemination eliminates these exceptions." For example, in some of the Savo dialects a word-final sequence Vns has become VVs, so that a word like vákens 'his people' (cf. standard Finnish väkensä) has become väkees in these dialects. "A surface exception to the rule of gemination has been created in the Savo dialects, but the rule of gemination has applied to give väkkees..." (p. 570). Thus, Skousen goes on, "there are phonetically-plausible regularities that speakers can capture." The form of this argument thus appears to be that illustrated by the reconstruction given in (32), which is more or less the classical *modus ponens* form.

- (32) If a (phonetically-plausible) rule applies to the output of a rule which enters the language after the rule in question, then it has been captured by speakers. Gemination applies to the output of a later rule. Therefore, gemination has been captured by speakers.

Skousen's first example of a "phonetically-plausible" rule which speakers do not capture concerns the well-known phenomenon of consonant gradation in Finnish, for which most generative phonologists have posited a rule which 'weakens' stops at the beginning of a closed syllable. The alternations to be accounted for (cf. Skousen 1972:571) are exemplified in (33).

(33)	p/v	tapa/tavan	'custom'
	t/d	pato/padon	'dam'
	k/φ	sika/sian	'pig'
	pp/p	piippu/piipun	'pipe'
	tt/t	lantti/lantin	'coin'
	kk/k	kirkko/kirkon	'church'

The above rule can be seen to account for these data: whenever there is an open syllable, the left member of the above pairs appears, while that on the right (the 'weakened' variant) shows up when there is a closed syllable. But this does not give an answer to the question which interests Skousen (p. 571), that of whether "speakers actually learn that gradation takes place in a closed syllable." He maintains that the answer is the negative, and that "speakers learn that stems take the weakened form when certain specific suffixes are added; for example, they memorize that the genitive suffix n and the inessive suffix ssä take the weak form of the stem without ever perceiving that both suffixes close the syllable." Skousen's argument in favor of his position has to do first of all with a change found in several of the western dialects of Finnish, whereby the inessive suffix ssä became sä, apparently underlyingly as well as on the surface. In these dialects, standard kädessä shows up as käresä (where r, rather than d, is the weakened form of t in these dialects). "This underlying inessive ending sä does not close the preceding syllable. Nevertheless, speakers of these dialects continue to use the weak form käre with this ending rather than the strong form käte. There is no tendency to change käresä to kätesä (p.571). "This suggests," Skousen continues (pp. 571-2), "that throughout Finnish, speakers are simply memorizing that the inessive suffix takes

the weak form of the stem--no matter how the inessive suffix may change." He goes on (p. 572) to give a similar example concerning the possessive suffixes, which, he claims, "are always added to the strong stem," regardless of how these suffixes affect the phonological environment of the relevant consonant in the stem, in many dialects, including the standard language. "There is absolutely no evidence that speakers ever change the system so that [suffixes which begin with a single consonant] take the weak stem." (I do not give the data with which this example is concerned, since it is so similar to the preceding example--see Skousen 1972:572 for the details). It is not clear from the discussion cited here what precise form Skousen intends his argument to take. However, he is more explicit in Skousen (1975:60) concerning a different example having to do with consonant gradation, and it is this more explicit presentation of the argument form which I will turn to now. I repeat this discussion here, replacing Skousen's "present passive form" (that being discussed in the passage under consideration) by 'inessive and possessive forms', which are the focus of the discussion at issue. Skousen states: "if the generative-phonological solution is correct, then the speaker would view the [inessive and possessive] forms as exceptional to the environment of gradation. And if the speaker must memorize this exceptional fact, we would expect some speakers...to change [these forms] so that [they] would conform to the environment of consonant gradation...However, there is no evidence for such changes..." The argument thus appears to be reconstructible as in (34).

- (34) If speakers regard gradation as taking place in closed syllables, then the inessive and possessive will cause changes to conform to this environment.
Such changes do not occur.
Therefore, speakers do not regard gradation as taking place in this environment.

This reconstruction, of course, has a *modus tollens* form. However, Skousen appears to want a considerably stronger conclusion, in addition to this one--that particular suffixes determine whether or not gradation takes place. In order to warrant this further conclusion, an additional premise to the effect that the "generative phonological solution" and Skousen's proposal are the only possible ones is presumably needed. If this is added--and Skousen seems to have operated under at least roughly such an assumption--then the additional conclusion does of course follow.

Let us now consider Skousen's argument against the second "phonetically-plausible rule" usually posited by generative phonologists, that raising word-final e to i. In some western dialects, word-final k has been completely lost (p. 573), so that "any word ending in ek would now be a surface exception to the rule of e-raising since the k would be missing. Yet in no case does the purported rule $e \rightarrow i / _ \#$ apply to eliminate this surface exception." Another example of this type is "the allative suffix lle, which originally ended in a consonant that has now been deleted..." The final e is not raised here, either. If this argument is intended to be of roughly the same form as the previous one (and cf. also Skousen 1975:67-8), then it can be reconstructed as in (35).

- (35) If speakers have captured the rule of final e-raising, then surface exceptions to it will be eliminated.
Such exceptions are not eliminated.
Therefore, speakers have not captured this rule.

This argument thus again appears to be of the modus tollens type. Since Skousen does not propose an alternative analysis in this case, we need not be concerned with strengthening it as in the last argument.

The final argument to be considered concerns the standard generative rule which converts t to s before i. Skousen argues (p. 573) that "internally-created words such as neiti 'Miss' is [sic] never changed to neisi. Onomatopoetic words like lotina 'splashing' and tippa 'drop' have been created since the historical rule $t \rightarrow s/ _ i$ applied, yet speakers never allow a rule like the historical rule to apply...Another example is the conditional ending isi, which originally came from ńsi, where ń represents a palatal consonant...The palatal ń was later changed to the high i vowel, but after the historical rule $t \rightarrow s/ _ i$ had applied. Consequently, a verb like pote 'to be sick' has the conditional form potisi. Speakers never change this form to posisi." Since this argument is apparently intended to be of the same form as the previous ones (although there is no direct evidence in either of Skousen's works that this is indeed the case), I will not attempt a reconstruction of it.

It can thus be seen that all three of Skousen's arguments are apparently intended to have the same (modus tollens) form; this form is indicated schematically in (36).

- (36) If the standard analysis is correct, then certain changes would occur.
No changes occur.
Therefore, the standard analysis is incorrect.

4.4.2. Kiparsky's reply. Kiparsky begins by claiming (p. 92) that "a more thorough look at the problem indicates that...the rules which Skousen questions are very much in evidence as real synchronic processes of Finnish phonology." Before attempting to establish this claim, however, he first argues that "Skousen is surely right when he says that a phonological rule is real if 'surface violations' of it...tend to get eliminated. But the converse claim, also made by Skousen, that a rule is not real if surface violations of a rule do not tend to become eliminated, is too strong." For one thing, "the failure of a specific change to take place in a specific language at a specific period means nothing, since no one has been able to show conditions under which a change, however natural, must take place...the failure of surface violations of a rule to be eliminated cannot be used as proof that the rule is a linguists' figment" (pp. 92-93). Moreover, he continues (p. 93), "all 'surface violations' of a rule need not be exceptions to it, and if they are not, there is no reason why they should become eliminated." Thus, in the case of the $t \rightarrow s/ _ i$ rule discussed above, although the examples Skousen adduces are indeed correct, "it is in those cases which are necessarily memorized, namely the morpheme-internal cases, that the $t \rightarrow s$ rule does not apply." Kiparsky's justification for the "necessarily" part of this claim is his argument earlier in his paper (cf. especially pp. 65-7) that "non-automatic neutralization processes apply only" when the environment is not "met already in the underlying representation of a simple morpheme." Since Skousen's examples are of this type, Kiparsky's proposal with respect to possible synchronic systems disallows the changes which Skousen maintains are necessary to establish the existence of a rule which has been "captured" by speakers; therefore, if Kiparsky's proposal is correct, "there is no reason to expect" Skousen's examples to undergo the $t \rightarrow s$ rule.

Kiparsky argues further, as alluded to at the beginning of this subsection, that there is external evidence for the rules which Skousen argues against. One rule which he discusses in connection with this concern is that of gradation. In the closely related language of Votic, claims Kiparsky (p. 94), "the inessive -ssa also went to -sa" (as in the dialects discussed by Skousen). Here the part of the gradation rule which affects nongeminates works as in the cases which Skousen described, but "the degemination rule has 'caught up' with the new form of the suffix and fails to apply, e.g., /nokka+sa/ nokkaza 'in the beak' (not *nokaza, which would be the expected form)."

A similar sort of evidence, Kiparsky continues, is available concerning the t → s rule. In support of this contention, he argues (p. 94) that "there is one form class where the rule does seem to have been extended to new cases meeting its structural analysis which arose after the rule entered the language: the past tense form of vowel stem verbs." The stem vowel contracts (pp. 94-5), "under certain conditions, with the past tense suffix -i into i..." This i triggers the rule in some verbs, does so optionally in others, and does not allow it in still others. Kiparsky maintains that this contraction rule historically postdated the t → s rule, and that the latter has therefore applied, at least in some cases, to the output of an historically later rule. Since most historians of Finnish hold that the opposite sequence was the one which actually obtained (and, if so, it cannot be maintained that there is an historically later rule whose output the t → s rule affects). Kiparsky must argue against these contentions; he offers "five reasons why this view of the historical developments is the more likely one."

I will detail here only the fifth argument, which Kiparsky apparently feels is the most convincing, since he states (p. 99) that he believes that it alone would "suffice to establish the point." It concerns "the behavior of t before contracted i elsewhere than in the past tense." He states that "in all of these cases, not only is t the rule, but there are no traces of any kind, either in standard Finnish, or in dialects, or in the older literary documents, of the s which the standard theory claims must once have existed in them." He then asks "why this discrepancy between contracted i (in these cases) which never triggers t → s, and the contracted i of the past tense, which normally triggers s?" He continues that while

...the customary chronology offers no explanation; if instead we assume that t → s preceded contraction, the reason is clear. The difference between the past tense, where t → s was extended to new i's, and all other cases, where t → s was not extended, is that the change t → s before underlying /i/ happens to occur only in the past tense...Prior to contraction, therefore, the process t → s was applicable in the past tense, but in no forms in the other categories....After contraction, the situation was, from a surface point of view, that t went to s always before i from final e..., sometimes before i in the past tense ..., and never before i's in other categories (where all i's came from contraction)...Hence, a 'model' for the extension of t → s to new i's existed only in the past tense.

(Kiparsky 1973a:99-100)

That is, while verbs must be marked in the past tense as to whether or not they undergo t → s, and we thus expect elimination of these markings,

in the other cases there is a general (presumably morphological) characterization of the forms which do not undergo this process, and so we do not expect eliminations of 'exceptions' which do not really exist.

Kiparsky gives no discussion of the final e-raising rule.

4.4.3. Discussion. As noted in section 4.4.2, Kiparsky has no quarrel with Skousen's argument as reconstructed in (32). This suggests that the major premise in this argument should be considered as a potential 'universal' of historical change (see section 5 for further discussion).

But Kiparsky does of course take exception to Skousen's argument concerning consonant gradation. The issue in this case appears to boil down to whether or not the major premise in (34) is in fact true. First of all, there is the question of Kiparsky's proposed constraint on the applicability of "non-automatic neutralization processes" mentioned above. If this constraint is in fact legitimate, and if the only examples in which gradation has not been extended to conform to its putative "phonetically-pausible" environment do conform to the type of case contraindicated by Kiparsky's constraint, then it seems clear that we cannot expect the changes in many cases which Skousen's major premise in (34) predicts, and so this premise would be (at least partly) false. I know of no obvious counterexamples to Kiparsky's proposal, and my knowledge of Finnish is sufficiently limited that I have nothing further to say concerning either of the conjuncts of the if-clause in the above sentence.

Furthermore, as Kiparsky suggests, there does indeed seem to be some reason to question this premise even apart from this constraint, especially if there is no time frame referred to in this argument. However, since there seems to be an implicit such time frame in Skousen's argument, this consideration is open to question. That is, it is possible that an argument containing a premise to the effect that it is likely that certain changes would occur given a fair amount of time will be somewhat more legitimate than the one which Kiparsky argues against. Discussion of this issue, however, will be held off until section 5.

Moreover, given that Kiparsky objects even to the minor premise, further discussion of this issue seems to relegate the former question to a position of relatively minor importance at this point. The question thus currently concerns the truth or falsity of the minor premise. That is, has the gradation rule in fact been generalized, contrary to Skousen's claims? Kiparsky's argument that the answer to this question is in the affirmative, as noted above, concerns evidence from (p. 94) a "closely related language." There is thus some question as to exactly how closely related this language (Votic) really is to Finnish, but I can offer no further comment on this issue, since the reference which Kiparsky cites in this regard is written in Finnish. It should be noted in this respect, however, that if Votic is sufficiently closely related to Finnish in the relevant respects, then the minor premise in (34) does indeed appear to be false.

As for the rule converting t to s before i, Kiparsky does not comment on it with respect to his proposed constraint in his reply to Skousen. He does, however, use this rule as a supporting example when he argues for this constraint (pp. 61, 64), so he presumably would consider his objection discussed above as being relevant in this case as well. He also, of course, claims that this rule has in fact been extended to apply to the output of an historically prior rule, thus apparently contesting the minor premise

of Skousen's argument. This issue is, as Kiparsky points out, highly controversial, and in fact depends crucially on Kiparsky's account of the chronology of the rules involved. There is thus some question as to whether Kiparsky's contention in this respect is correct.

How reasonable is Kiparsky's account of the chronology, then? It seems to me that it is indeed correct that "the customary chronology offers no explanation" of the facts cited in the argument discussed above. Moreover, the other arguments which Kiparsky offers (which were not discussed here) also seem fairly convincing, for the most part. That is, the traditional account does not appear to have much to recommend it. However, it should be pointed out that Kiparsky's account requires that the historically later contraction rule be added to the grammar of Finnish in such a way that it precedes the $\underline{t} \rightarrow \underline{s}$ rule. Since the latter is, according to Kiparsky, a "non-automatic neutralization rule," this account violates King's (1973) constraint against "rule insertion." Not only must there be such a rule insertion on this account, but all forms which would meet the structural description of $\underline{t} \rightarrow \underline{s}$ as a result of the operation of the inserted rule must be marked as exceptions to the $\underline{t} \rightarrow \underline{s}$ rule, since it is potentially fed by contraction in these cases (it is true, of course, as Kiparsky points out, that some of the exceptional cases are generally characterizable, but I am not sure how relevant this fact is in this context). A situation such as that just described would only be possible, as Kiparsky notes, if synchronic phonological theory is altered fairly substantially so that the more or less obvious synchronic description, assuming the possibility of extrinsic ordering of rules, of the state of affairs before the extension of the $\underline{t} \rightarrow \underline{s}$ rule (i.e., ordering contraction after this rule) is not the most highly valued description in a case such as this. It should be noted in this regard that ruling out in principle counterfeeding orders (along with, e.g., Koutsoudas, Sanders and Noll 1974), an alternative which Kiparsky appears to give serious consideration, would entail that all of the "rule reorderings" out of counterfeeding order which Kiparsky and others have advocated cannot be the correct account of the changes involved. There are, however, as Kiparsky has also noted, other alternatives. In sum, Kiparsky's account seems to be less than totally convincing.

What is more, there seems to be a perfectly plausible way of accounting for the pattern of extension of the $\underline{t}/\underline{s}$ alternations even if the $\underline{t} \rightarrow \underline{s}$ rule is no longer alive synchronically, namely that the traditional concept of "analogy" is involved. That is, the $\underline{t}/\underline{s}$ alternation is being extended only in the past tense because it is only here, as Kiparsky notes (see above), that there is "a 'model' for the extension of" this alternation. Formally, what is going on under this approach is sketched in (37), where \underline{V} represents any vowel other than \underline{i} and \underline{Q} is the morphological category involved, with the obvious solution to this proportional equation being ' $x = s$ '.

$$(37) \quad tV : si]_{\underline{Q}} :: tV : xi]_{\underline{Q}}$$

Notice that the only occasion in which the left hand side of this formula will occur is when \underline{Q} is past tense--there are no instances of $[si]$ (only $[ti]$) in the other morphological categories discussed by Kiparsky. Thus the facts which he cited in the argument discussed here seem amenable to an alternative account, if the chronology is as he argues it is. It should be noted that this solution does not require rule insertion, since

it does not require the existence of the $t \rightarrow s$ rule and therefore of course does not require any particular ordering relationship between it and the contraction rule. It seems to me that the other data cited by Kiparsky would be amenable to a precisely analogous account.

I have no way of telling whether the Votic facts discussed in connection with the gradation rule are also amenable to this type of account, since Kiparsky presents very little data, and since it would be difficult, it not impossible, for me to find further data.

It should be pointed out that this type of account entails rejecting the premise that a rule is synchronically valid if surface violations of it tend to get eliminated. Further discussion of this question will be given in the next section.

5. Conclusion. It should be clear that all of the arguments examined in section 4, at least if they are interpreted perhaps somewhat charitably, can be considered as instances of elementary logically valid forms of inference--classical modus ponens and modus tollens (or, probably more realistically in many cases, the 'almost' variants of these mentioned in section 3), and the Bayesian form schematized in (2).

This situation, in my experience, is quite characteristic of arguments given by linguists (although there are, perhaps inevitably, some exceptions), and I suspect that it is generally true of arguments given in any science. This should not seem surprising, since linguists and other scientists are, at least indirectly (and sometimes directly), schooled in logical analysis. In fact, it seems to me that it is a good rule of thumb for a methodologist to follow that if it appears that an investigator is committing an elementary logical fallacy, then the methodological analysis itself is not unlikely to be faulty and that it therefore merits considerable scrutiny. I would suggest also in this respect that it is in general more likely that any lack of persuasive power felt concerning the arguments in question is probably due to a corresponding lack of belief in the truth of the premises involved in a logically valid argument form. That is not to say, of course, that investigators are never guilty of such logical fallacies (there are undoubtedly a number of quite genuine cases of this kind), but this need not compromise the value of this suggested rule of thumb as such. I would thus regard the putative discovery of an ever-increasing number of new types of fallacies (concerning mainly arguments in favor of extrinsic rule ordering--cf., for example, Koutsoudas 1972) as cases of somewhat misguided methodological analysis, and would maintain that what is actually involved here is disagreement concerning the truth of (implicit) premises in the arguments in question. This should not be taken as implying that such methodological studies have no value, for they have often provided good reasons for questioning the premises at issue (at least in my opinion), but terming such arguments 'fallacious' adds a rhetorical effect to the criticism that does not seem to me to be appropriate in such cases; investigators can differ with respect to beliefs about the truth of premises without one of them necessarily being found lacking in a necessary skill of the field--that of constructing logically valid arguments--which is not of course the case if he is in actuality guilty of a logical fallacy. Adopting such a rule of thumb, moreover, appears to be quite in keeping with the approach of many contemporary philosophers of science (cf., e.g., Suppe 1977), who place considerable weight on the actual practice of working

scientists. The work of the 'best' scientists is, not too surprisingly, held to be of the most importance by such investigators, and I would have to go along with them in this respect as well: we should be doubly skeptical when our methodological analysis entails that a classical argument (i.e., one which has met with considerable acceptance over a fairly long period of time) involves a logical fallacy.²⁴

What is important as far as the evaluation of the arguments discussed above is concerned is thus the truth (or likelihood) of the premises involved. I would like to turn now to a brief consideration of this issue. This is an especially important concern, since at least some of these premises (or generalizations or revisions of them) are relevant not only to these arguments, but also potentially to other arguments which invoke diachronic data.

The major premise (5) of Kiparsky's argument in favor of the brace notation is, unfortunately (especially given the amount of attention it has received), not one of these. It appears to be quite specific to this particular argument, and I can see no obvious way of generalizing it. There is a lesson which can be gained from a consideration of this premise, however: the perceived cogency of an argument from historical change (or any other, for that matter) will vary with the degree of belief of each individual investigator concerning the truth of the premises involved. In this case, Kiparsky (and Chomsky and Halle) apparently found the relevant historical changes to indeed be quite unlikely unless the brace notation was psychologically real; Stampe, on the other hand, did not and was thus unconvinced by the argument--despite his acknowledgment of its "beauty". As pointed out above, evaluating the truth of this premise depends ultimately on (one's degree of belief in) the correctness of Kiparsky's (and Stampe's) synchronic analyses of Old English and Early Middle English. Moreover, in this case--and in many others, I would maintain--there is no clear way of establishing the required "correctness"--the problematic forms adduced by Stampe could always be given an (ad hoc) explanation, especially by a staunch syllable opponent. Whether or not a given investigator finds this argument (and other, perhaps all, arguments as well) convincing thus depends in the final analysis on essentially subjective factors.

As for the Hooper-Harris debate, the major premise (12') appears to be more readily generalizable, although even here what is probably the most obvious generalization--namely, that leveling not in the direction of the putative underlying form is problematic for a diacritic theory--is not terribly general. Here again the ultimate evaluation of this premise depends on subjective factors. The diacriticist could, for example, maintain that there is an implicit ceteris paribus clause attached to the prediction of direction of leveling, and that in the case(s) at issue there are other essentially irrelevant factors which are responsible for the 'failure' of the prediction. (In this particular case, in addition, the apparent similarity of the prediction made by the Hooperian analysis would likely mitigate the force of an objection based on this premise.) I can see no really objective way of determining whether or not these other factors are in fact irrelevant. The differences in directionality of leveling with respect to different dialects are suggestive, however, and I will return to this question below.

In the case of the Schuh-Leben controversy, what appears to be behind the major premise in (21) is something which is very similar to the principle

just mentioned: 'underlying' forms are favored in cases of analogical change. Similar sorts of questions can of course be asked about this principle. Leben's particular response concerning the directionality involved does have its problematic aspects, but this is not to say that there is no other analysis which does not require inverse rules which can make this directionality seem reasonable; the force of this argument depends on a subjective degree of belief in the likelihood of the possibility of such an alternative analysis. Similarly, an evaluation of the major premise in (24) (which does not seem amenable to generalization) depends on subjective factors, although, as pointed out above, Leben's particular synchronic analysis of the phenomenon in question is not without its problematic aspects. Here, too, however, nothing precludes a successful noninverted analysis. Since it appears that nothing new can be learned from a reconsideration of the remaining arguments given by Schuh, I will not discuss them here.

As noted in section 4.4.3, the major premise in (32) deserves serious consideration; I give a somewhat generalized form of this premise below in (38).

- (38) If a rule applies to forms which have appeared in the language, then it is psychologically real (i.e., "has been captured by speakers").

Unfortunately, it does not appear that it can be maintained. For one thing, how are we to tell if the if-clause in (38) is satisfied? The fact that a change in a form could be the result of the (productive) application of a synchronic rule is not sufficient to establish that the putative rule is indeed present in the language at the stage in question, as long as there is the possibility that there may be no rules involved in either the change or the synchronic alternations which preceded the change. It does not seem at all unlikely, moreover, that this possibility should be taken seriously, given that the direction of leveling does not appear to be predictable. That is, the different directions of leveling found in Spanish dialects (cf. section 4.2) could be taken to indicate, as Hooper suggests, that different speakers can have different rules, or that speakers have no rules at all for the phenomena at issue; I know of no clear way of distinguishing these two alternatives empirically (cf. note 16). To be more explicit about this new possibility, what I am suggesting (and I emphasize that at this stage it is only a suggestion) is that, except for productive processes and external sandhi alternations, speakers capture no psychologically real regularities (to use Skousen's terminology); i.e., other than in the cases just mentioned, speakers simply memorize the words of their language in roughly their classical phonemic representations, and they learn no rules (either morphological or phonological) to relate these words.²⁵ If a speaker does not know a word, then he quite literally uses analogy (cf. section 4.4.3) to come up with the required 'new' word.²⁶ Often the result of this use of analogy will be the word which already exists in the language, but sometimes speakers will not be 'successful' in their use of analogy, in that they come up with something that has heretofore not been accepted as a word of the language. In such cases, and if the new word meets with acceptance, there will have been an analogical change.

This possibility depends crucially, of course, on the nonpredictability of the direction of leveling, and I would like to briefly present some further evidence for such a nonpredictability.²⁷ The Spanish facts are

not isolated ones, and Skousen (1975) gives some discussion of some similar examples from the history of French. For example (p. 36), "the verb boire, according to the normal historical development, should have a future-conditional stem buvr-", and it did have for some time in the history of French, but it "has now been replaced by the stem boir-", so that the alternation in the stem for this verb has been leveled out in favor of the infinitive. However, we find (pp. 38-9) that "in Old French, infinitive forms like ardoir and saillir were replaced by ardre and saudre. The future forms of these verbs were historically ardra and saudra." Thus, here again the stem alternation has been leveled out, but this time in favor of the future. Skousen interprets these data as evidence that some speakers "learn that the future-conditional stem is the same as the infinitive" (p. 36), while others "have learned the reverse pattern--that the infinitive is based on the future-conditional stem" (p. 37), but unless we can predict which speakers are going to make which analysis, and/or which of these analyses will 'win out' in cases involving leveling, we are still not going to be able to predict the direction of leveling (cf. also note 16). Skousen gives at least two further examples of this type (cf. the discussion of the development of parler and arraisonner, and of aimer and clamer (pp. 43-5)) as well.

Skousen's discussion of these levelings is of some interest, it seems to me, and I will digress here a bit to give some attention to this discussion. One of the main points of Skousen's monograph is (e.g., p. 41) that "...analogical changes [of the type discussed above] can be used as evidence for a psychologically real regularity between the infinitive and the future-conditional stem." Furthermore, he has (p. 41) "been using a restricted theory of analogy. In particular, [he has] assumed that analogical changes are not random, but occur in directions...Analogical changes occur when speakers remove exceptions to psychologically real regularities." He continues (p. 42) that "in this sense, analogical change is viewed not as the speaker's attempt to create a surface regularity, but rather as an attempt to eliminate surface exceptions to a surface regularity that has already been captured." Thus, Skousen appears to be proposing something quite contrary to what I have just suggested--any analogical change, for Skousen, suffices to establish the existence of a rule (= "psychologically real regularity"). This seems to be a stronger claim than Hooper's concerning similar facts (cf. note 16 and the related discussion), and also an incorrect one, at least if the concept "psychologically real regularity" is given a nontrivial interpretation.

Consider, for example, the development of the now quasi-productive -burger morpheme in English (cf. Jeffers and Lehiste 1979). There was of course originally a single word hamburger which consisted of two morphemes, hamburg and -er, but after this word had been (mis)analyzed as having the morphological composition ham+burger (presumably due to the presence of the word ham in the language), the new morpheme -burger has been extended so that there are now in English such words as cheeseburger and fishburger (cf. also Burger King, Burger Chef). Note that this development began with a single form, so unless we are willing to claim that a "psychologically real regularity" can be derived from a single example (and what could constitute an irregularity if this were the case?), it appears that Skousen's claim cannot be upheld. Another example of this type concerns the recently developed -(a)holic morpheme, which again appears to be the result of a misanalysis of a single form as alc+oholic (we now have words like workaholic,

gumaholic, etc.). (For discussion of a somewhat similar case concerning the development of the Latin infinitive, cf. Jeffers and Lehiste 1979.)

The upshot of all this is that (38), at least in its present form, cannot be maintained. But how about a somewhat weaker version? One could suggest, for example, that analogical changes which could be the result of the application of productive or external sandhi rules are evidence for the psychological reality of such rules. But even this has its problematic aspects. What is the difference between a productive rule and a nonproductive one? And how can we be sure that our proposed rule, and not another (more or less) empirically equivalent rule, is responsible for the change? Even the strongest tenable version of (38) must apparently be rather weak with respect to such factors, probably weaker than that just suggested, and will thus be of correspondingly little use in evaluating synchronic analyses.

One final alteration of this principle will be suggested here, one which in fact appears quite plausible to me, although I know of no conclusive reasons for maintaining it. (Here again, then, the success of arguments incorporating this principle as a premise apparently will ultimately depend on a subjective evaluation of the truth of this principle by each individual investigator.) The alteration involves the kind of changes involved, in particular whether or not they are across-the-board ones. The final version of this principle is given more explicitly in (39) below.

- (39) If a putative rule applies to every possible input for the rule which arises after the postulated entry of the rule into the language, then it is psychologically real.

In fact, it could be that Skousen had something of this nature in mind, rather than the (reconstructed) major premise in (32), since he claims (Skousen 1972:569) that "in every case, the rule of gemination eliminates [the relevant] exceptions" (emphasis added).

Let us turn now to an examination of the major premise in (36). This version undoubtedly should be generalized even further if it is to be considered as a universal principle of language change; a further generalization is given below in (36').

- (36') If a putative rule is psychologically real, then (relevant) changes will occur.

As noted above, Kiparsky rejects this principle, and it is sufficiently vague anyway (when will the changes occur?) that it would require revision in any event. It seems to me that a principle such as (36''), in which (36') has been made probabilistic and a time frame has been added, deserves serious consideration, although again it is not clear to me what sorts of evidence would be relevant to determining its truth or falsity.

- (36'') If a putative rule is psychologically real, then (relevant) changes are likely to occur within several centuries.

As noted above (section 4.4.3), Skousen may have implicitly maintained such a time frame. It is perhaps worth noting that even Kiparsky may subscribe to a principle such as this, since he did not stop at proclaiming that the major premise in (36) was false, but proceeded to argue that the

minor premise in this argument is also false. The answer to questions concerning the relevance of historical data to determining the structure of a synchronic grammar thus again appears to depend in large part on the subjective degree of belief of each individual researcher in the premises of the argument at issue.

The emphasis which I have given to the role played by subjective factors in explicating the relative convincingness of an argument will undoubtedly be disturbing to some, and I must admit that I am somewhat uncomfortable about it myself. But if it is believed that it is desirable that the philosophy of science reflect in large part the actual practice of scientists (as, again, is now widely held--cf. Suppe 1977 and the references cited there), then such a conclusion seems to me to be inescapable. Again, I do not mean to claim that investigators are never guilty of assigning truth values to premises in an intuitively undesirable manner, but I can see no nonarbitrary way of proscribing such 'undesirable' behavior. A fair amount of disagreement about the reality of received 'knowledge' in any given field thus appears to be inevitable, unless there is widespread agreement in the field concerning the premises which play a substantial role in the arguments advanced in these fields (i.e., unless there is a "paradigm" in a fairly strong sense of this term as used by Kuhn 1970). This is perhaps a somewhat pessimistic view to take of scientific knowledge, but if the cases considered here are at all representative of linguistic research, and if linguistics is a more or less representative science (as the non-linguistic arguments discussed in Churma 1979:Chapter II at least suggest), then such an approach seems to be inevitable. The truth (or likelihood) of at least the major premises in most scientific arguments (not just arguments from historical change) simply is not decided on in practice by 'objective' means, and for the most part is not in principle decidable in such ways. But the situation is not as chaotic under this view as the subjective nature of the evaluations might lead one to believe, and in fact the framework points to an intuitively reasonable procedure in cases of disagreement about the force of an argument. With the isolation of the crucial premises involved, proper attention can be given to research concerning these premises; an investigator who is convinced of the falsity of such a premise can seek out counterexamples, and an adherent of one can attempt to marshal evidence which makes the premise more likely. But we will never have certain knowledge in any empirical field, I would suggest.

It is worth noting in this regard that if the truth of the premises in the Bayesian arguments (and probably the others as well--cf. Churma 1979:Chapter II, Appendix) is objectively ascertainable, then the much discussed 'problem of induction' (cf., e.g., Salmon 1967) will have been solved. Given the lack of success in the solution of this problem in the past few hundred years, one could, it seems to me, reasonably look at a putative solution to this problem with a fair amount of reservation. The subjectivity required in my interpretation thus seems to me to be in some sense a virtue of such an approach, rather than a liability; while one might in principle desire certain knowledge, it seems that, as suggested above, in practice it is not possible, and that therefore an approach which claims to provide such knowledge is questionable solely on this basis. Thus, I would maintain that it is not possible to reduce inductive inference (in particular, Bayesian inference) to deductive inference, and that to hold that it is possible (as in establishing with certainty the probabilities involved in a Bayesian argument) involves a fundamental misconception. (For further discussion of this point, cf. Churma 1979 Ch. II.)

One final positive note deserves mention in this respect. This concerns the fact that, as Salmon (1967:122) puts it, in spite of the 'problem' that "the prior convictions of reasonable people can vary considerably," Bayes' Theorem entails that "as these individuals accumulate a shared body of observational evidence, the differences of opinion will tend to disappear and a consensus will emerge." That is, differences in prior probability assignments are more or less irrelevant, as long as the "shared body of observational evidence" is large enough; as long as enough evidence can be found, rational investigators will eventually be forced to have roughly the same degrees of belief concerning 'well-supported' hypotheses. Until such a body of evidence is accumulated, however, the main determinant of relative persuasiveness of a given argument will be the degree of belief of each individual investigator in the truth of the premises of that argument, and this degree of belief may, of course, vary from person to person. The premises of arguments which invoke data from historical change are no different from any others in this respect, and so the value of such arguments for the purpose of assessing the descriptive adequacy of proposed grammar fragments apparently must be tied correspondingly to the relevant subjective degrees of belief.

Footnotes

*This paper is a revised version of parts of my dissertation (Churma 1979), mostly Chapter III. I would like to thank Fred Householder, Wayne Redenbarger, David Stampe, and Arnold Zwicky for helpful discussion of the issues involved.

¹I extend here (naturally, I feel) Chomsky's concept of descriptive adequacy so that a partial grammar is descriptively adequate if it is (or would be) part of a complete descriptively adequate grammar. It is perhaps appropriate that I give here at least a working definition of what I consider a descriptively adequate grammar to be: following Chomsky and Halle (1965:99), such a grammar "...gives a correct account of the speaker's 'tacit knowledge'" of his language. The term 'tacit knowledge' here should prevent any misunderstanding which might arise if Chomsky's (1964:63) term "linguistic intuition" (Chomsky apparently considers the two expressions to be equivalent) is employed--the latter could create the impression that the issues at stake in phonological controversies may be settled by direct appeals to the "linguistic intuition" of native speakers about the constructs in question. On the contrary, most (if not all) contemporary phonologists appear to agree that intuitive judgments about the validity of these constructs cannot be made by speakers with any degree of reliability. The 'tacitness' of the knowledge involved is thus crucial.

²In my experience, it is almost always the truth of the premises involved which is at issue when disagreements as to the force of an argument arise--even if one of the adversaries claims that the other is guilty of a 'logical fallacy.' In addition, the implicitness of the premises in many arguments will often add to the difficulties. For further discussion of this question, see Churma 1979, Ch. III.

³For a similar discussion of the areas of child language acquisition and word games, see Churma 1979. For an essentially complete list of the types of data used as external evidence in the literature, see Zwicky 1975.

⁴Thus, the present study can be considered as an attempt to provide what Botha (1979b) calls "bridge principles" for relating synchrony and diachrony in phonology. It is worth pointing out here that this latter work (cf. also Botha 1979a) presents what appears to me to be a rather distorted interpretation of the views concerning external evidence expressed by Chomsky (1976). Contrary to Botha (p. 39), who interprets passages from Chomsky (pp. 5-6) as entailing that Chomsky feels that external evidence is not necessary "for the validation of mentalistic claims", it seems that what Chomsky actually intends here is that (p. 12), given the current state of linguistic research, internal evidence is likely to be more valuable, given the essential lack of relevant "bridge principles" in the case of external evidence. As Botha would undoubtedly be the first to admit, the relevance of external evidence to synchronic analysis is far from clear without such principles. And external evidence would be necessary, of course, in cases where internal evidence alone does not allow for a choice between two alternative accounts. Botha has also made some other rather disturbing misinterpretations of Chomsky's views on psychological reality and external evidence, in my opinion, but this is not the place to go into them.

⁵The only works of this type of which I am aware are the long (and for the most part misguided, in my opinion--cf. section 4.2.3 for details) discussion of an argument from historical change given by Kiparsky (1968) in Botha 1973 and the brief discussion of arguments from historical change in Sommerstein 1977. The latter discussion, though perhaps of some relevance to the concerns of this study, will not be examined here, since it appeals only to hypothetical examples; the cases examined here all come from actual arguments in the literature in favor of or opposed to a given analysis of natural language data.

⁶The "at least partially" qualification is intended to allow for 'supra-segmental' or 'prosodic' features of speech, i.e., ones which are not of the same character as the more 'ordinary' features traditionally made use of in linguistic analysis, such as those employed in Firthian prosodic analysis or Harrisian 'long components', and more recently in the work of Leben (1973) and Goldsmith (1976).

⁷It has often been argued that there exists a level of representation intermediate to the levels of lexical representation and systematic phonetic representation. Since these arguments do not make use of the type of evidence at issue here, as far as I know, they will not be considered here, nor will the issue be pursued further.

⁸The "somewhat" qualification is meant to reflect the fact that, if the probability that A is true = \underline{p} , the probability that B is true given that A is true = \underline{q} , and the probability that B is true = \underline{r} , then $\underline{r} \leq \underline{p} \cdot \underline{q}$. See Churma 1979, Appendix for a proof.

⁹Note that, if expressed in this way, the argument appears to say nothing at all about the psychological reality of braces, but only about whether or not they are "part of linguistic theory." This is a rather trivial problem (although Botha (1973) apparently does not agree, as will be seen in section 4.1.3), since it is clear from the last part of the

first paragraph quoted above that Kiparsky feels that the theory is making claims about psychological reality. In other words, a construct is part of linguistic theory if, and only if, it is psychologically real. When we combine this with the conclusion of the preceding argument (that the brace notation is supported as being part of linguistic theory), we of course have the conclusion that its psychological reality is supported.

¹⁰The requirement that E be at least fairly likely if H is true is not explicit in (2). In fact, as (2) is stated, this requirement can only be arrived at as the result of some sort of Gricean implicature (cf. Grice 1975). Such a requirement seems intuitively quite reasonable, and moreover appears to be supported by the more explicit version of (2) mentioned above. See Churma 1979, Ch. II for details.

¹¹It is not clear that any data are of this type; see, for example, Kuhn 1970 .

¹²"Rule 19" referred to by Harris is a lexical redundancy rule which expresses the fact that "only third-conjugation verbs have the mid/high and mid/high/diphthong alternations" (Hooper 1976:161). It is given formally in (i).

$$(i) \left\{ \begin{array}{l} \left[\begin{array}{l} -\text{syll} \\ +\text{high} \end{array} \right] \quad \left[\begin{array}{l} +\text{syll} \\ -\text{back} \\ -\text{high} \end{array} \right] \\ \left[\begin{array}{l} +\text{syll} \\ -\text{high} \end{array} \right] \\ \left[\begin{array}{l} +\text{syll} \\ +\text{high} \end{array} \right] \end{array} \right\} \text{Co]}_{\text{verb}} \text{ IS ALWAYS [+ 3rd conj.]}$$

It should also be pointed out that Harris's description of the change in the form of rule (1) appears to be different from that of Hooper, since she implies that rule (2) above is the only relevant rule which remains after the change, so that the entire second case, not just its environment, has been lost from rule (1).

¹³Rule 14 is a lexical redundancy statement which is roughly equivalent to part of rule (i) in note 12, and indicates (Harris 1978:47) that the diacritic "[HM] appears only on the last nonlow vowel of a relatively small number of third-conjugation stems." It is given in (i).

$$(i) \text{ If [HM], then } \#X \left[\begin{array}{l} \text{---} \\ +\text{syll} \\ -\text{low} \\ 3 \text{ conj} \end{array} \right] \text{ C}_0 \#.$$

¹⁴For a very similar formulation, cf. Harris (1978:56).

¹⁵There remains Harris's point that "nothing in Hooper's account reflects the fact that diphthongization is lost in one particular morphological subclass..." This appears to be quite true, but since there appears to

be nothing to prevent Hooper from adding to the grammar of the innovating dialects a "special statement" analogous to that of Harris to this effect, I will not pursue this point further.

¹⁶There is of course also the possibility that, as Hooper puts it (p. 167), "different speakers may arrive at different analyses of the same data...", so one could claim that a somewhat different rule--one in which the diphthong is in fact the elsewhere case in the relevant part of the rule--is present in the dialects which level in favor of the diphthongs. If this is true, however, it is impossible to disconfirm the theory with facts from historical change, given the essential lack of constraints on the form which rules may take, since it will always be possible to formulate a rule in which the favored variant is the elsewhere case. Thus, the theory would be making essentially no empirical claims, despite the invocation of 'external evidence,' as well as having no predictive power concerning the direction of leveling.

¹⁷For a fuller examination of the issue of whether "rule inversion in Chadic" has in fact occurred, see Churma (ms.)

¹⁸As Leben (1974:269) points out, this rule (or another one) must apparently also convert stem-final -e in verbs to ə between a stop and a sonorant consonant. Leben treats it as being part of the epenthesis rule, formulating it as in (i).

(i) $\left\{ \begin{array}{c} \phi \\ e \end{array} \right\} \rightarrow \text{ə} / [-\text{son}] \text{C} [+ \text{son}]$

(Schuh is not very explicit on any of the rules in this part of his article, and it is somewhat difficult to tell exactly what he intends.) It is not clear that these phenomena actually can be treated as a single rule, since there are cases where -e alternates with ψ (see below).

¹⁹Following Leben, I modify here the Hausa orthography used by Schuh to reflect the w output of Klingenheben's Law. The orthographic u is retained in direct quotations from Schuh, however, so it is of some importance to remember that his uu and au correspond to my (and Leben's) uw and aw.

²⁰Leben (1977a, b) has recently given revised analyses of Hausa plurals in the framework of "upside-down" phonology (cf. Leben and Robinson 1977). Since this is not relevant to the present discussion, it will not be considered here. For some general criticisms of the upside-down framework, see Churma 1980b, Janda 1980.

²¹This is an example of the type alluded to in note 18, where the -e must be deleted rather than converted to schwa. It is not clear to me from Newman's discussion what the precise conditions on -e deletion are.

²²For a perhaps equally plausible alternative account of these developments in the Hausa plural system within an inverted rule framework, cf. Churma (ms.).

²³ Since Leben's main concern seems to be to restrict as much as possible the range of permissible analyses of language data (by, in this case, eliminating the "middleman" of inverse rules "from the realm of possible phonological systems"--cf. Leben, pp. 265-6), it is not clear that Leben's own proposal does not violate the spirit of his enterprise; in what sense is linguistic theory restricted if inverse rules are disallowed, but "competing underlying forms" are not?

²⁴ Sadock (1976) is thus quite rightly bothered by the fact that his analysis of Halle's (1959) argument against the phoneme requires that it not correspond to a logically valid form of inference. For further discussion of Halle's argument and Sadock's critique, see Churma 1980a

²⁵ This suggestion shares some aspects of other currently proposed 'concrete' theories of grammar, in which words are listed in the lexicon in roughly their surface form (cf. Vennemann 1974b, Leben and Robinson 1977, Pollack 1977, Farrar 1978). My suggestion seems to differ from these proposals in (1) distinguishing nonproductive word-internal processes from all others, and (2) claiming that in such cases there are no rules at all, whether "via," "upside-down", or "redundancy."

²⁶ It has often been pointed out (cf., e.g., Jeffers 1974, Skousen 1975) that the term "analogy" has frequently been used carelessly in historical linguistics. However, if we are careful to make clear exactly what the analogy being appealed to is, it seems to me that this concept is not nearly as 'dangerous' as it has often been claimed to be (cf. King 1969: 139ff.). For further discussion along these lines (but in the context of historical linguistics), see Jeffers 1974

²⁷ It is true that there appear to be regularities (at least tendencies) concerning which forms are innovated (cf. Kurylowicz 1949, Watkins 1969, Hooper 1979). At least some of these tendencies (e.g., Kurylowicz's 'second law' that analogy proceeds from "formes de fondation" to "formes fondées") can perhaps be seen in terms of the framework suggested above to be the result of the (putative) fact that a speaker is more likely, based on considerations of relative frequency and perhaps other similar factors to know certain kinds of forms than he is to know other kinds (e.g., he would be more likely to know a "forme de fondation" than a "forme fondée").

References

- Anderson, J. and C. Jones, eds. (1974). Proceedings of the First International Conference on Historical Linguistics. Vol. 2. Amsterdam: North Holland.
- Botha, R. (1973). The Justification of Linguistic Hypotheses. The Hague: Mouton.
- Botha, R. (1979a). External evidence in the validation of mentalistic theories: a Chomskyan paradox. Stellenbosch Papers in Linguistics 2:1-38.
- Botha, R. (1979b). Methodological bases of a progressive mentalism. Stellenbosch Papers in Linguistics 3:1-115.

- Brown, E. and R. W. Rieber, eds. (1976). *The Neuropsychology of Language*. New York: Plenum Press.
- Bruck, A., et al., eds. (1974). *Papers from the Parasession on Natural Phonology*. Chicago, IL: Chicago Linguistic Society.
- Campbell, R. J., et al., eds. (1974). *Linguistic Studies in Romance Languages*. Washington, D.C.: Georgetown Univ. Press.
- Cole, P. and J. Morgan, eds. (1975). *Syntax and Semantics, Vol. 3: Speech Acts*. New York: Academic Press.
- Chao, Y.-R. (1934). The non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology, Academia Sinica* 4:363-397. Reprinted in Joos (1957).
- Chomsky, N. (1964). Current issues in linguistic theory. In Fodor and Katz (1964).
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1976). On the biological basis of language capacities. In Brown and Rieber (1976).
- Chomsky, N. (1978). An interview with Noam Chomsky. (S. Saporta, interviewer). *Linguistic Analysis* 4:301-319.
- Chomsky, N. and M. Halle (1965). Some controversial questions in phonological theory. *Journal of Linguistics* 1:97-138.
- Chomsky, N. and M. Halle (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Churma, D. (1979). *Arguments from External Evidence in Phonology*. Unpublished Ph.D. Dissertation. Columbus, OH: The Ohio State University.
- Churma, D. (1980a). A further remark on the 'Halleian syllogism.' This volume.
- Churma, D. (1980b). Some further problems for upside-down phonology. This volume.
- Churma, D. (ms.) Rule inversion in Chadic: a closer look. Columbus, OH: The Ohio State University.
- Dinnsen, D., ed. (1979). *Current Approaches to Phonological Theory*. Bloomington, IN: Indiana University Press.
- Donegan, P. and D. Stampe. (1979). The study of natural phonology. In Dinnsen (1979).
- Dressler, W. and F. Mares̃, eds. (1975). *Phonologica 1972*. München: Wilhelm Fink Verlag.
- Farrar, C. (1978). *Phonological regularities in the lexicon*. Unpublished paper. Columbus, OH: The Ohio State University.
- Fodor, J. and J. Katz, eds. (1964). *The Structure of Language*. Englewood Cliffs, N.J.: Prentice Hall.
- Fujimura, O., ed. (1973). *Three Dimensions of Linguistic Theory*. Tokyo: TEC Company, Ltd.
- Goldsmith, J. (1976). *Autosegmental phonology*. Bloomington, IN: Indiana University Linguistics Club.
- Grice, H. P. (1975). Logic and conversation. In Cole and Morgan (1975).
- Halle, M. (1959). *The Sound Pattern of Russian*. The Hague: Mouton.
- Harris, J. (1974). Morphologization of phonological rules: an example from Chicano Spanish. In Campbell, et al., (1974).
- Harris, J. (1978). Two theories of non-automatic morphophonological alternation. *Language* 54:41-60.
- Hooper, J. (1976). *Introduction to Natural Generative Phonology*. New York: Academic Press.

- Hooper, J. (1979). Substantive principles in natural generative phonology. In Dinnsen (1979).
- Hudson, G. (1974). The representation of non-productive alternations. In Anderson and Jones (1974).
- Hudson, G. (1975). Suppletion in the Representation of Alternations. Unpublished Ph.D. Dissertation, Los Angeles, CA: UCLA.
- Janda, R. (1980) 'Upside-down' phonology: regenerative or degenerative? Bloomington, IN: Indiana University Linguistics Club.
- Jeffers, R. (1974). On the notion 'explanation' in historical linguistics. In Anderson and Jones (1974).
- Jeffers, R. and I. Lehiste (1979). Principles and Methods of Historical Linguistics. Cambridge, MA: MIT Press.
- Joos, M., ed. (1957). Readings in linguistics I. Chicago, IL: University of Chicago Press.
- Juilland, A., ed. (1977). Linguistic Studies Offered to Joseph Greenberg. Saratoga, CA: Anma Libri.
- Kenstowicz, M. and C. Kisseberth, eds. (1973). Issues in Phonological Theory. The Hague: Mouton.
- King, R. (1969). Historical linguistics and generative grammar. Englewood Cliffs, N.J.: Prentice Hall.
- King, R. (1973). Rule insertion. *Language* 49:551-578.
- Kiparsky, P. (1968). Linguistic universals and linguistic change. In Bach and Harms (1968).
- Kiparsky, P. (1973a). Phonological representations. In Fujimura (1973).
- Kiparsky, P. (1973b). Productivity in phonology. In Kenstowicz and Kisseberth (1973).
- Klingenheben, A. (1928). Die silbenauslautgesetze des Hausa. *Zeitschrift für eingeborenen Sprachen* 18:272-297.
- Koutsoudas, A. (1972). The strict order fallacy. *Language* 48:88-96.
- Koutsoudas, A., G. Sanders, and C. Noll. (1974). The application of phonological rules. *Language* 50:1-28.
- Kuhn, T. (1970). The structure of scientific revolutions, second edition. Chicago, IL: University of Chicago Press.
- Kuryłowicz, J. (1949). La nature des procès dits analogiques. *Acta Linguistica* 5:15-37.
- Leben, W. (1973). Suprasegmental Phonology. Unpublished Ph.D. Dissertation, Cambridge, MA: MIT.
- Leben, W. (1974). Rule inversion in Chadic: a reply. *Studies in African Linguistics* 5:265-278.
- Leben, W. (1977a). Doubling and reduplication in Hausa plurals. In Juilland (1977).
- Leben, W. (1977b). Parsing Hausa plurals. In Newman and Newman (1977).
- Leben, W. (1979). The phonological component as a parsing device. In Dinnsen (1979).
- Leben, W. and O. Robinson. (1977). Upside-down phonology. *Language* 53: 1-20.
- Mandelbaum, D., ed. (1949). Selected writings of Edward Sapir. Berkeley, CA: University of California Press.
- Newman, P. (1974). The Kanakuru language. (West African Monographs 9) Leeds: University of Leeds Dept. of Modern Languages.
- Newman, P. and R. M. Newman, eds. (1977). Papers in Chadic Linguistics. Leiden: Afrika Studiecentrum.

- Peranteau, P., et al., eds. (1972). Papers from the Eighth Regional Meeting of the Chicago Linguistic Society. Chicago, IL: Chicago Linguistic Society.
- Pollack, J. (1977). Lexical features in phonology. Unpublished Ph.D. Dissertation. Columbus, OH: The Ohio State University.
- Sadock, J. (1976). On significant generalizations: remarks on the Hallean syllogism. In Wirth (1976).
- Salmon, W. (1967). The Foundations of Scientific Inference. Pittsburgh, PA: University of Pittsburgh Press.
- Sapir, E. (1933). La réalité psychologique des phonèmes. *Journal de Psychologie Normale et Pathologique* 30:247-265. Reprinted in English in Mandelbaum (1949).
- Schuh, R. (1972). Rule inversion in Chadic. *Studies in African Linguistics* 3:379-97.
- Schuh, R. (1974). A comment on "Rule inversion in Chadic: a reply". *Studies in African Linguistics* 5:279-280.
- Skousen, R. (1972). On capturing regularities. In Peranteau, et al. (1972).
- Skousen, R. (1975). Substantive Evidence in Phonology. The Hague: Mouton.
- Sommerstein, A. (1977). Modern Phonology. Baltimore: University Park Press.
- Stampe, D. (1973). A Dissertation on Natural Phonology. Unpublished Ph.D. Dissertation. Chicago, IL: University of Chicago.
- Suppe, F., ed. (1977). The Structure of Scientific Theories. Urbana, IL: University of Illinois Press.
- Vennemann, T. (1972). Rule inversion. *Lingua* 29:209-242.
- Vennemann, T. (1974a). Restructuring. *Lingua* 33:137-56.
- Vennemann, T. (1974b). Words and syllables in natural generative grammars. In Bruck, et al. (1974).
- Watkins, C. (1969). A further remark on Lachmann's Law. *Harvard Studies in Classical Philology* 74:55-66.
- Zwicky, A. (1975). The strategy of generative phonology. In Dressler and Mareš (1975).

A Further Remark on the "Hallean Syllogism"*

Donald G. Churma

In the past few years, there has been a rebirth of interest in Halle's (1959) classic argument against the phoneme; Sullivan (1975), Christie (1976), Lamb and Vanderslice (1976), and Lockwood (1977) have all proposed reanalyses of Halle's Russian data, and Sadock (1976) and Sommerstein (1977) offer methodological critiques. Most of these discussions have been somewhat negative in tone, and given the rather widespread agreement¹ among contemporary phonologists concerning the success of Halle's argument, as well as the considerable use of arguments roughly of the form of Halle's in later linguistic arguments, it seems to me that this argument deserves further methodological examination. This paper will present such an examination (section 1), together with a discussion of the above mentioned reanalyses of the Russian data (section 2) and methodological critiques (section 3) in the light of the methodological reanalysis presented in the first section.

1. A methodological reanalysis of the "Hallean syllogism".

It is worth considering Halle's argument itself in some detail here; neither Sadock nor Sommerstein cites actual passages from Halle, and it is perhaps not unlikely that this is the reason why they have (in my view) not fully understood the structure of Halle's argument.

Halle (1959:21-3) argues that the level of the (classical) phoneme should be rejected as a valid level of linguistic structure. His argument is essentially that the following "requirement has played a particularly important role in the development of American linguistics" (numbering Halle's):

Condition (3a): A phonological description must include instructions for inferring (deriving) the proper phonological representation of any speech event, without recourse to information not contained in the physical signal.

If this condition is accepted, Halle maintains, it follows that in Russian, where "voicing is distinctive for all obstruents except /c/, /č/ and /x/, which do not possess voiced cognates, we would have to give the following analysis of voicing assimilation (note that in Russian, essentially,² all obstruents are voiceless word-finally "unless the following word begins with a voiced obstruent, in which case they are voiced"): the phonetic forms [mók l, i], [móg b̄], [žéc l, i] and [žéj̄ b̄] would be represented phonologically as /mók l, i/, /móg bi/, /žéc l, i/ and /žéj̄ bi/, respectively. He continues as below:

- (1) Moreover, a rule would be required stating that obstruents lacking voiced cognates--i.e. /c/ /č/ and /x/--are voiced in position before voiced obstruents. Since this, however, is true of all obstruents, the

net effect of the attempt to meet both Condition (3) and (3a) would be a splitting up of the obstruents into two classes and the addition of a special rule. If Condition (3a) is dropped, the four utterances would be symbolized as follows: {mók l, i} {mók bi} {žéc l, i} {žéc bi},³ and the above rule could be generalized to all obstruents, instead of only {č} {c} {x}. It is evident that Condition (3a) involves a significant increase in the complexity of representation...If Condition (3a) can be dispensed with, then there is also no need for the 'phonemic' representation.

It is my contention that, contrary to Sadock and Sommerstein, there is nothing at all wrong with the form of this argument, and that if we accept that Halle maintained implicitly what would undoubtedly be for him an obviously true minor premise, this argument has the form roughly of classical modus tollens.⁴ It will be reconstructed as a first approximation, as in (2).

- (2) If "Condition (3a)" (roughly, classical phonemics) is adopted, then a significantly complex representation results. The representation should not be so complex. Therefore, Condition (3a) must not obtain.

As just suggested, the minor⁵ premise in this reconstruction is never explicitly stated by Halle, and should not be so complex, but it is fairly clear from the rest of his discussion of this issue that he would subscribe to this view. In fact, however, the force of the argument probably comes not so much from the relative complexity of the representations required as from the "splitting up of the obstruents into two separate classes." Although Halle again is not explicit on the matter, this too is something which should not be done. That is, the result of Condition (3a) is treating a unitary phenomenon as two separate phenomena. Let us reconstruct (2) in these terms, then, as (2'):

- (2') If Condition (3a) is adopted, Russian voicing assimilation must be two separate phenomena.
Russian voicing assimilation is a unitary phenomenon.
Therefore, Condition (3a) must not obtain.

Thus, if we accept the first two statements in (2'), we must conclude that Halle's argument is quite successful in demonstrating the weakness of theories which entail a classical phonemic level of representation. What is especially interesting about this argument, it seems to me, is that most phonologists appear to agree about Halle's assessment of the situations as embodied by the premises in (2')⁶ (further discussion of this point is given in the next two sections).

It is worth pointing out that Halle's argument does not establish (even if we accept his premises), namely, the lack of any level between Chomsky's (1964) "systematic phonemic" level and his "systematic phonetic" level. Thus, Halle's argument is not relevant to theories which posit

an intermediate level which does not correspond to the classical phonemic level,⁷ including theories which make use of (variants of) Praguean archiphonemes or Firthian prosodies, or which include an additional level abstract enough to allow for Halle-type representations at this level, such as the level of lexical representation in the "natural phonology" of Stampe (1973). In view of the comment in Halle (1959:21n), it would appear that Halle, unlike Chomsky (1964, 1966, 1967), never intended to establish such a further conclusion via this argument.

We need not accept Halle's premises, of course. What is peculiar about this argument, again, is that almost everybody has accepted it (with the qualification given in note 6), even Halle's critics. The remainder of this paper will be devoted to a consideration of the issues which have in fact been raised with respect to Halle's argument by his critics.

2. The reanalyses.

There have been several attempts to reanalyze Halle's data so that an intermediate 'phonemic' level can be maintained, mainly by stratificational grammarians, although the nonstratificationalists Johns (1969) and Christie (1976) have also presented sketches of reanalyses. It is noteworthy that all of these reanalyses explicitly depart from classical phonemic theory. Thus, Lamb (1966) and Sullivan (1975) make use of what appears to be a variant of Firthian prosodic analysis, while Johns (1969) and Lockwood (1972) seem to appeal to something along the lines of a Praguean archiphoneme.⁸ That is, none of these linguists has argued that the solution that Halle presented as the classical phonemicist solution is in fact the correct one. Moreover, they all appear to agree that the solutions that Halle argues against is in fact the standard classical phonemic analysis. Lamb (1966:544), for example, states that "the phonemic analysis which Halle criticizes is the traditional one." The fact that all of these investigators, despite their opposition to Halle's solution, agree on these points (i.e., essentially that the premises in (2') are true) would appear to provide strong support for the claim made above that most linguists agree with the premises of Halle's argument. Furthermore, I know of no published work that would indicate that its author disagrees with these premises.⁹ However, as the stratificationalists point out, since their analyses do not follow classical phonemic principles, Halle's argument is not relevant to these reanalyses.

3. Sadock on the "Hallean syllogism".

I will discuss for the most part only Sadock's critique here, since it is more strictly methodological than Sommerstein's, and is considerably more complete (Sommerstein states, for example (1977:121), that Halle's conclusion does not follow from his premises without ever giving what he takes Halle's conclusion to be.) Sadock's critique concerns both Halle's original argument and later arguments which have the form of Halle's, although only the former is discussed in detail. He intends (1976:85) "to establish that arguments of the form of Halle's should not be used in the way that they have been." Sadock gives an "outline" of Halle's argument (pp. 85-6), repeated below in (3):

- (3) ...Halle showed that autonomous phonemics imposed on the grammarian a treatment of Russian in which two separate, but complementary, voicing assimilation rules are required.

In a grammar without an autonomous phonemic level, however, he showed that it was possible to describe the voicing alternations in Russian in terms of one general, and hence simple, voicing assimilation rule. From these facts Halle concluded that there is no level of autonomous phonemics.

He goes on to the effect that "as the argument stands, this conclusion is clearly a nonsequitur." However, two sorts of "background assumptions" would be able to patch up the argument: "(a) that we have a priori knowledge that the general solution is correct in Russian, or (b) that general descriptions are always the correct descriptions of selected data in natural language." He then rejects (1976:86-8) both of these assumptions, and ends up by deciding (p. 88) that "the most that can be made to follow from Halle's argument is the very much weaker conclusion that (all other things being equal) the theory in which the general solution is possible is to be preferred," perhaps on the grounds that it is more "falsifiable" in Popper's (1965) sense. He concludes that Halle's analysis is indeed more falsifiable than the "phonemic" analysis, and therefore (p. 91) "should be examined as a working hypothesis before the less general treatment is" on the basis of the following considerations (pp. 89-90). Halle's theory is essentially that all obstruents assimilate in voicing to following obstruents, while the "phonemic theory" is that "some obstruent morphemes assimilate in voicing to following obstruents" and "some obstruent phonemes occur as the voiced allophones before voiced obstruents." But if Russian were just like it is except that "[c], say, failed to alternate, Halle's theory would be disconfirmed...", while the phonemic theory would not. Hence, "the power of Halle's argument, it seems to me, lies partly in the gross difference in testability between his treatment and the phonemic treatment." A final bit of evidence (p. 91) for his "contention about the lack of force of the Hallean syllogism" is that neither side in the generative semantics-lexicalist debate recognizes the other's arguments as being damaging to its position, despite the fact that these arguments often are of the form of Halle's.

What Sadock apparently means, in terms of the reanalysis presented above, when he rejects "background assumption...(a)" is that, roughly, he rejects the minor premise in (2'). (He evidently finds nothing to quarrel about in the major premise.) But, as noted above, he also makes a "contention about the lack of force of the Hallean syllogism," which I take as meaning he is unhappy with the form of the argument (cf. also the first quotation above). These are two separate issues, although it seems to be Sadock's unwillingness to accept Halle's premise that leads him to his conclusions about the argument's force, and I will attempt to keep them apart insofar as possible in the following discussion.

As I pointed out above, Sadock appears to be very much in the minority as far as his misgivings about accepting the minor premise are concerned. His reasons for rejecting "background assumption" (a) above (which would be essentially our minor premise if we replace "the general solution" by 'a general solution') are that (p. 86) accepting it "would reduce linguistic analysis to vacuity since, if we had a priori knowledge of the correct description of natural language facts, all we would need to do in describing a language would be to examine our intuitions as to the correct description." But Sadock has oversimplified the issue, and also

appears to be confusing his "background assumptions" (a) and (b). What is required to establish our minor premise is not a priori knowledge but a reasonable degree of certainty (cf. note 4), and not about correct analyses, but about incorrect ones (i.e., any analysis which makes it appear as if Russian voicing assimilation is two processes is incorrect). This says nothing more than that linguists have learned something by their investigations of language (note also that this makes the "knowledge" involved not a priori, but a posteriori). I can see no reason why having fairly clear intuitions about what kinds of analyses are incorrect in certain cases "would reduce linguistic analysis to vacuity," or even why such intuitions about correct analyses in a small number of cases should have this effect (although this kind of case would presumably be much rarer than that involving incorrect analyses). It is important to note in this regard that it is not necessary to have any particular degree of certainty about the correctness of Halle's analysis, as the existence of alternative analyses indicates--all that Halle's argument requires is a fair degree of certainty about the incorrectness of the classical phonemicist solution (or any solution which makes Russian voicing assimilation appear to be two separate phenomena).

Having rejected this premise, however, Sadock is faced with the fact that Halle's argument has nonetheless been found quite convincing by most linguists. It is evidently this fact which led him to analyze it in terms of falsifiability. But the relative degree of falsifiability of the two analyses appears to have little, if anything, to do with their acceptability, as the following considerations indicate. Suppose that we change the "phonemic theory" slightly so that instead of Sadock's formulation we have something like 'all obstruent morphophonemes assimilate in voicing to following obstruents' (this is essentially Halle's formulation of the rule in the "phonemic theory"--see (1) above) and 'all obstruent phonemes occur as the voiced allophones before voiced obstruents.' This formulation will get the right results as long as it is not necessary to have a morphophonemic representation to correspond to every phonetic (and phonemic) representation, since we could simply not set up morphophonemic representations for forms involving the voiceless phonemes which do not have voiced counterparts. The first rule would then take care of 'morphophonemic' voicing assimilation and the second would take care of the rest. I know of no injunctions by classical phonemicists against such an analysis, and it appears to conform (at least as well as Sadock's does) to their actual practice in a fair number of cases. With this modification, the two analyses would appear to have the same degree of falsifiability; whatever would falsify one would also falsify the other. Yet, it seems to me, the revised analysis is no more tenable than that given by Sadock, and I suspect that Halle's critics (cf. section 2) would agree on this point. (At any rate, their reanalyses would lead one to believe that they would, since such an analysis has never been proposed by any of them). The reason is that it still makes Russian voicing assimilation look like it is two separate processes. Thus, it would appear that degree of falsifiability has nothing to do with the success of Halle's argument. Neither, it would seem, does simplicity (cf. Sadock's outline of Halle's argument), for the stratificationists do not appear to be at all bothered by the fact that they require two rules, one to get from the morphophonemic level to the phonemic level, and one to get from there to the phonetic level. What is important to them, and to Halle's argument, is that voicing assimilation is done by one rule.¹⁰

The difference between the degree of success of Halle's argument and the syntactic arguments cited by Sadock is undoubtedly due to the fact that the general agreement about the premises in Halle's case is not present in the case of these other arguments. This should not seem terribly surprising--we know a good bit more about phonology than we do about syntax.

Footnotes

*This paper is a revised version of a section of my Ohio State University Ph.D. dissertation (Churma 1979). I would like to thank Fred Householder and David Stampe for their helpful comments on a preliminary version.

¹As will be seen below (cf. especially section 2), this agreement is probably more widespread than might at first glance appear to be the case.

²There is a slight complication (Halle 1959:63): "{*v} functions as a sonorant and as an obstruent if followed by an obstruent." (*v) represents the incompletely specified version of the morphophoneme {v}).

³The braces are used to denote what later came to be called "systematic phonemic" representations (cf. Chomsky 1964).

⁴Actually, the argument form probably should be what I have termed "almost modus tollens", since the premises in the argument are known only probabilistically, and not with certainty (cf. Churma 1979, Ch. 2 for details). For purposes of exposition, I will treat the argument as an instance of actual classical modus tollens; nothing crucial appears to be lost by such a simplification.

⁵This lack of explicitness in the statement of the premises of arguments presented is apparently not at all uncharacteristic of arguments given by working linguists (cf. Churma 1979), or, I suspect, those given by practitioners of other sciences.

⁶Actually, condition (3a) is probably not sufficient to require the analysis that Halle presents as that embraced by classical phonemics (cf. the references cited in section 2 for discussion). However, this is not crucial for Halle's argument to be considered a modus tollens-like refutation of the brand of classical phonemics which would advocate an analysis like the one presented by Halle. There is some condition, which apparently has not been explicitly formulated by its advocates, which would (be likely to--cf. note 4) entail such an analysis. Halle's argument then suffices, if we accept his minor premise, to refute any theory which embraces this analysis. For purposes of ease of presentation, I will continue to act as if it is condition (3a) which is responsible for this analysis.

⁷The use of the definite article here may be somewhat misleading, since there were after all several versions of what might fairly be called "classical phonemic theory." The "long components" of Harris (1951), for example, would not be countenanced by Bloch, Trager, Smith, etc.,

and yet all of these could be considered classical phonemicists, as could the Praguians. Not all of these theorists would advocate the analysis which Halle argues against, so "the" classical phonemic theory should be interpreted in the context of this discussion as that version which would advocate the analysis in question, i.e., that of Bloch et al.

⁸Lockwood (1977) claims that his 1972 analysis did not make use of the archiphoneme, and that Sullivan's (1975) analysis did (Sullivan also makes the latter claim), but I feel that the interpretation I just gave is closer to the truth. Sullivan's analysis does have some archiphonemic characteristics but his postulation of a "phoneme" of voicing seems to put him closer to the Firthians than to the Praguians. Comparison of stratificational analyses with other types of analyses is complicated by the stratificationalists' use of "singulary" features (cf. Lockwood 1972) instead of the more or less standard binary features. Fortunately, the only crucial issue is whether or not the analyses in question depart from classical phonemic theory, and there appears to be general agreement that they do.

Chomsky's (1966, 1967) claims that Lamb's analysis is a notational variant of Halle's appear to be due to a natural enough misinterpretation of that analysis to the effect that the output of Lamb's "single rule" is the phonetic representation, rather than the phonemic representation as Lamb apparently intended. (There must then be, of course, an additional rule or rules which Lamb does not mention to derive the phonetic representation, which is apparently what led to Chomsky's confusion). Sullivan's analysis, which is quite similar in other respects to Lamb's, does in fact contain such a rule.

⁹The only possible exception of which I am aware is Ferguson (1962: 288), who presents some considerations in favor of the classical phoneme, and may be hinting that he does not feel that Halle's solution is "more natural" than the one he criticizes. However, Ferguson is not explicit concerning his views on the matter. (F. Householder (personal communication) also explicitly states that he does not accept Halle's minor premise.) It is also possible that Sadock rejects Halle's minor premise (see below).

¹⁰Whether or not voicing assimilation is in fact a single phenomenon is in principle an empirical question. Thus, if the voiceless obstruents without voiced counterparts could be shown to behave differently than the others (say in loan words), there would be good reason to suspect that what is involved is not a unitary phenomenon at all. I know of no discussion in the literature along these lines, however.

References

- Chomsky, N. (1964) Current issues in linguistic theory. In Fodor and Katz, *The Structure of Language*. Englewood Cliffs, NJ: Prentice Hall.
- Chomsky, N. (1966). *Topics in the Theory of Generative Grammar*. The Hague: Mouton.
- Chomsky, N. (1967). Some general properties of phonological rules. *Language* 43:102-128.

- Christie, W. (1976). Another look at classical phonemics. *Language Sciences* 39:37-9.
- Churma, D. (1979). Arguments from external evidence in phonology. Unpublished Ph.D. dissertation. Columbus, OH: The Ohio State University.
- Ferguson, C. (1962). Review of Halle (1959), *The Sound Pattern of Russian*, The Hague: Mouton. *Language* 38:284-298.
- Harris, Z. (1951). *Methods in structural linguistics*. Chicago, IL: University of Chicago Press.
- Johns, D. (1969). Phonetics and generative phonology. In Binnick, et. al., eds., *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*. Chicago, IL: Chicago Linguistic Society.
- Lamb, S. (1966). Prolegomena to a theory of phonology. *Language* 42:536-573.
- Lamb, S. and R. Vanderslice (1976). On thrashing classical phonemics. In P. Reich, ed., *The second LACUS forum 1975*. Columbia, SC: Hornbeam Press.
- Lockwood, D. (1972). *Introduction to Stratificational Grammar*. New York: Harcourt Brace Jovanovich.
- Lockwood, D. (1977). Phonological "neutralization" in classical and stratificational theories. Paper presented at the winter meeting of the Linguistic Society of America.
- Makkai, A. and V. B. Makkai, eds. (1975). *The first LACUS forum*. Columbia, SC: Hornbeam Press.
- Popper, K. (1965). *Conjectures and refutations*. New York: Harper and Row.
- Sadock, J. (1976). On significant generalizations: remarks on the Hallean syllogism. In J. Wirth, ed. *Assessing Linguistic Arguments*. Washington, D.C.: Hemisphere Publishing.
- Sommerstein, A. (1977). *Modern Phonology*. Baltimore, MD: University Park Press.
- Stampe, D. (1973). *A Dissertation on Natural Phonology*. Unpublished Ph.D. Dissertation. Chicago, IL: University of Chicago.
- Sullivan, W. (1975). Alternation, transformation, realization, and stratification revisited. In Makkai and Makkai (1975).

Some Further Problems for Upside-Down Phonology

Donald G. Churma

Leben and Robinson 1977 (henceforth L&R) have presented a theory of phonology in which phonological rules function, not to derive surface pronunciations from underlying representations as in standard generative theories, but 'to permit the morphology to relate words that superficially are phonetically dissimilar' (L&R:1). This is accomplished by 'undoing' the rules of standard treatments, subject to certain constraints specified by L&R. This theory, which is termed 'upside-down' phonology (hereafter UDP), is claimed to have attributes which make it significantly more attractive than other currently proposed theories. In this paper, it will be argued that, from a synchronic perspective, UDP is both too weak, in that it either does not allow words which should be related to be related or is forced to treat an apparently uniform phenomenon as two or more separate phenomena, and too strong, since it allows words to be related which should not be. It will also be argued that UDP, at least in a form which can handle certain synchronic facts which it could not otherwise account for, does not have many of the diachronic advantages claimed for it. Finally, it will be suggested that psycholinguistic considerations present significant difficulties. Before proceeding to these arguments (sections 3, 4 and 5), however, I will first present an outline of the theory in its various manifestations (section 1), and then give a brief discussion of some previous criticisms of UDP, especially the long critique given in Janda 1980 (section 2).

1. The theory.

According to L&R (p. 1), 'the central function of phonological rules is to answer this question: "Given two words whose pronunciation and meaning are in the lexicon, are A and B related morphologically?"' Thus, in the case of their example sane/sanity, the phonological rules apply so as to allow these words to be related by morphological rule (1) (cf. L&R:1).

- (1) Word-Formation Rule: -ity Attachment

[ADJ - ətiy]_N

By altering the phonological shape of the stem, the phonological rules eventually allow sanity to be parsed as sane plus -ity. L&R's conventions on rule application (p. 2) are given in (2), and a sample derivation relating sane and sanity in (3).

- (2) Conventions on rule application.

- a. If, in a conventional generative treatment, a form is derived by three rules A, B, C, applying in that order, they apply in our account in the reverse order, C, B, A, except as provided by (2c) below.

- b. A rule of the form $X \rightarrow [-F]/Y_Z$ is undone by replacing $[-F]$ with $[+F]$ on segment X in the environment Y_Z . Analogously, a rule of the form $\emptyset \rightarrow X/Y_Z$ is undone by deleting X from the context Y_Z .
- c. A rule is blocked if undoing it would not increase the compatibility of forms A and B with respect to Word-Formation Rule R.

(L&R:2)

(3) Sample upside-down derivation of sane/sanity (adapted from L&R:3):

	Word A	Word B	Morphology
LEXICAL FORMS	[sēyn] _A	[sænətīy] _N	[ADJ-ətīy] _N
a. Vowel Shift	[sæyn] _A	_____	_____
b. Diphthongization	[sæn] _A	[sænətī] _N	[ADJ-ətī] _N
c. Laxing	_____	[sænətī] _N	_____

Once stage (3c) is reached, the stems are phonologically non-distinct, and sane and sanity are recognizable as being related by rule 1.

As noted by Janda, condition (2c) is an extremely important part of this theory, especially since it is this condition which is mainly responsible for the ability of UDP to relate pairs which require exception features in standard theories 'without resorting to ad-hoc means' (L&R:4). Unfortunately, as stated, it is somewhat vague in that neither 'compatibility', as Janda has pointed out, nor 'form' has been explicitly defined.

To remedy the first vagueness problem, Janda (pp. 8-20) suggests two possibilities. The first would require (p. 8) that to increase compatibility 'would be either to increase the number of shared segments, either by changing one pre-existing segment to match another pre-existing segment, or by adding a new segment to match a pre-existing segment and/or to decrease the number of unshared segments, by deleting a pre-existing segment that has no match in the other form, in corresponding positions in the two forms A and B'. This interpretation is pretty clearly not what L&R intended, since several of their derivations violate it (cf., for example, the relating of [jōr] and [järi] on p. 13), and I myself had never even thought of this possibility until I heard Janda 1977. Since Janda also gives another example which requires violation of condition (2c) if compatibility is interpreted in this way, such an interpretation seems clearly undesirable.

Janda's other suggestion in this respect (p. 14) seems much closer to what L&R intended. Here compatibility would be defined 'in such a way as to allow individual feature-values to remain--or even to become--different so long as there is a rise in the overall number of shared feature-values'. Janda has some objections to using this "'weak" definition' (p. 15). First, it is 'intuitively repugnant' in that it entails that the Finnish words 'virsi and virren are more compatible because their [s]~[r] contrast has been altered to an [s]~[t] contrast'. Secondly, 'the only real reason for undoing' a rule which increases compatibility in this sense is to allow a later rule to be undone, thus requiring

considerable 'global' power if derivations in accord with this interpretation of compatibility are allowed. It is also (pp. 16, 18) 'ad hoc,' but must be adopted anyway because of the empirical inadequacies of the 'strong' version (p. 18). All that I can say in this respect is that I do not share Janda's intuitions about the repugnance of this interpretation--[t] does indeed appear to me to be more like [s] than [r] is--and that I therefore do not agree about the 'real reason for undoing' rules in cases of this nature. I cannot agree, either, that such an interpretation makes the definition ad hoc, although I am not entirely sure that I understand what Janda intends by this term here. I will thus assume 'compatibility' to be defined in terms of shared feature specifications, rather than whole segments, in the remainder of this paper.¹

It is also somewhat unclear what the term 'form' in condition (2c) should be taken to refer to; it could be either 'word' or 'morpheme'. L&R's elaborations on this point in the course of their illustrations of this condition do not seem to be of much help in resolving the vagueness; they in fact suggest both interpretations at different points in their discussion. Thus, they note (p. 3) that 'once we get to stage (3c), we see that words A and B have representations of the stem san- that are non-distinct...' (emphasis added in all of these citations). However, two sentences later, they state that 'convention (2c) enjoins us from bothering with [further rules--DGC], since further applications would not increase the compatibility of words A and B with respect to Rule 1'. On p. 4, the implication is again that it is morphemes, not words, which are involved: 'we have proposed that rules are blocked if they do not increase the similarity between two allomorphs in a derivation'. Their examples also appear to point in contradictory directions with respect to this issue; thus we find (pp. 8-9) that rules which affect affixes are undone in relating Ketill and Kötlum, but not in the case of jaki and jökli, which are claimed not to be directly relatable by L&R. But they are in fact relatable, as long as it is permissible to undo rules which affect affixes. According to Anderson (1974), whose analysis L&R follow, jökli 'glacier' (dat. sg.) is underlyingly /jak+ul+e/ (where -ul is some kind of stem-forming suffix), and the surface form is derived via rules of syncope (which deletes vowels in contexts which are not relevant to this discussion) and u-umlaut (whereby a becomes ö when followed in the next syllable by u), as well as the rule which accounts for the change of final e to i and is irrelevant here. It should be clear that undoing syncope on jökli increases its compatibility with jaki with respect to the -ul- suffixation rule (and the rules which account for inflectional affixes). And once it has been undone, u-umlaut can be undone, giving jak+ul+i, and the two forms can thus be identified as being morphologically related according to the WFRs in question. It is thus not at all clear what L&R intend in this respect. However, since there appear to be a significant number of cases which would require the 'word' interpretation (including many of those discussed below), it seems clear that this interpretation must be adopted.²

One further point deserves mention in this respect. As L&R have noted (p. 10), morphologically related words do not always involve only a single WFR, contrary to the implications of condition (2c). Such a situation is not restricted to 'polysynthetic language(s) with many layers of morphology built into a single word' (L&R:10), but will also occur (at least) whenever the words involved belong to an inflectional paradigm and have

phonologically non-null inflectional endings. While in most cases examined here the required extension of condition (2c) seems intuitively clear, in some cases the intended UDP interpretation is not at all obvious (cf. section 3.1.4 for further discussion).

Two refinements on the conditions should also be mentioned here. The first is somewhat trivial; Leben (1977 a, b, c, 1979) has expanded condition (2b) to include the obvious specification of how to undo a deletion rule-- the \emptyset output is replaced by the input. The modification suggested by Robinson 1977 is not at all trivial, however, and in fact entails a considerable shift in the class of possible phonologies from the UDP perspective. Robinson never explicitly formulates his revision (which apparently is intended to replace, at least in part, both conditions (2b) and (2c)), but what he intends seems clear enough. I repeat below in (4) Janda's (pp.25-6) reconstruction of Robinson's revision:

- (4) i. Check whether a given morphological rule M can apply to relate Words A and B. If it can, the derivation stops; if not, then--
- ii. Attempt to undo phonological rules, in the following way:
 - a. Check whether either or both of Words A and B is a possible output of the first relevant phonological rule (e.g., has on some segment/s in the correct environment, feature-values non-distinct from those present--explicitly or implicitly--in the structural change of the rule). If not, proceed to the next phonological rule and repeat this step; if so, then--
 - b. Substitute variable feature-values for the specific values of exactly those features on the segment identified as a possible output of the rule in question, which explicitly appear in the structural change of that rule.
 - c. Check whether the segment in corresponding position in the other form bears specific values for these same features. If not, return to (i) above; if so, then--
 - d. Substitute these specific values for the variable values of the features in the segment changed by (b). Then return to (i).

The intended effect of this revision, if it is not already clear from this statement, should become so when it is applied to actual examples below. But let us turn now to an examination of some previous critiques of UDP.

2. Some problems and some non-problems for UDP.

This section will be devoted to an examination of previous criticisms of UDP, especially that of Janda 1980. Some of the specific objections raised are found not to be serious problems for UDP, while others appear to be genuinely problematic for the theory. It is these latter objections which I will consider first.

2.1. Problems.

The first problem, discussed by Janda (pp. 16-7, 32-4), concerns an aspect of UDP not yet brought up. L&R, in connection with the alleged impossibility of directly relating jaki and jökli discussed above, suggest that it is also possible to relate forms indirectly by relating both of them to a third form (L&R:10). There is no question that such a provision will allow a considerable number of forms to be correctly related which could not otherwise be. But there is some question whether it can take care of all of the morphologically related words which cannot be shown directly to be so related. Particularly troublesome would be (Janda 1980: 17) 'a defective paradigm--one either riddled with accidental gaps, or adulterated by suppletion', as well as, I would add, paradigms which inherently have only two members (e.g., singular/plural). Although Janda has given no examples which would indicate that this is an actual empirical problem (and not just a potential problem), it seems clear that UDP will eventually have to deal with cases like this--there is no reason at all to believe there will always be a third form which 'comes to the rescue' when two forms cannot be directly related (cf. also the discussion in section 3.1.4 below). It should also be pointed out that allowing for such a possibility would make for a situation in which, as L&R put it in tentatively rejecting an alternative to condition (2a) (p. 8), 'the parsing procedure provided by the grammar would be much less determinate', since there is no guarantee that the 'correct' third form will be chosen on the first try.

This brings up a somewhat related problem, that of how to tell, in a determinate fashion, when two forms are not morphologically related (cf. Janda 1980:36-7). The problem is that whenever two semantically related words (e.g., depart/left) are not also morphologically related, it is possible that several morphological rules must be tried out in an upside-down attempt to relate the two words. This would be the case in instances where a language has different ways of performing a morphological process, depending on the (arbitrary) lexical class to which the lexical items in question belong, such as the rules which form past tense in English.³ In the case of depart/left, then, all of the past tense formation rules would have to be checked, attempting with each such rule to relate the two words by undoing the phonological rules of English. Clearly, in this case it would take some doing to show that depart and left are in fact not related by any of the past tense formation rules of English.

Janda also notes (p. 20) that, at least in the case of a hypothetical example he constructs (pp. 18-9), UDP 'must...have global phonological rules that look farther ahead than at their own output' (and not simply at their own output, as required by condition (2c)). This is true not only of this hypothetical example, moreover, but also of a real example which will be discussed below (section 3.1.1).

Another problem noted by Janda (pp. 41-3) concerns condition (2b) when certain kinds of neutralization rules are involved. The rule of English which reduces unstressed lax vowels to schwa, for example, creates severe difficulties for this condition if it is extended in the natural way to cover (rightside-up) rules which alter more than one feature value of the input (i.e., replace each feature value mentioned in the output by the opposite plus/minus value). Since schwa is, among other things, [-high, -low], this rule, when undone according to this procedure, would result in a segment which is [+high, +low]--a physical impossibility in the

system of Chomsky and Halle 1968 (as well as most others)--and it is certainly undesirable from the standpoint of UDP in any event, even if it were a physical possibility. This appears not to create a serious difficulty for the theory, however, since it is a trivial matter to revise condition (2b) so that rules are undone in a quite literal fashion--the output of this rightside-up rule is replaced by its input, whether it is a deletion, insertion, or feature-changing rule. The problematic aspects of L&R's original condition (2b) uncovered by Janda can thus be seen to be fairly readily remedied.⁴

Janda has also pointed out (pp. 46-8) an apparent inconsistency in L&R's claims about the 'complicatedness' of forms. On the one hand, they indicate (L&R:3) that obesity is not 'complicated', since it 'can be parsed at the surface without any consultation of the phonological rules', despite its 'exceptionality...with respect to laxing'. On the other, in the context of a discussion of historical change from the perspective of UDP, they claim (p. 19) that 'apparent underapplication of a rule is resolved by changing the deviant item to support the threatened rule'. That is, obesity is predicted to change (and thus introduce allomorphy), despite the fact that it is not 'complicated'. Even if the claimed uncomplicatedness of obesity and its susceptibility to change can be reconciled, there appears to be a problem concerning the diachronic predictions made here. The change in fact appears to be proceeding in just the opposite direction--the pronunciation [ōwbīysətīy] appears to be an innovative one, replacing older [ōwbəsətīy] (which is the only pronunciation given in the OED, and a variant pronunciation in Webster's Third). That is, the underapplication of laxing is being introduced, and not eliminated to 'support the threatened' laxing. (Thus (Janda, p. 48) 'it is...incorrect to claim that there is nothing complicated about a form like obesity and that [standard theory] errs in marking it as exceptional, and thus strongly predicting a change in its pronunciation'--the change, again, is apparently in the opposite direction). Such cases are not at all rare, and in fact the theory of 'suppletive' lexical representations of Hudson 1974, 1980, has been specifically designed to deal with them. But there are cases which are in accord with L&R's prediction, such as that of Swiss German umlaut (L&R:18) which is supported by a change in pronunciation, so Hudson's theory has its problems as well. In fact, I know of no theory which makes all the right predictions in cases like these, so this problem is not unique to UDP. But such cases do seem to make it less 'clear' (L&R:19) 'that the notion of opacity as a motivation for change finds a much more comfortable home in our theory than in the standard framework'.

The final objection raised by Janda (pp. 46-9) to be discussed in this subsection concerns the apparent failure of one of the diachronic predictions of the theory. L&R state (p. 19), as noted above, that 'apparent underapplication of a rule is resolved by changing the deviant items to support the threatened rule', but Janda adduces an example from Yawelmani where this seems not to be the case. It involves the notorious rule of vowel harmony (see Janda for references), with respect particularly to the passive-aorist suffix. This suffix appears phonetically as -it in most cases, but after stems with -u- and, crucially, some stems with -o:-, it shows up as -ut. Since UDP does not have to worry about exceptions to rules (cf. L&R:3-5, and below), the obvious way of accounting for the suffix alternation would be via a vowel harmony rule which rounds

and backs suffixal i when it follows a rounded vowel. Such a rule will allow the variant forms of this suffix to be related by simply undoing this rule whenever an -ut form is being compared with an -it form, and in the other cases, they can be directly related without undoing any rules. But, Janda points out, given such a synchronic analysis, there is 'apparent under-application' of the vowel harmony rule (it does not apply after some mid vowel stems), and UDP therefore predicts an increase in mid vowel stems which trigger harmony. But the actual change, given the rather meager amount of evidence which bears on this issue, appears to be heading in the opposite direction--there is now a mid vowel stem which takes a suffixal -i- as well as the earlier -u-. The only way out of this problem, Janda maintains, is to adopt a diacritic analysis in which only high rounded vowels trigger harmony. Such an analysis would undoubtedly not be acceptable to proponents of UDP. But there may in fact be another way out for UDP in this case: the vowel harmony rule could be restricted so that it applies only after high rounded vowels, and -u- suffixes after mid stem vowels could be related to the corresponding -i- suffixes not directly but via a third form which contains an -u- suffix after a high stem vowel. The two -u- suffixes would match without undoing any rules, and the one following the high suffix vowel could be related to the -i- suffix by a straightforward undoing of the restricted vowel harmony rule. It is not clear to me whether such an alternative would be acceptable to the proponents of UDP, since they nowhere to my knowledge discuss attempts to establish the morphological identity of affixal allomorphs (this whole issue appears to merit further inquiry in this respect). Robinson (1980:132) implies that the use of a third form might be limited to 'derivationally related forms', and would not be allowed for 'comparing the members of a single paradigm', and Leben (personal communication) has made comment, which suggest that he might subscribe to a similar view. The case at hand does not really fit into either of these categories, since the suffix in question is not a derivational one and the forms involved are not 'members of a single paradigm'. I know of no discussion in the literature which is relevant to the issue of the categories with which cases such as this are most closely related (assuming this is a legitimate question to ask), although it seems to me that affixes which show allomorphy are much less subject to pressure for change than stems within a paradigm. It is thus unclear whether the suffixal allomorphs in question could (or should) be related via a third form.

2.2. Non-problems.

While I agree with Janda that the above criticisms are, at least to some extent, genuine problems for UDP, this is not the case with respect to all of his criticisms. I will devote this subsection to a brief outline of the reasons for my disagreement; in so doing, I hope to clarify the real issues as far as an evaluation of UDP (or any other theory) is concerned, and thus to indicate why I feel that UDP deserves sufficiently serious consideration as a theory of phonology that it requires the further criticisms given below.

Janda's first criticism of UDP is that (p. 7) 'all rules in UDP must crucially be global', since condition (2c) requires that the potential output of a rule be examined before it can be determined whether or not it is applicable. This is quite true, and it is also true that rules in UDP are inherently transderivational (cf. Lakoff 1973), since their outputs must be compared to outputs in different derivations. It is not

at all clear to me why Janda brings up this point, since he claims (p. 8) that 'such "local" globality does not increase the class of possible grammars..., and so it is hard to characterize it as objectionable...'. He does not provide any arguments to support his claim about the class of possible grammars entailed by UDP, and since similarly glib statements concerning the relative generative capacity of various revisions of UDP can be found throughout, it is worth a brief digression to say a few words about this question.

Globality (and transderivationality) have gotten a lot of bad press recently (cf., for example, Baker and Brame 1972, Langendoen 1975), and it is tempting to believe that such characteristics are inherently bad. But the reason that they were held to be objectionable had nothing to do with their inherent qualities--the arguments go that they (unacceptably) increase the size of the class of possible natural languages. Such an increase, or lack thereof, must be demonstrated, and cannot be established by fiat. It is not always easy to provide such a demonstration, and if one cannot provide one, then the only rational thing to say in such a situation is that the relative power of the theories in question is not known. One should be quite clear, moreover, on what is required for a demonstration that one theory is more powerful (i.e., less restrictive) than another. Since linguistic theories typically generate an infinite number of possible languages, it must be shown not only that there are languages which the putatively more powerful theory can generate and the other cannot, but also that there are no languages which are generable by the latter and not by the former.⁵ In many cases, it will not be possible to do so, and the two theories may well be incommensurable (incomparable) in the mathematical sense. But they need not be incomparable from a linguistic point of view. As long as the theories make different claims about what the (infinite) class of possible natural languages is, they can be assessed on the basis of the correctness of their claims. Thus, in the case of incommensurable theories, if it can be shown that one of the theories can generate an impossible natural language, or cannot generate a possible (preferably actual!) language, while this is not the case for the other theory, then the latter is to be preferred to the former. It should also be pointed out that in cases where one theory can in fact be shown to be more powerful than another, it must also be shown that such an increase in power is unacceptable in that the additional languages generable are not possible ones.⁶ Janda has nowhere, as far as I can tell, addressed this issue directly. It will thus be the main focus of my own criticism of UDP from a synchronic perspective--I will argue that UDP is both too weak in that there are possible languages which UDP cannot generate, and too strong in that there are also impossible languages which UDP can generate.

Janda also fails to address this issue when he claims (p. 29) that the revised procedure for rule application suggested in Robinson 1977 ((4) above) has 'excessive power'. The only possibly relevant attempted justification for this claim (an irrelevant one will be discussed immediately below) is that this procedure 'reduces the derivation of Finnish virsi/virren to a single step...'. Since these two forms are clearly morphologically related, as Janda (pp. 10-1) agrees, I cannot see how this fact can show that UDP can generate impossible languages. This procedure, like L&R's original procedure ((2) above), does in fact appear to be too strong, as argued in section 3.2, but nothing in Janda's exposition would lead one to believe that this is the case.

A seemingly relevant observation by Janda (p. 29) in this respect is that 'the phonological rules in an UDP incorporating [4]...are in danger of relating scene to sanity. But this is only possible if it is not the case that, as Janda puts it (p. 30), 'one stimulates that a semantic comparison of some kind...is performed before any morphological or phonological rules in a derivation are undone...'. Such a stipulation is rejected by Janda, since 'the fact that semantics, and not phonological rules (by virtue of not being able to apply), is what is necessary to prevent the relating of, e.g., scene and sanity...must certainly count as a further strike against this version of UDP. But surely it is in fact semantic factors which are responsible for the knowledge that speakers of English have that scene and sanity are not morphologically related, and not phonological ones. And it is undoubtedly the case that whenever two words are semantically unrelated (at least, if they are as unrelated as these two), speakers of the language in question will judge them to be morphologically unrelated as well, and precisely because they are not related semantically. As the example cost/caustic, given by Leben 1979:185 (and cited by Janda), clearly shows, even phonological identity of what is putatively the same stem is not enough to establish morphological relatedness, and the only way of preventing these two words from being related is 'to equip...morphological rules with semantic characterizations that must be satisfied by words related by such rules' (Leben 1979:185). Neither can I see why Janda apparently feels (p. 29) that it would be desirable that 'semantics...not come into play in derivations in UDP until morphological rules can actually apply...'. Speakers are not even tempted to suspect that Janda's microorganism and lick are related, because of their semantic disparity, and so it is entirely appropriate that the grammar, which is a model of speakers' knowledge, not also be 'tempted' to undo Vowel Shift on the first vowel in this pair, again because of their semantic disparity. Far from being a defect of UDP, attributing the morphological unrelatedness in such cases to semantic factors seems clearly to be desirable, since it is just such factors which are behind speakers' judgments about morphological relatedness. In fact, any theory which claimed that scene and sanity were not related for phonological reasons would surely merit a good bit of scepticism.

McCawley 1979:295 raises what appears to be a similar objection (his example in moth/mother). It is thus subject to the same kinds of criticisms as Janda's objection, although I agree with McCawley that it might be worthwhile 'to consider the possibility that different morphemic identities may have different psychological status', and should thus have formally distinct representations in a theory of morphology.

Janda also discusses (pp. 33-5) what he feels to be a problem which is brought about by the possibility of relating two forms via a third form--this possibility (p. 35) 'greatly multiplies the number of incorrect derivations produced by UDP.' Thus, in the case of the Finnish triple virsi/virren/virsia, 'UDP will render the incorrect verdict that only the first and the last of these three forms are morphologically related' (if Janda's 'strong' version of condition (2c) is used). But there is no reason to believe that verdicts about unrelatedness in UDP should be arrived at as quickly as the above statement implies. A perfectly acceptable definition of unrelatedness is that there be no possibility of relating the forms in question directly and that there be no other form which serves to relate them indirectly. Under this interpretation, no verdict at all

about relatedness is reached until either a match is found (in which case the forms are morphologically related) or all (semantically related) words have been checked as possible 'third forms' and no match has been found (they are not related). There is thus no obvious problem related to (p. 34) 'generating all and only the correct output [sic]' which is due solely to the 'third form' possibility, and the only reasonable way of going about determining whether there is in fact a problem is the empirical one (cf. the discussion of globality above): does allowing 'third form' derivations correctly characterize the class of possible natural languages, or does it not? This question is simply not addressed when one flatly claims, as Janda does (p. 34), that 'for UDP to be an interesting theory of phonology, an unsuccessful derivation must mean that the forms in it are (all) morphologically unrelated...'

In a similar vein, Janda (pp. 22-4) criticizes the 'unconstrained' character of the proposal to abandon condition (2c) and to allow rules to be undone optionally, with an unrelatedness verdict given only when all possible combinations of actual applications of rules have been tried and found not to produce a match. He even claims (p. 22) that such a procedure 'effectively immunizes UDP against ever being faced with a counterexample'. In this case, Janda does attempt (p. 22) to give some justification for this claim, but it is not at all clear to me that this attempt succeeds. For example, as long as there is (p. 22) 'a linear list of phonological rules', then if this linear ordering entails that the undoing of one of the rules counter-feeds (cf. Newton 1971, Koutsoudas, Sanders and Noll 1974) another, and the counter-fed rule must be undone to relate some pair of forms, then the forms will not be relatable by this procedure--the environment of the counter-fed rule will never be met, since the counter-feeding rule cannot be undone until after the counter-fed rule has been. If such a situation should be uncovered, it would in fact be a counterexample with respect to a version of UDP which incorporates such a procedure. It should be pointed out in this respect that this kind of revision to the theory does indeed appear to make the revised version more powerful than that proposed by L&R, since the optionality of undoing rules appears to give the same effect as does condition (2c) in blocking 'bad' undoings, and the possibility of undoing rules even when compatibility is not increased allows for languages which could not be generated if condition (2c) is maintained. But in order for this to count as a defect of the theory, it must be shown that this increase in power is undesirable, i.e., that it permits the generation of impossible languages.

It is interesting in this respect that the kind of evidence mentioned above would not be a counterexample to the proposal considered (and tentatively rejected) by L&R (pp. 7-8), since the 'random ordering' (apparently intended to mean that forms are relatable if any order of the rules succeeds in relating them) suggested there always allows a potentially feeding relationship to be actualized--if some rule potentially feeds another in a given derivation, then it will feed it for some order of the rules. It thus may well be that Janda's suggested proposal is more restrictive than L&R's, although this remains to be shown (it has not been demonstrated whether the latter can generate languages which the former cannot--I suspect that this is not the case). Thus, even if the random ordering proposal is rejected, this does not imply that Janda's suggestion should be.

It is also worth noting here that neither L&R nor Robinson 1980 'reject the principle of random rule-ordering', (Janda, p. 22), but rather that

they only 'disfavor' it (Janda, p. 24). L&R, for example (p. 8), 'conclude that linear ordering, insofar as it is tenable, is desirable' (my emphasis). If linear ordering should be shown not to be tenable, then random ordering, or perhaps Janda's proposal, could be tentatively adopted. It is not clear to me why L&R are only tentative in their rejection of random ordering, since Leben 1979:183 and Robinson 1977:9 take some pains to point out that UDP is not (non-inherently) global, and unrestricted globality and random ordering appear to be equivalent in generative power from the perspective of UDP. (This has not been shown to be the case, but neither has any proponent of UDP shown that it is not the case; globality has indeed gotten 'bad press' (see above), and this is presumably the reason that Leben and Robinson have been so quick to renounce it.)

Another not terribly damaging problem discussed by Janda (pp. 40-1) concerns the 'abstractness' of UDP. Although L&R (pp. 5-6) mention this issue only indirectly, it is pretty clearly a direct concern of Robinson 1977, and so merits some discussion here. The final representations arrived at in an UDP derivation do indeed appear to be 'abstract' in many cases, and UDP is thus not much different from, say, Chomsky and Halle 1968 in this respect. What is more, a less abstract rightside-up theory (e.g., one incorporating an 'alternation condition--cf. Kiparsky 1973) would undoubtedly decrease any difference in abstractness between such theories and UDP. On the other hand, I cannot agree with Janda (pp. 26, 56-8) that the constraint against absolute neutralization brought about by Robinson's revised procedure for rule application ((4) above) is not more intrinsic to UDP than Kiparsky's 1973 'alternation condition' is to standard theories; if (4) is indeed intended to replace fully conditions (2b) and (2c), as it appears to be, then this constraint does in fact follow directly from independently motivated principles of the theory. That is, this decrease in abstractness does seem to be an integral part of UDP (or at least the Robinson 1977 version).

A related point is Janda's (pp. 39-40) discussion of the 'solid body of evidence that phonological theory must countenance at least some abstractness.' First of all, it is not at all clear exactly how solid this body is (Janda refers here to so-called 'external evidence'). The discussion in Sommerstein 1977, Churma 1979:ch. 5, and Manaster-Ramer 1980 (although the latter is undoubtedly overcritical) indicates that the kinds of evidence which Janda cites must be taken with a grain of salt. What is more, a theory of UDP such as that proposed in Pollack 1977a, b, and Leben 1979, in which lexical representations are fairly abstract--more so than classical phonemic representations--would take care of at least some of this evidence. Adopting such a theory, however, does result in diachronic problems, as will be shown below (section 4).

Similarly, the fact that some phonological rules must apparently apply productively and in rightside-up fashion (pp. 51-2) need not be problematic for a theory of the type just mentioned, as long as the inputs to the rightside-up rules are not too abstract (and in Janda's example, they apparently are not), although, again, diachronic problems would result from the adoption of such a theory.

3. Some further synchronic problems for UDP.

It should be clear from the discussion given above that UDP deserves serious attention as a possible theory of phonology. Furthermore, since

even the genuine problems discussed above appear not to be totally devastating ones, it might be tempting to some to (at least tentatively) adopt UDP. I will argue in this section that this should not be done; in particular, I will argue that UDP is inadequate as a theory of phonology (and morphology) because it fails to characterize appropriately the class of possible morphophonologies of natural languages. Thus, for example, it will be maintained that there are natural languages whose morphophonology cannot be adequately characterized within the theory (i.e., that UDP is too weak). It will also be maintained that there is at least one example for which UDP can provide a straightforward characterization which does not correspond to a possible morphophonology of a natural language (UDP is too strong). If UDP in its present form(s) does in fact fail in both these respects, then this at least suggests that minor modifications of the theory will not succeed in remedying such failures: any obvious decrease in the restrictiveness of a principle of UDP (thus alleviating the 'too weak' problem) would, if anything, aggravate the problem of being too strong, and vice versa.

This is not to say that no modification of UDP could succeed. In fact, the modification given in Robinson 1977 both allows some morphophonologies that the L&R version did not and does not allow some which the latter did (thus making the two versions incommensurable with respect to generative capacity). One could legitimately question whether such a radical revision really leaves us with the 'same theory' we started with,⁷ however, and I will thus not seriously consider the possibility of making such radical changes in the theory except when they have actually been proposed; to do so is an impossible task at any rate, since the number of possible such changes is infinite. Specifically, I will not consider the possibility suggested by Janda in several places of allowing exception features in UDP. I agree with Janda (p. i) that condition (2c) 'is crucial to (the spirit of) UDP', and since perhaps the principal claimed virtue of UDP is that this condition allows UDP to do away with exception features, the lack of such features is correspondingly crucial.⁸

One further issue deserves some discussion before I proceed to the task at hand, one which may appear to compromise this entire task. I will depend for the most part on specific analyses of various languages, and it might be objected that there is no assurance that these analyses are the correct ones (cf. Fn. 6). This is quite true, and perhaps unfortunate in the best of all possible worlds but in this world such a situation appears to be unavoidable. There simply are no neutral 'empirical data' which can be used to falsify UDP, or any other scientific theory (cf., for example, Kuhn 1970). Since a discussion of a large number of examples would clearly be impractical, I will limit myself to a relatively small number. I will thus depend, as any scientist must, on the assumed relevance of the 'data' discussed to the question at issue. I would expect, however, that the not at all insignificant number of examples offered, together with the existence of numerous parallel cases from other languages (specified in more detail below), in the following discussion will suffice to convince most investigators that my claims are in fact well supported.

3.1. UDP is too weak.

I will first offer several examples which indicate that UDP, in one or more of its actually proposed forms (and perhaps other versions as

well), is unable to provide a morphophonological grammar for all possible (in fact, it is claimed here, actual) natural languages. I would like to reiterate that more or less trivial loosening of the restrictions of the theory will undoubtedly not be relevant here, since this would result in an increase in the number of languages generable by a theory which can already generate impossible languages.

3.1.1. Let us first consider an actual case in which it appears that UDP cannot (directly) relate two actually related words without, as Janda puts it (p. 20) 'global phonological rules which look farther ahead than at their own output'. As pointed out above, both Leben and Robinson seem quite opposed to allowing this kind of globality, and at any rate, since allowing such globality would only permit the generation of further languages not previously generable, in addition to those generable by the original, doing so would, if anything, aggravate the problem that UDP is already overly strong.

The case in question comes from Icelandic, and was originally discussed in Anderson 1969, 1974. L&R:8-9 discuss this example as well, but fail to recognize that the relevant words can be related only by violating condition (2c), or by making it global in the sense that rules can be undone if compatibility is thereby increased at some indeterminate future stage of the derivation. The rules involved include u-umlaut and syncope (described in section 1 above), i-umlaut, which converts a to e when followed in the next syllable by i, and the l/r rule given below, together with the required derivation, in (5).

(5)	Word A	Word B	Morphology
LEXICAL FORMS:	[ketill] _{ns}	[kötllum] _{dp}	[N-r] _{ns} ; [N-um] _{dp}
a. <u>l+r</u> → <u>l+l</u>	[ketilr] _{ns}	_____	_____
*b. <u>i</u> -umlaut	[katilr] _{ns}	_____	_____
c. <u>u</u> -umlaut	_____	[katlum] _{dp}	_____
d. syncope	_____	[katilum] _{dp}	_____

The asterisk indicates a rule which can only be undone by violating the non-global version of condition (2c). Note that the starred stage (5b) causes the vowel affected to differ from the one with which it is being compared by three distinctive features, roundness, backness and height, whereas it only differed from its mate by one feature, roundness, before i-umlaut was undone. Note further that attempting to undo u-umlaut first will not help matters, since this would lower and back the ö, thus adding another feature difference to the already existing one.

One possible way of remedying this situation would be to require that all rules be undone simultaneously, if possible, checking after each set of simultaneous applications for compatibility until no more rules can apply.⁹ However, this approach would make it quite difficult to check to see if condition (2c) were being violated; all we would know after checking the output of a block of rules which resulted in a violation of (2c) is that at least one of the rules in this block was responsible

for the violation. We would thus have to continue trying all possible combinations of rules, leaving out suspected culprits, until we either found a combination which (2c) permits or had exhausted all possible combinations. This, of course, would make for a situation in which the parsing procedure would be, as L&R put it (p. 8), much less determinate. This approach, then, does not look very appealing.

Another possibility is to relate these words via a third word, and in this case one could in fact succeed in relating them via katli (dat. sg.). Since Robinson 1980:132 apparently wants to bar such a possibility for 'comparing the members of a single paradigm' (cf. the discussion in sec. 2.1 above), I will not pursue the matter further. Note also that allowing the use of this procedure in cases like this would take away the diachronic advantage of distinguishing 'leveling in paradigms and leveling across paradigms' which Robinson claims the theory to have.

It should be noted that even random ordering would not help here, as long as condition (2c) is retained. (Random ordering--or even linear ordering--would work if this condition were eliminated, but it is, as noted above, a crucial part of the theory.) The reason for this is that undoing either of the umlaut rules results in a, a vowel which is less like either ö or e than they are like each other. Thus, no matter what the order is, undoing an umlaut rule will result in a decrease in compatibility.

The procedure suggested in Robinson 1977 does in fact allow the unlauded vowels to be related, as illustrated in (6).

(6)	Word A	Word B	Morphology
LEXICAL FORMS	[ketill] _{ns}	[kötlum] _{dp}	[N-r] _{ns} ; [N-um] _{dp}
a. <u>l+r</u> → <u>l+l</u>	[ketilr] _{ns}	_____	_____
b. <u>i</u> -umlaut	[kVtilr] _{ns}	_____	_____
	[ketilr] _{ns}	_____	_____
c. <u>u</u> -umlaut	_____	[kVtlum] _{dp}	_____
	_____	[ketlum] _{dp}	_____
d. syncope	_____	[ketilum] _{dp}	_____

To 'undo' i-umlaut according to procedure (4) above, we first substitute variable coefficients for those features mentioned in its output, here [-low, -back]; I indicate the result of this operation by V, no matter what features have variable coefficients (I trust that no undue confusion will result from this practice). Comparing the corresponding vowel in the other word, we find that it is [-low, -back], and so replace both variables by a minus sign, thus converting the 'archivowel' (back) to e. To undo u-umlaut, three features must be given variable coefficients, since its output would be [-low, -back, +round]. Examining the corresponding vowel in the first word, we find the specifications [-low, -back, -round], and so substitute these values for the variables, with the result being e.

But undoing rule (6a) appears to violate the conditions of this procedure, since it requires (Robinson 1977:7-8) that 'there must be evidence, from the items being compared, for the insertion of a specific feature value different from that found in a given segment...' (my emphasis; cf. also 4.iii above). Here, however, the 'evidence' which would permit the

replacement of [α lateral] by [-lateral] (i.e. taking surface l back to r) comes not from the items being compared, but from what is called for by the WFR's in question, and thus the full undoing of this rule is blocked. One might suggest that I am being over-literal in my interpretation of Robinson's intentions in this respect, and that it was just an oversight on his part that he failed to mention the possibility of using evidence from WFR's to replace the variable feature coefficients. But this is apparently not the case, since not permitting such evidence is crucial to successfully ruling out the undoing of the rules which this procedure is designed to rule out.

Let me illustrate this with one of Robinson's examples, based on the analysis of Nupe suggested in Hyman 1970. This analysis, which Robinson wishes to show is not permitted by UDP (in its corresponding upside-down form, of course), contains rules (7) and (8) below (cf. Robinson 1977:8).

$$(7) \begin{bmatrix} V \\ +low \end{bmatrix} \rightarrow \begin{bmatrix} +back \\ -round \end{bmatrix}$$

$$(8) \begin{bmatrix} C_i \\ -cons \\ +syl \\ 2tone \\ +high \\ \alpha round \\ \alpha back \end{bmatrix} \rightarrow \begin{bmatrix} C_i \\ V \\ \alpha round \end{bmatrix} X]_V]_N$$

Rule (7) is a phonological rule of 'absolute neutralization' which converts o and e (and a, vacuously) to a, and rule (8) is a WFR which forms participles from corresponding verbs by reduplicating the initial consonant cluster and producing a mid toned high vowel with the backness and roundness features of the vowel which follows this cluster (i.e., i if it is unrounded, u if it is rounded). But when Robinson discusses the attempt to relate [t^wá] 'to trim' and [t^wütwā] 'trimming', which are putatively related by rule (8), he claims (p. 8) that, by virtue of rule (7), 'we may in fact substitute [α back] and [β round] on the [á]'s' of these forms. However, 'we cannot go on to insert the values [+back] and [+round], since there is no evidence for those specific feature values on those vowels', despite the fact that 'the morphological rule...makes crucial use of the feature [+round]'. These remarks can only be interpreted, as far as I can see, to mean that the required 'evidence' in Robinson's revised procedure cannot come from WFR's; if it could, then [+round], at least, could in fact be substituted, since it is crucially called for by rule (8). Thus, Robinson's revised procedure for undoing rules appears to encounter very real difficulties. They do not involve its 'excessive power,' as Janda puts it (see section 2.2 above), at least in cases like that from Icelandic just discussed, but in fact just the opposite--there are languages which clearly should be generable by the theory, but cannot be generated as long as procedure (4) is maintained. And the reason for these difficulties is not (Janda 1980:31) that this procedure would 'deprive the notion "phonological rule" of its essential content as a pairing of input and output...' (and note that there is no reason necessarily to expect the 'content' of this notion to remain unchanged in a theory which posits the radically different function for phonological rules), but rather that it would deprive phonological rules of the possibility of increasing the compatibility of two forms with respect to a given WFR whenever such an

increase results from altering a segment so that it looks more like what is called for by this WFR. This is thus a quite general problem for this procedure; the example given would not be an isolated one, and in fact this kind of problem would arise whenever, in a rightside-up treatment, a phonological rule affects an affix. Note also that making the obvious revision, so that the WFR's can also be checked to see if there is any increase in compatibility, would allow not only the l/r rule to be undone, but also some rules¹⁰ of absolute neutralization, the banning of which was the motivation for the revised procedure. It thus seems clear that Robinson's revision is no improvement over 'standard' UDP, and that the revised revision just suggested violates the spirit of Robinson's proposal, since it does not rule out absolute neutralization rules. What is more, both of these revised versions of UDP are still too strong (i.e., can generate impossible languages), as will be shown in section 3.2. This example thus appears to be quite damaging for UDP in any currently proposed version, as well as in several suggested further revised versions. It should be noted that there is no reason to believe that this example is an isolated one (although I cannot at present suggest any further examples of this type), since, as Janda (pp. 24-5) has pointed out, rightside-up derivations need not always decrease the compatibility of morphologically related words, and it is such a decrease which causes undoing a rule to increase compatibility.

3.1.2. The next example involves the interaction of a rule of reduplication and two phonological rules in Tagalog. This example has been fairly widely discussed (cf. especially Wilbur 1973, Anderson 1975, Carrier 1975 and Herbert 1977), and not all investigators agree as to the formulation of the phonological rules involved. The differences are not crucial to this discussion, however, and I adopt here Herbert's quite convincing analysis. The phonological rules involved, then, will be one of nasal assimilation, which assimilates a nasal to a following obstruent with respect to point of articulation, and one of obstruent deletion, which deletes an obstruent when preceded by a nasal. The interesting thing about these processes is that in rightside-up theories the morphological rule of reduplication (which copies the first vowel of the stem and the consonant which immediately precedes it) appears to apply to the output of the phonological rules, so that /maŋ+bigay+REDUP/ 'give' (fut.), for example, becomes first /mamigay+REDUP/, and finally [mamimigay]. Consider now what an upside-down derivation involving these rules, and relating mamimigay and bigay, would look like, given in (9).

(9)	Word A	Word B	Morphology
LEXICAL FORMS:	[bigay]	[mamimigay]	[maŋ+C _i V _i + [C _i V _i X]] _V fut
a. Obstruent del.	_____	[mambimigay]	_____
b. Nasal assim.	_____	[maŋbimigay]	_____

The WFR in question forms the future of verbs by prefixing maŋ and copying the first CV of the verb stem. Note that, even after both of the relevant phonological rules are undone,¹¹ the two words cannot be related, and there would be no other rules which could convert the last m back to a b, given a system which is even close to traditional rightside-up accounts.

This sort of problem will arise, it seems to me, whenever a morphological rule gives the appearance of applying to the output of a phonological rule in a rightside-up treatment (cf. note 12 below for examples from the literature). The only way I can see of getting around this problem is to allow morphological rules to be undone (perhaps, following Aronoff 1976, only in certain places in the block of phonological rules) as well as phonological rules. This would entail a considerable modification of the theory, since currently there is a radical difference between phonological rules and WFR's in UDP, with the latter being 'redundancy rules' in the sense of Jackendoff 1975. Since the latter do not apply generatively, they cannot be (generatively) undone either.¹² It should also be noted that such a revision would occasionally lead to stems which never occur in isolation on the surface, as in Janda's (p. 13) example *pi- from pious/piety, and the status of such a representation in UDP is quite questionable, as it is in rightside-up theories which entail such representations (cf., for example, Halle 1973, Jackendoff 1975, Aronoff 1976). What is more, as Janda suggests in several places, allowing such a possibility (at least together with other revisions such as getting rid of condition (2c), etc.) could make for a theory which is essentially a notational variant of standard theories. Thus, examples of this type appear to present significant difficulties for the theory.

3.1.3. UDP will also encounter difficulties whenever what appears to be a single rule applies both word-internally and across word boundaries in a rightside-up treatment. Since L&R state (p. 2) that 'only fully regular morphological and phonological processes will apply in the formation of new words from existing words, or in the derivation of external sandhi variants', there is no explicit provision in their theory or any elaboration of the theory by Leben or Robinson that I know of (but cf. Pollack 1977a for an extensive discussion of this issue), for the treatment of external sandhi phenomena when the processes involved are not fully regular.¹³ It is not clear to me why they have restricted the operation of external sandhi rules in the way that they have, but even in the cases which they have provided for, the theory appears to encounter serious difficulties. The problem arises when, as noted above, what would appear to be a single rule in rightside-up theories applies in both internal and external sandhi environments, since (if it is 'fully regular') it will apply rightside-up in the external sandhi cases, since all words are entered in the lexicon in their 'surface representations' (p. 1). Note that L&R apparently do not want to allow for rightside-up derivations of existing words even if the processes involved are fully regular, since such processes are said to apply in the formation of new words (see the above quotation). We thus seem to be forced by the theory to treat Russian voicing assimilation, for example, as two separate phenomena,¹⁴ one when it occurs word-internally, and another when it occurs across word boundaries. Thus, while rightside-up theories would analyze the phonetic forms in (10) (data from Sullivan 1975 and Halle 1959) by positing the corresponding underlying representations in (10) and deriving the phonetic forms by a rule of voicing assimilation, UDP would have to enter the first two forms in the lexicon in their phonetic representations and undo the voicing assimilation rule in (10b) to relate the two stems.

- (10) a. [podojtʲɪ] 'approach' (perfective): /pod+ojtʲɪ/
b. [potxodʲɪt] 'approach' (imperfective): /pod+xodʲɪt/
c. [mokl,i] 'was (he) getting wet': /mok#l,i/
d. [mogbʲɪ] 'were (he) getting wet': /mok#bʲɪ/

In the case of the last two forms, however, it is not possible to enter them in the lexicon in their phonetic forms, since they are made up of two separate words, and therefore each requires two separate lexical entries (one for each word), as they would in rightside-up theories, with the voicing assimilation rule applying rightside-up in such cases to derive the phonetic form, again as in rightside-up theories. Thus, although the 'same rule' is involved in both types of examples, the fact that it applies in 'normal' rightside-up fashion in one of the cases, but has to be undone in the other, indicates that the analysis is making the claim that there are really two different phenomena going on here, one which is handled in the traditional way, and one which is treated by undoing traditional rules. Such cases do not appear to be at all rare. English flapping, for example, seems to be another obvious candidate for an example of this type, since it too occurs both word-internally and across word boundaries.

Pollack 1977a, 1977b has also noticed this problem, and has proposed that lexical representations in UDP should be made considerably more abstract than L&R propose, so that they correspond to the lexical representations of 'natural phonology' (cf. Stampe 1973, Donegan and Stampe 1979).¹⁵ If this approach is adopted, then the cases just discussed can be treated just as in rightside-up theories, since the lexical representations would then be the same as in such theories, and the phonetic forms would be derived by the live 'natural processes' of voicing assimilation in Russian and flapping in English. It is only those alternations which cannot be characterized in terms of natural processes (i.e., are due to the operation of rules in Stampe's sense), Pollack proposes, which are to be handled in upside-down fashion. But, while this approach would allow these phenomena to be treated in the unitary manner which seems to be required, its good points are not unaccompanied by bad points as far as UDP is concerned, as will be argued below (section 4).

Furthermore, even adopting the Pollack/Leben revision cannot take care of all problems of this type. Rather, any case which requires a lexical representation which is more abstract than those allowed for by natural phonology still would appear to present a problem for UDP. The particular case to be discussed here, French liaison, is especially troublesome in that it does not seem possible to give an analysis similar to that given for the Russian case. Here the standard rightside-up rule must apparently apply upside-down to produce the sandhi variants, and it must do so in the same way that rules apply in rightside-up theories--generating a phonetic form which differs from its lexical representation--a possibility which is not envisioned in any treatment of UDP which I am aware of, and one which appears to be quite undesirable as long as rightside-up rules can also apply productively in rightside-up fashion as well. It should be noted that this kind of productive rightside-up application of a rule appears to be necessary whenever a deletion or insertion rule, rather than a feature-changing rule, is involved, so this problem is not a peculiarity of this particular example.

Consider now the forms in (11):

- (11) petit ami [pətitami] /pətit#ami/
petit garçon [pətigarsõ] /pətit#garsõ/
petits amis [pətizami] /pətit+z#ami+z/
petits garçons [pətigarsõ] /pətit+z#garsõ+z/
petite amie [pətitami] /pətit+ə#ami+ə/
petite fille [pətitfiy] /pətit+ə#fiy+ə/
petites amies [pətitsami] /pətit+ə+z#ami+ə+z/
petites filles [pətitfiy] /pətit+ə+z#fiy+ə+z/

Standard rightside-up analyses (e.g., Schane 1968) posit lexical representations of the type given on the right, and the corresponding phonetic forms are derived by a rule which, roughly, deletes consonants preconsonantly and prepausally, together with one which deletes schwa in environments which need not concern us here. But this kind of lexical representation would presumably not be allowed in natural phonology, since consonant deletion is no longer a live process in French (cf. sac [sak], etc.), and so UDP could not adopt the standard analysis even if lexical representations were allowed to be as abstract as they are in natural phonology.

It is not entirely clear to me what the UDP analysis of these data would be, but it does seem clear that there will have to be at least two different ways of handling this apparently unitary phenomenon. If we accept L&R's proposal (p. 1) that 'lexical representations of words are expressed in their surface-phonetic isolation forms',¹⁶ then petit and petits would be lexicalized as /pəti/, while petite and petites would be /pətit/. Relating the masculine and feminine isolation forms (and the preconsonantal forms) is a relatively straightforward affair in UDP: the standard rule relating masculine and feminine forms of adjectives, given in (12), would be employed, together with the previously mentioned phonological rules, as in (13).

(12) Word-Formation Rule: ə-Attachment

[Amasc -ə] Afem

(13)	Word A	Word B	Morphology
LEXICAL FORMS	[pəti] _{Amasc}	[pətit] _{Afem}	Rule 12
a. <u>ə</u> -deletion	_____	[pətitə] _{Afem}	_____
b. C-deletion	[pətit] _{Amasc}	[pətitə] _{Afem}	_____

The plural isolation forms can be accounted for in much the same manner. However, nothing said so far accounts for the prevocalic variants, which also must have the isolation forms as their lexical representation. It is these cases which apparently require the productive upside-down use of the standard deletion rule, for we must somehow obtain phonetic [pətit] from lexical /pəti/, for example, in prevocalic environments. It

should be noted that, in order to determine which consonant to insert ([t] for singular petit, [n] for bon, [z] for mauvais and all plurals, etc.), the rule presumably requires transderivational or translexical power (cf. Lakoff 1973) in that the final consonant present in the feminine singular form will (usually) be the one to appear in the masculine prevocalic variants; otherwise, unless each masculine form is lexically marked for the consonant to be inserted, it is simply impossible to tell which consonant to insert. Condition (2c) is, as noted above (section 2.2), inherently translexical in nature (as is procedure (4)), but while this may be quite natural when two words are being compared to determine whether or not they are morphologically related, it seems much less clear that translexical power should be allowed in deriving surface forms from lexical forms.

Let us now briefly examine the possible revision of UDP mentioned in Fn. 16, that of listing the variants in question in the lexicon. Under this approach, petit would be entered as something like /pəti~pətit/ in the lexicon, or perhaps, following Hudson 1974, 1980, as /pəti{t}/, where the braces indicate that the t is not always realized in the phonetic form. Similarly, petits would be /pəti{z}/, and petites /pətit{s}/. Petite, at least on the basis of the data given above, shows no variation, and would be simply /pətit/, although in a full treatment it would contain a final {ə}, since a final schwa is present before so-called 'aspirate h' words. In addition, rules would be required to derive the phonetic forms from these; I will assume for the purposes of discussion that the consonant deletion rule mentioned above will do the job, although I am not sure how technically feasible this actually is. Whatever the actual form of the rules, however, this approach seems to entail a quite counterintuitive claim, namely that the lack of a t in the lexical representation of petits and the presence of a {t} in that of petit are two quite unrelated phenomena. There is no t in the former because there is none in any of its phonetic variants, and braces are required in the latter because one variant has a t and the other does not. Similarly, petite has a simple t in its lexical representation because it always shows up in the pronunciation. All of these phonetic facts have a straightforward explanation in standard rightside-up theories, namely that in all of these cases the morpheme /pətit/ is involved. Whenever the final t in this morpheme precedes another consonant or a pause, it is deleted; there is never a t in the pronunciation of petits because the underlying t is always followed by another consonant, the plural morpheme /z/, and it only sometimes shows up in petit because whether or not it is followed by a consonant depends on the initial segment of the next word. But the version of UDP under discussion does not relate the total lack of t in the phonetic variants of petits to the invariant presence of the consonantal plural morpheme, despite the fact that it does claim that in other cases the presence or absence of t is due to the nature of the following segment. As long as one feels, as I do, that the real reason for the lack of t in phonetic forms is the same in both cases (i.e., the presence of a following consonant), then the inability of this version of UDP to express this reason can only be considered a significant failing of the theory.

Thus, unless lexical representations are made even more abstract than those in natural phonology, UDP appears to be forced to treat cases such as French liaison as at least two separate phenomena.¹⁷ Such a degree of

abstractness would further weaken any claims of the theory about explanatory power with respect to sound change (cf. section 4 below).

3.1.4. The final example which indicates that UDP is too weak comes from Kasem. There is considerable disagreement about the optimal treatment of the Kasem facts, as presented in Callow 1965, 1968 (cf. Chomsky and Halle 1968, Howard 1969, 1970, Anderson 1969, 1974, Phelps 1975, 1979, Goyvaerts 1978, and Halle 1978); I will base my discussion on Anderson's analysis, since it is this analysis which is cited by L&R (p. 9) as being 'particularly troublesome' with respect to rule ordering, but most of the issues raised will be relevant to all of the published analyses of Kasem nominals that I know of. I will argue that the Kasem data are 'troublesome' with respect to more than just rule ordering.

Consider the singular and plural forms for the word for 'sheep', pia (sg.) and pe (pl.). Anderson analyzes these forms as being underlyingly /pia+a/ and /pia+i/, respectively, where -a is the singular marker and -i the plural marker for this class of nominals. The derivation of the phonetic forms involves three rules, one of metathesis which interchanges the first two of a sequence of three vowels, one of truncation which deletes one of a sequence of identical unrounded vowels, and one of contraction which converts, for example, /ai/ to /e/. The rightside-up derivations which Anderson posits are given in (14).

- | | | | |
|------|----|---------|-------------|
| (14) | a. | /pia+a/ | Input |
| | | pia | Truncation |
| | b. | /pia+i/ | Input |
| | | paii | Metathesis |
| | | pai | Truncation |
| | | pe | Contraction |

The ordering relationships in the above derivations are determined according to the principles of Anderson's 'local ordering', but are of no relevance to the point considered here (see Anderson 1974 for further discussion). Let us now attempt to relate the surface forms using Anderson's rules in an upside-down derivation, as in (15).

- | (15) | | Word A | Word B | Morphology |
|------|----------------|-----------------------|-----------------------|---|
| | LEXICAL FORMS: | [pi+a] _{sg} | [pe] _{pl} | [N-a] _{sg} ; [n-i] _{pl} |
| | a. Contraction | _____ | [pa+i] _{pl} | _____ |
| | *b. Truncation | [pia+a] _{sg} | [pai+i] _{pl} | _____ |
| | c. Metathesis | _____ | [pia+i] _{pl} | _____ |

It is somewhat unclear whether or not condition (2c) will permit the undoing of Contraction in (15a), since, while undoing it makes Word B look like the plural of something with respect to the plural WFR given in (15), it does not make it look like the plural of Word A. It is thus not at all certain that undoing this rule would 'increase the compatibility of forms A and B with respect to [this] Word-formation Rule', as is required by condition (2c).

This problem is not unique to (the upside-down version of) Anderson's analysis in fact, it remains in all of the analyses of Kasem cited above, and it is perhaps worthwhile to digress a moment to take a brief look at the analysis which undoubtedly has the least in common with the others, that of Phelps 1975, in this respect. The relevant rules are one of "Truncation", which deletes the second of a sequence of two vowels agreeing in backness and roundness (p. 313), and a rather unusual rule (cf. Halle 1978:181) called "Vowel Height Exchange", which among many other things, converts ea to ia (cf. Phelps 1975:314 and the revised version in Phelps 1979:37). These rules apply (in this order) to convert her underlying singular /pe+a/ to the required surface pia, via the latter rule, and the plural /pe+i/ to pe, by applying the former. The corresponding upside-down derivation is given in (16).

(16)	Word A	Word B	Morphology
LEXICAL FORMS	[pi+a] _{sg}	[pe] _{pl}	[N-a] _{sg} ; [N-i] _{pl}
a. V Height Exchange	[pe+a] _{sg}	_____	_____
b. Truncation	_____	[pe+i] _{pl}	_____

The undoing of (16a) is problematic in much the same way that the undoing of (15a) was. Here, the (putative) stems are made to look more like each other, in fact identical. But it is not easy to see that the words in question are thereby made more compatible with respect to the relevant WFR's, since Word B (still) does not look like the plural of anything. That is, as far as the WFR's are concerned, the two words are (still) totally incompatible. An obvious way of resolving this problem in a way which seems favorable to UDP is to define 'compatibility' as in (17):

- (17) The compatibility of Words A and B with respect to Word-Formation Rules R_1, R_2, \dots , is increased by undoing a given phonological rule if
- i. the putative stems, or an affix called for by more than one of the R 's, in the two words are thereby made to look more similar, in that there is an increase in feature specifications shared by corresponding segments, at least one of which was present prior to the undoing of the rule,¹⁸ or
 - ii. one of the words is made to look more like something called for by one of the WFR's, i.e., is made to look more like a member of the morphological category called for by this WFR, in the sense specified in i.

Such a definition would indeed permit the undoing of the rules under discussion, as well as that mentioned in note 11. It is not clear to me whether this is in fact what was intended by condition (2c), since 'compatibility' is nowhere explicitly defined by the proponents of UDP. If it is adopted, of course, it would, if anything, only exaggerate the difficulties faced by UDP in virtue of the fact that it is already overly strong.

But let us return to the question of derivation (15). Undoing Truncation in (15b) is required to obtain the eventual match, but disallowed even if condition (2c) is interpreted in accordance with (17): the words already look exactly like a singular and a plural, respectively, and so (17ii) is not met, and the stems are clearly not made more compatible by (17ii) (if anything, they are made less compatible, since there are now two pairs of incompatible segments, whereas before there was only one). What is more, it seems to me, undoing Truncation will not make the stems any more compatible under any obvious reading of this expression, and so is blocked by any reasonable interpretation of condition (2c). Note that no other linear order will help here, and, moreover, that even applying the rules in different orders for the words will not allow these words to be related.¹⁹ Thus, pia and pe cannot be related in this theory, despite their seemingly clear morphological relationship, given an Anderson-type analysis, or, it will be noted here (without justification, due to space limitations), in any of the analyses mentioned above, aside from those of Phelps, which, as hinted at earlier, are rather suspicious. It can thus be seen that the Kasem facts, unless they are subjected to substantial further reanalysis, present quite severe difficulties for UDP.²⁰

3.2. UDP is too strong. Since the possibility of using exception features is at least potentially a genuine drawback of standard theories, in that theories which permit their use may be able to describe impossible languages and thus are themselves overly strong, UDP's ability to do without such features makes it appear to be a quite attractive alternative. However, if it can be shown that UDP has only transferred the (putatively) objectionable power of exception features to some other device of the theory, then it can be seen that it is not nearly as attractive in this respect as it might at first appear to be. This, together with the fact that UDP in its currently proposed forms cannot generate some languages which it should, thus requiring some modification in the direction of being able to generate languages which it now cannot (cf. the discussion at the beginning of section 3), makes it doubly important to give careful consideration to this issue.

I will argue here that UDP has in fact merely transferred at least some of the power of exception features to some other aspect of this theory. It may even have more objectionable power than some standard theories, since such theories, even with exception features, appear to be unable to handle the following example, given a constraint on abstractness such as that of Kiparsky 1973. The reason this is so is that what would be exceptions to a rule in rightside-up theories correspond to the non-necessity of undoing a rule in UDP. Whether there are exceptions to a given rule in a rightside-up account, or how many exceptions it has, is thus totally irrelevant from the standpoint of UDP. That is, a rightside-up rule which applies in 99.99% of the cases in which its structural description is met is not distinguished in any way which I can see from one that applies only 0.01% of the time. It is this characteristic of the theory which allows it to generate impossible languages and thus makes it too strong.

Let us consider now an example which demonstrates this. Within UDP, pairs such as father/paternal, mother/maternity, etc., which most linguists would undoubtedly maintain are at best only distantly related synchronically,²¹

can be related 'without resorting to ad-hoc means' (L&R:4), just as caprice and capricious can. Consider the rules given in (18). Rule (18a) relates not only the [ð] of father and the [t] of paternal (via /θ/ by (18b)), but also singular/plural pairs such as elf/elves, path/paths, etc., and so presumably cannot be ruled out by the evaluation measure.²²

- (18) a. $\begin{bmatrix} +cont \\ +obstr \end{bmatrix} \rightarrow [+voi]/V_ [+voi]$
 b. $\begin{bmatrix} +obstr \\ -cont \end{bmatrix} \rightarrow \begin{bmatrix} +cont \\ -voi \end{bmatrix}$
 c. $n \rightarrow \emptyset / \begin{bmatrix} +cons \\ +son \end{bmatrix} _ \#$

(18b) is involved in at least ten alternations, six [t]/[ð] alternations (mother/maternal/maternity and the corresponding alternations with father and brother, again in concert with (18a)), and two [b]/[f] and [f]/[p] pairs of alternations in the father and brother sets,²³ so it too would apparently be sanctioned by the evaluation measure. It is perhaps worth mentioning here that the fact that evidence for (18b) comes only from the items in question which are in fact claimed not to be morphologically related, does not prevent the evaluation measure from sanctioning this rule in UDP, since it does not 'know' this. All that is available to the evaluation measure is the fact that such pairs are semantically related, and the ten 'alternations' in question can be 'accounted for' by this rule. (18c) will handle hymn/hymnal, column/columnar, damn/damnation, etc., as well as the six [n]/∅ 'alternations' involved here. We will also need the clearly independently motivated rules assigning stress and that of vowel reduction (cf., for example, Chomsky and Halle 1968). The derivation relating father and paternal using these five rules, all well-motivated from the standpoint of UDP, is given in (19).

(19)	Word A	Word B	Morphology
LEXICAL FORMS	[fáðər] _N	[pətʌrnəl] _A	[N-əl] _A
a. Rule (18a)	[fáθər] _N	_____	_____
b. Rule (18b)	[pátər] _N	_____	_____
c. Rule (18c)	[pátərn] _N	_____	_____
d. Stress	[patərn] _N	[pətʌrnəl] _A	_____
e. V Reduct.	[patʌrn] _N	[patʌrnəl] _A	_____

Note that in this derivation none of the stages violates condition (2c), so that father and paternal can be readily related, requiring only two more steps than the sane/sanity derivation, and the same number of steps as that relating caprice and capricious. Evidently, in this theory the father/paternal-type pairs are at least no more distantly related than caprice and capricious. This is in spite of the clearly extravagant nature of (at least) rule (18b) and the lack of such extravagance in L&R's rules.

One further point deserves mention in connection with this discussion. It may have been noticed that rule (18b) is context-free, i.e., it converts all stops to voiceless fricatives, regardless of their phonological context. Clearly, this rule has an immense number of exceptions to it, or at least it would have in a rightside-up theory. But, as noted above, this is simply irrelevant from the standpoint of UDP. Leben himself has noticed (Leben 1979:199) that there is nothing in the version of UDP given in L&R to prevent rules from being 'absurdly general', but suggests that the 'recoverability principle' he proposes there 'limits the generality of the rules'. I cannot see why this would be so; in the other cases discussed by Leben (pp. 191-8) for which there is a recoverability problem, this problem is solved not by eliminating or altering some rule which is causing the problem, but by changes in the forms themselves which permit their relationship to become more recoverable with respect to the existing rules. I thus fail to see why Leben feels that this principle has the effect he claims it does with respect to 'absurdly general' rules. What the principle actually seems to predict in the case at hand is that rule (18b) is a part of the grammar of English (for the reasons given above), and that, due to the recoverability problems caused by this rule, forms will change so as to increase their recoverability with respect to it. (The forms under discussion would not be the ones predicted to change, it should be noted, since here there is no recoverability problem in Leben's sense--see section 4 for discussion). Thus, the fact that 'exceptions' to rightside-up rules 'don't count' in UDP represents a serious flaw in the theory. That is, even if the possibility in UDP of having 'incomplete' rules (McCawley 1979) is for other reasons an attractive one, there are also problems caused by this possibility, and these are not gotten around by Leben's 'recoverability principle'.

It should be noted that the procedure suggested in Robinson 1977 (cf. 4 above) will also allow the successful relating of these words, since in this case the 'evidence...for the insertion of a specific feature value' (Robinson 1977:7) does indeed come 'from the items being compared', and not from what is called for by a WFR, as was the case in the Icelandic example discussed above (section 3.1.1). Thus, not only does this procedure make it impossible to relate, for example, the Icelandic forms, but it also allows father and paternal to be just as easily related as in the original version of the theory. It should be noted that this relating involves the use of a rule of absolute neutralization, (18b), and that it is thus not the case that Robinson's revised procedure would (p. 9) 'rule out abstract analyses'; while it may well rule out more or less plausible such analyses (although I am not certain that it would), it fails to disallow implausible ones--analyses which no rightside-up account would have seriously considered. The reason is, again, the 'exceptions' simply do not matter in UDP.

Finally, note that the possibility of having rules like (18b) also would appear to make a characterization of rule naturalness or markedness very difficult for UDP, since the rules would lack the phonological contexts which often influence naturalness in rightside-up theories.

4. Some further diachronic considerations.

As noted above, UDP is claimed to have 'diachronic advantages' over rightside-up theories (L&R:11). I would like to turn not to a brief examination of this claim. It is argued (L&R:14) that UDP offers an explanation for the fact that 'sound change affects only surface items' because of 'the purely derivative nature of non-phonetic representations' in the theory, whereas 'in standard theory it must be stated as an extrinsic

constraint on change (cf. King 1973)'. As mentioned above (section 3.1.3), however, this advantage disappears from versions of the theory, such as those proposed by Pollack and Leben, in which lexical representations are not 'surface items'. Notice that lexical representations in the Pollack/Leben model can be quite abstract (i.e., relatively different from the phonetic representations). In the Russian example, of course, they would be more abstract than the 'classical phonemic' representations, as Halle 1959 has pointed out, and they undoubtedly would be in the English case as well. But if UDP claims to explain the putative fact that sound change affects only phonetic representations (i.e., L&R's 'surface items') by having such representations in the lexicon (and, presumably, claiming that phonological change affects only lexical representations), it cannot allow for such a revision. Once any degree of abstractness is permitted in lexical representations, moreover, UDP is in precisely the same position as any rightside-up theory as far as permitted phonological changes are concerned; the putative diachronic explanatoriness can be maintained only by maintaining fully concrete lexical representations. Thus, even if a position such as the one suggested (p. 2) by L&R that some less abstract, but still not fully phonetic, level is the level of lexical representation should be adopted, the diachronic consequences of the theory will not be as claimed by L&R. Precisely analogous problems would arise concerning their claim (p. 19) that their theory makes it possible 'to make room for lexical gradualness' in diachronic change if the Pollack/Leben proposal (or any proposal with 'abstract' lexical representations) is adopted. Thus either this model must be given up (thus leading to the synchronic problems mentioned in section 3.1.3) or these claims of the theory about diachronic explanatoriness cannot be maintained.

Let us now turn to a brief examination of the diachronic implications of Leben's 'recoverability principle' (cf. section 3.2). This principle, together with UDP's blindness toward exceptions to rightside-up rules, and the evaluation measure, makes some startling (and obviously incorrect) claims about change. For the evaluation measure, unless perhaps it is revised to reflect 'recoverability' somehow, would dictate that every rule in UDP should be context-free, because of the features 'saved' by getting rid of environmental restrictions. This, of course, would result in wholesale recoverability problems, since 'the established chain of rules' (Leben 1979:198) would be broken frequently with many cases of these 'absurdly general' rules not requiring application. The corresponding diachronic prediction is wholesale changes in pronunciation to alleviate these problems.

It thus would appear that a closer look at Leben's 'recoverability principle' is in order. Unfortunately, I am not fully confident that I understand what Leben intends in this respect, since he nowhere explicitly defines this term, and the relevant discussion is not a model of clarity. It is therefore open to question whether the definition given in (20), which is pieced together from various places in Leben 1979, accurately reflects his intentions.

(20) Recoverability Principle: If

- (i) 'from looking at a given surface form' one cannot 'immediately tell' (p. 194) what the surface form of the corresponding 'base word' should be (i.e.,

a segment in a 'derived word' potentially corresponds to more than one segment in 'a base word' (p. 197),

and

- (ii) there is no 'orderly chain of segments that are successively tried in searching for an appropriate... word' (p. 196) by virtue of 'a chain of successfully applied steps' (p. 197) in a UDP derivation,

then the difficulty of recovering the morphological relationship between the two forms is significantly great, and the pronunciation of some forms is predicted to change in order to alleviate this difficulty.

Particularly puzzling to me are statements which seem to indicate that only one of (20i) or (20ii) need be satisfied in order to create a recoverability problem. Thus, for example, Leben (p. 197) outlines 'the principle suggested here, that recovery is hindered when there is more than one path that can be followed in taking a segment in a derived word back to a representation compatible with the corresponding segment in a base word...', while he notes on the next page that the nonnecessity for undoing Vowel Shift in relating caprice and capricious 'breaks the established chain of rules, pointing to a recovery problem under the principle proposed above.' This suggests that satisfying (20ii) will suffice to cause a recoverability problem. He even seems to imply (e.g., p. 197) that he feels that the two conjuncts of (20) are synonymous. Nevertheless, since in his discussion (pp. 195-6) of an English example, he indicates that he feels that (20i) is satisfied (but (20ii) is not), and that yet there is an 'absence of any sign that...pressure [for change in this case] is being felt', it seems most likely that Leben intended a conjunction, and not a disjunction, of (20i) and (20ii).

Assuming that this is in fact what he intended, two questions immediately present themselves. First of all, why is it that recovery is 'hindered' when a 'derived word' potentially corresponds to more than one 'base word' but not vice versa? It seems clear that the 'vice versa' situation is not all rare--one cannot, for example, tell immediately from looking at the German surface form [bunt] whether the corresponding derived word should be [buntə] or [bundə]--and I can see no reason why such a situation should be any less problematic for recoverability than that described in (20i). It could be that not allowing for recoverability problems to result from this kind of thing represents a simple oversight on Leben's part, and he might well have intended that the 'recoverability principle' be extended to cover situations like this. But if he did, then it is difficult to see that the 'recovery problem' was 'rectified' (p. 194) by the Polish changes discussed by Leben, and so let us take a brief look at these changes. In Old Polish, there were pairs such as radość 'joy' / rado[śn]y 'joyful', zawiść 'grudge' / zawi[stn]y 'begrudging', post 'fast' / po[śn]y 'lenten', and głos 'voice' / gło[śn]y 'loud' (cf. pp. 189-90, 193). Leben attributes the changes from rado[śn]y to rado[sn]y and from po[śn]y to po[stn]y to the fact that they (p. 197) 'aided recovery of the segments in question'. This is probably true, since there is now only one 'base form' which corresponds to 'derived forms' in [śn], while in Old Polish there were three.²⁴ But notice that this change had no

effect on what would also have been a recoverability problem under this interpretation--there are still two kinds of derived forms which correspond to base forms in śc, just as there were in Old Polish (although one of these derived forms is different than it was in Old Polish). It should also be noted that this change, while it has eliminated the recoverability problem with respect to derived forms in [śn], has also introduced one with respect to derived forms in [stn], since there are now two corresponding base forms instead of the single base form in Old Polish. Note further that (20ii) is still not satisfied after the change, since C Drop (cf. note 24) must be blocked by condition (2c), and the 'chain of successfully applied steps' will thus be broken, in the relating of głos and głos[śn]y (assuming, in accord with condition (2a), that the C Drop rule precedes the one responsible for the s/ś alternation in a UDP derivation). That is, the changes in question had a rather minimal effect on recoverability, and, what is more, entail a rather substantial alteration in the grammar of Polish.

Leben never gives a clear picture of what the grammar is like at either stage, but judging from Gussmann's account, there would have to be a rule reordering and the loss of another rule,²⁵ both of which appear to carry over to the UDP account. In Gussmann's account, some lexical items must acquire a diacritic marking them as exceptions to C Drop (p. 302) as well, but this will of course not carry over to the UDP account. This would appear to be a rather extreme reaction to the recoverability problem, given the meager amount of resulting improvement in terms of recoverability (see above). This is especially true from the standpoint of a theory like UDP, where speakers memorize pronunciations and are only rarely called upon to recognize morphological relationships. In fact, from the standpoint of such a theory, it is not all clear why speakers should be particularly bothered by recoverability problems, certainly not enough to make the drastic changes in the grammar required in the Polish case. Moreover, one can very well question the relevance of a state of affairs like that described in (20i) in an account which (Leben, p. 198) 'makes crucial use of the characteristic...of sometimes factoring complex alternations into a sequence of simple ones...'; if this possibility is so important, then how can the ability to 'immediately tell' (i.e., without a sequence of steps) what kind of 'complex alternation' there is also be of critical importance? Why go through the sequence of rules if you already know the answer to your question?

The second question mentioned above concerns the fact that all that is predicted by the principle is that some forms will change. Nothing at all is said about which forms will change, or about the direction in which change will proceed. Without some refinement in the predictive power of the theory, the Polish changes, which certainly seem rather unlikely from the standpoint of standard theories (at least, changes like that from rado[śn]y to rado[sn]y), would appear to be regarded as rather unlikely from the standpoint of Leben's revised version of UDP as well.²⁶ There was, after all, another way of alleviating the recoverability problem, namely doing away with allomorphy completely by changing rado[śn]y to rado[ścn]y, and indeed, something precisely analogous happened in the case of po[śn]y. Moreover, this is not the only change which would seem to be viewed as more likely than the actual change from the standpoint of the theory. A change to rado[stn]y, for example, would appear to give

all the recoverability advantages of the actual change, and also decrease allomorphy in some sense, since there would be a consonant cluster in the derived form corresponding to the stem-final cluster in the base form. (In addition, this would alleviate the problem of nonpredictability of derived forms from a given base form, if the recoverability principle is taken to cover such cases--see the discussion above.) That is, without some such refinement, the recoverability principle cannot be taken to offer a very good explanation of the changes in question.

The diachronic objections raised by Janda and discussed and elaborated on above (section 2.1) are also relevant here, but I see no need to discuss them further.²⁷

5. Psycholinguistic issues.

L&R:3-4 suggest that psycholinguistic evidence might provide further support for their theory. It seems to me, however, there is a fair amount of evidence of this type which would create problems for the theory.

First of all, it is well known that children tend to 'over-generalize', apparently using rules when they should not from the standpoint of the adult system, even when they know the correct adult form (cf., for example, McNeill 1970). This would appear to be very good evidence that they are in fact using word-formation rules and phonological rules productively. Thus, unless we are willing to accept that adult grammars are radically different from child grammars, in that in the case of the latter most rules apply rightside-up and productively, while for the former they usually apply upside-down (if at all) and surface pronunciations are for the most part memorized (recall that few rules can meet L&R's criterion of being 'fully regular'--cf. Fn. 13), it would seem that adults too should use productively and in rightside-up fashion processes which are not fully regular in L&R's sense. There is some more direct evidence which points to the same sort of conclusion as well, such as the classic Berko study (cf., again, Fn. 13). It is worth pointing out that some of the rightside-up uses of rules suggested by Berko's study require at best marginally productive rules in anybody's theory, such as those which relate sing, sang and sung; this fact would appear to be quite troublesome for the theory no matter how 'fully regular' is interpreted.

A further kind of direct evidence (of at least a semi-psycholinguistic nature) for the productive rightside-up use of WFR's which are not fully regular is that from the study of speech errors (cf., for example, Fromkin 1971). Slips such as groupment for grouping and concludement for conclusion (p. 45) can only be explained, as far as I can see, as the productive use of the WFR which forms nouns from verbs by adding -ment.²⁸ Yet, this rule is not 'fully regular' in L&R's sense, due to its semantic idiosyncrasies (cf. government), as well as, at least for some speaker's, phonological ones (no [n] in government). This WFR thus could apparently not be used by L&R's criterion.

One final bit of semi-psycholinguistic evidence should be mentioned here. It concerns the apparent ease (noted in Hetzron 1975:870) with which speakers can tell whether or not they have previously heard a word which is the result of a process of derivational morphology (or one of radically irregular inflectional morphology, I would add), as contrasted with one which results from a regular inflectional process, in which case speakers are much less clear about whether they have heard it before.

This seems to suggest a different kind of lexical storage in the two kinds of cases. While in the case of the first type, it might well be possible to maintain that words are stored as wholes (i.e., as L&R:1 put it, 'the surface representations of words'), this hardly seems a legitimate possibility in the latter type of case. To maintain otherwise would appear to be equivalent to saying that speakers have memorized the (surface representations of) words in both cases, but that speakers have conscious knowledge of these memorized forms only in cases of the former type, and in cases of monomorphemic words, but not in the other cases. Even if one were willing to go along with this (in my view) absurd position, it seems to be clearly incumbent upon an adherent of such a position to provide an explanation for the difference in question. None has ever been offered, to the best of my knowledge. The implications of this as far as UDP is concerned are that, at least in cases of the second type, it apparently cannot be held that (L&R:1) 'the lexicon...excludes all but the surface representations of words', or even some more abstract representations of words; it must contain at least some bound morphemes.

Let us now briefly consider a different psycholinguistic issue, that of language acquisition, from the standpoint of UDP. In order to acquire competence in the phonological and morphological aspects of the language being acquired, a child must at least learn the following: (1) lexical representations of words; (2) morphological rules; and (3) phonological rules. In addition, the child must also learn the order in which the phonological rules apply if extrinsic ordering is permitted, and, if the Pollack/Leben model is taken somewhat loosely (so that the Stampean 'natural processes' are not considered to be innate), natural processes. For the purposes of this discussion, I will assume that the acquisition of all of the above is unproblematic except in the case of the phonological rules (but see the preceding few paragraphs for some discussion of morphological rules). These rules, judging from the examples provided by Leben and Robinson in their work, are given in rightside-up form, and therefore must presumably be learned in such a form by the child if UDP is to have anything at all to say about language acquisition. Yet, at least in the case of rules which are not 'fully regular,' they are never used in the way which the form of the rule suggests (i.e., rightside-up). Such rules are always undone--never does a rule of this type apply to the 'input' to produce the 'output'. This leads one to question how a child could acquire a rule which will never be used in the form in which it is acquired. And, although it might not be easy to imagine how a child could acquire the rules required in standard theories (they are, after all rather complex and abstract), it seems even more difficult to understand how rules could be acquired if these rules are never used as such, as appears to be the case if UDP is adopted.³⁰

At this point, it might be suggested that UDP should not be committed to the rightside-up rules of standard theories, but rather that the phonological rules in a given analysis need have nothing to do with the corresponding rules in standard theories. That is, UDP should be positing rules which are 'upside-down' only from the standpoint of standard theories, but are actually in the form which phonological rules 'should' take (i.e., are rightside-up from their own perspective); the rules given in current treatments are only for the purposes of exposition (cf. Leben 1979), and the rules in a fully worked-out UDP might take a rather different form.

If this suggestion is adopted, however, it seems that at least two new problems will be created. First, there is the problem that some rules must apparently be allowed to apply productively in rightside-up fashion, and rewriting the rules in a form which makes them look more acquirable does not appear to allow for this possibility. Moreover, such a revision of the form which rules take seems to make it quite difficult to impose any well-motivated set of constraints on the form which phonological rules may take, at least not without making reference to the form which they may take in standard theories (there would be no necessity, for example, for rules to be phonetically plausible, and in fact they would typically be just the opposite, perhaps even context-free). The lack of such constraints would subject UDP to the same criticisms which Leben 1979 raises against Natural Generative Phonology; there would be no conceivable rule, no matter how extraordinary, which is not viewed as a potential phonological rule of some natural language. No matter which alternative is chosen, then, UDP appears to be in an uncomfortable position as far as language acquisition is concerned.

6. Conclusion.

In this paper, I have argued that the theory of UDP as represented in various versions proposed in the literature, as well as in conceivable alternative versions, creates at least as many difficulties as it alleviates with respect to 'standard', rightside-up theories of phonology and morphology. In particular, I have argued (cf. section 3 above) that UDP is faced with extreme problems from a synchronic point of view, in that it incorrectly characterizes the class of possible natural languages; what is more, the characterization provided is incorrect in both of the possible respects--the theory fails to provide adequate descriptions for possible (and actual) languages, and allows for the straightforward description of impossible ones--which suggests strongly that no minor modification of the theory can fully remedy this situation. I have also argued (section 4) that, at least in versions of the theory which can alleviate some of these synchronic problems, UDP is in much the same situation as rightside-up theories with respect to phonological change, i.e., that UDP has no 'diachronic advantages' over other theories, and that a version incorporating Leben's 'recoverability principle' in fact has diachronic disadvantages. Finally (section 5), I have argued, the characterization of the grammar given by UDP is implausible on psycholinguistic grounds. Thus, even though the previous criticisms of UDP considered in section 2 were found not to be terribly forceful on the whole, there are a number of genuine problems for the theory, at least some of which would appear to be insurmountable by any revision of the theory which retains the 'crucial' condition (2c). Since it is not at all clear that any other currently proposed theory of phonology/morphology is fully adequate, we will probably have to continue our search for an appropriate such theory. The results of this paper can be of some aid in directing this search, in that we know where not to look; it seems clear that such a theory will not be, in the sense intended by Leben and Robinson, 'upside-down'.

Footnotes

* This paper is an extensively revised and expanded version of a paper presented at the 1977 Summer meeting of the Linguistic Society of America. I would like to thank P. de Zeeuw, P. Gathercole, V. Gathercole, R. Janda, B. Joseph, W. Leben, J. McCawley, and A. Zwicky for their helpful comments on that version and on another preliminary version of this paper, as well as discussion of some of the general issues involved.

¹Janda's definition undoubtedly needs some refinement. As it stands, it implies, quite counterintuitively, that undoing a deletion rule (any deletion rule) on both forms being considered always increases their compatibility. Thus, for example, if there is a WFR which adds a suffix X to a stem to form some other morphological category, and a phonological deletion rule, then the compatibility of the forms Y and Y + X would be increased by undoing the deletion rule according to this definition, even though they already matched with respect to the WFR. Such considerations will have some relevance later on (cf. section 3.1.4); I will postpone further discussion until then.

²Leben (personal communication) has informed me that this interpretation is in fact what was intended. It is not clear, however, that such an interpretation would not allow for derivations which Robinson 1977 wants to rule out as too abstract, e.g., in cases where an absolute neutralization rule affects (in a rightside-up treatment) an affix and the underlying segments subject to this rule have an effect on segments in the stem. For further discussion of this question, see section 3.1.1.

³In Janda's discussion, he does not require that the words in question be semantically related, although he considers this possibility in a footnote (pp. 58-9), and he never considers the possibility of checking only those morphological rules suggested by the meanings of the words involved. But speakers clearly know, for example, that left is a past tense form and therefore presumably the result of a past tense formation rule, and so there is no reason not to build corresponding kinds of information into the parsing procedure. The situation in this respect is thus not nearly as bad as Janda's discussion implies, although the amount of work which must go into showing that depart and left are not related, even with the semantic restrictions suggested here (see below), seems clearly disturbing from the standpoint of UDP.

⁴Janda indicates (p. 43) that he feels that such a change is one which UDP 'should perform,' but apparently has not recognized that the alteration proposed above would in fact perform such a change, although he also states somewhat puzzlingly that all rules of all types' 'should be undone...by reversing the arrows in their respective rightside-up versions.' As a result, he comes up with a quite unwieldy revision of condition (2b). His revision, together with an apparent misinterpretation of L&R's requirements for 'matching', leads Janda to further irrelevant criticism (p. 45) of this condition; it is not 'the purpose of an upside-down derivation...to arrive at individual segments...that are identical...', but rather (L&R:3) to arrive at forms which are 'non-distinct' (presumably in the sense of Chomsky and Halle 1968:336). (It should be noted in this regard, however,

that Robinson 1977, 1980 does in fact require that segments eventually be identical (and fully specified) in order to 'match'.) Since Leben in his singly authored papers never indicates that he subscribes to a similar interpretation, it could very well be that there is some disagreement between Leben and Robinson on this count. Even so, it is not at all clear that Robinson's requirement for matching is subject to Janda's criticisms on this matter, since his 1977 paper gives a significantly different procedure for undoing rules (cf. procedure 4 above).

⁵This statement follows from the mathematics of infinite sets. For general mathematical discussion of this point, see, for example, Wilder 1965; for an application of this result to a comparison of theories of rule ordering, cf. Churma 1980c. Wilbur 1975 also contains relevant discussion of the non-inherent nature of the evils of globality.

⁶The situation is not quite as simple as this discussion might make it appear. It is always possible to contest a given analysis and, thus, the claim that this analysis corresponds to a possible natural language; and it is likewise possible to maintain that a claimed impossible language is in fact a possible one. Such a state of affairs appears to be not at all restricted to issues of this nature, or even to linguistics; I have argued elsewhere (cf. Churma 1979, 1980a) that the ultimate determinant of a given scientist's acceptance of most theoretical claims is his or her (degree of) belief in the truth of the premises of the argument which leads to the conclusion in question.

⁷Such a question is undoubtedly behind the difficulty which Janda 1980:29 has in seeing how the proposal of Robinson 1977 'could be considered to be "undoing" a rule', and behind his dissatisfaction with it. Janda never comes to grips with this theory (whatever it is) on its own terms, however, and whether or not the modified theory is still UDP has no bearing on the status of its claims about the class of possible morphophonologies. I will argue below that this theory, too, is unsatisfactory in this respect.

⁸The issue which Janda raises in several places (e.g., p. 53) of whether a version of UDP which incorporates exception features is 'a notational variant of' standard rightside-up theories is thus not a terribly interesting one. Given his beliefs about the crucial nature of condition (2c), it is somewhat puzzling why he finds it so.

⁹This is essentially the proposal which Koutsoudas, Sanders, and Noll 1974 (hereafter KSN) have made concerning the application of rules in rightside-up phonologies; i.e., rules apply whenever their structural descriptions are met. It would also give much the same results as the theory of rule application presented in Robinson 1980, although the latter could not handle cases like the one under discussion. It is worth pointing out in this respect that Robinson's theory, or the simultaneous theory just outlined, would also require something analogous to KSN's principle of 'proper inclusion precedence' to handle cases of mutual bleeding and what I refer to (Churma 1977) as transfusing-type relations (i.e., cases

in which the order of the rules has empirical consequence, but where the relation between the two rules is of neither the feeding nor bleeding type).

¹⁰It appears that in the Nape example discussed here, the two forms cannot in fact be related via rule (7) by this modification of Robinson's procedure, since there is 'evidence for' [+round] in rule (8), but none for the required [+back]. But this has nothing to do with the fact that (7) is a rule of absolute neutralization, and it seems that such a modified procedure could handle the German example of Bach and Harms 1970 discussed by Robinson.

¹¹It is not clear that they can be undone, since (9a), while it does put back a b into Word B, does not appear to put it into a position which 'corresponds' to that of the b in Word A. Furthermore, undoing this rule destroys the match which previously existed with respect to the reduplication part of the WFR (i.e., it looked like there was reduplication before the undoing of (9a), but not after). It is therefore difficult to tell whether or not condition (2c) would be violated here. Janda is thus quite correct (p. 4) that the term 'compatibility' requires further clarification, although not for the reasons he suggests (cf. section 1 above; for further discussion of the question at issue, see section 3.1.4 below).

¹²Despite this, Leben (personal communication) has suggested that all reduplication rules should in fact be undone (presumably because they look a lot more like phonological rules than do other WFR's). In fact, he allows a Hausa reduplication rule to be undone (cf. Leben 1977a:431, 432, 438). But the Hausa rule is not a WFR, as the formulation given by Leben (p. 429), and repeated here as (i), indicates.

$$(i) \quad [X(C)C]_R - VC - V \rightarrow [X(C)C]_R - VC - C - VC - V$$

$$\quad \quad \quad 1 \ 2 \ 3 \quad \quad 4 \quad 5 \quad 1 \ 2 \ 3 \quad \quad 4 \quad 3 \quad 4 \quad 5$$

There is no mention of any morphological function in this rule, and, as far as I can tell, its only function is to make things difficult for Hausa children. The two cases are thus quite different. And, of course, it is not the reduplicative nature of the rule in question that is the issue--any WFR whose relationship to the phonological rules was of this type would entail precisely similar problems (cf. Anderson 1975:48-50, 56 for two such examples of non-reduplication rules from Danish and Rotuman).

¹³It is not clear to me exactly what is intended by the term 'fully regular'. If taken literally (i.e., meaning 'exceptionless'), it would seem to entail that English plural formation, for example, could not be used 'in the formation of new words from existing words' because of pairs such as child/children, mouse/mice, etc. This seems clearly not to be the case (cf. Berko 1958). Note further that it is apparently not possible in this theory to make regularity a matter of degree as far as directionality is concerned, since productive (phonological) rules will apply rightside-up and nonproductive rules will apply upside-down, thus not allowing for any middle ground.

¹⁴There is a good deal of evidence that almost all linguists who have written on the subject feel that Russian voicing assimilation is in fact a single phenomenon, i.e., that the minor premise of the argument given by Halle 1959 is true (for justification, cf. Churma 1980b). Considerable argument would thus appear to be in order if it is felt that there is more than one phenomenon involved here.

¹⁵Leben 1979 has apparently independently come to an essentially identical conclusion about the nature of lexical representations, although not for these reasons.

¹⁶If we do accept this, then it would appear that the theory has no way of even talking about sound changes which affect the alternating parts of non-isolation external sandhi variants. They would thus presumably be forced into listing every occurring 'surface-phonetic' form of every word in the lexicon in order to retain their putative explanation of the nature of sound change. Further discussion of the possibility of listing external sandhi variants in the lexicon is given below.

¹⁷This criticism would apparently apply to all theories which require a level of lexical representation which is less abstract than that required here, such as that of natural phonology or of Natural Generative Phonology (cf. Donegan and Stampe 1979 and Hooper 1976, and the references cited there).

¹⁸The qualification that one of the segments which undergoes the increase in shared feature specifications must be present prior to undoing the rule is intended to remedy the problem mentioned in Fn. 1.

¹⁹Note also that the principles for rule application proposed in Robinson 1980 do not determine an order for contraction and truncation, since this is a transfusing order (cf. Fn. 9), and so does not fit into the feeding-bleeding taxonomy.

²⁰Robinson's revised procedure for undoing rules (cf. 4 above) would of course be even worse off, since it would straightforwardly disallow the undoing of (15a) and (16a).

²¹Not everybody would; Lightner, for example, would probably find them just as closely related synchronically as he does nation and pregnant (cf. Lightner 1975:617). Whether or not two forms are in fact synchronically related does not seem to be the kind of question that one can give rational arguments about, and so I will not attempt to do so. A Lightnerian would thus probably find this discussion irrelevant, or perhaps even supportive of the theory, just as I would find a discussion of the 'problem' of how to relate nation and pregnant in UDP or any other theory rather strange. I direct this discussion only toward those who agree that such forms as father and paternal should not be straightforwardly relatable phonologically in a synchronic grammar of English.

²²Note that any semantically related pair, such as depart/left, could be related (by the past tense-formation rule) if we did not appeal to the evaluation measure to rule out the 'phonological' rules needed to relate the forms in question. (No exception features would be needed if we did allow these rules, however.) Note also that there would appear to be nothing any more exceptional about the singular/plural pairs with this voicing alternation than the [s]/[z] alternation in the plural morpheme in this theory. That is, if the ordinary allomorph of the English plural morpheme is taken to be underlyingly /z/, so that the relevant word-formation rule is something like [N-z]_{Npl}, then the theory appears to make no distinction in complexity between alternations such as cat/cats and bus/buses, on the one hand, and leaf/leaves on the other. The latter would require rule (18a) to be undone, while the former would necessitate the undoing of a rule devoicing obstruents after voiceless segments and one inserting schwa between sibilants, respectively. The fact that there is no such alternation in the great majority of forms where there is the potential for one seems to be simply irrelevant as far as UDP is concerned, despite the rather extreme counterintuitiveness of these implications of the theory.

²³Note that this rule relates the p of paternal to the f of father, but the b of brother to the f of fraternity, so that the input to the rightside-up rule is part of the adjective form in the first case, but part of the noun form in the second. This, again, is quite counterintuitive, but irrelevant from the point of view of the theory, as far as I can tell.

²⁴There is a complication not immediately apparent which should be pointed out here. Old Polish post and po[śn]y cannot be directly related using the Leben/Gussmann rules (Leben's discussion is based on that of Gussmann 1976). The reason is that, while the rule of 'C Drop', which deletes a coronal consonant when it is between (among other things) s and n (Leben, p. 188), can be undone on po[śn]y to give something like po[śCn]y, the rule of palatal assimilation (which converts, e.g., s to ś before a palatal) cannot be undone since its structural description is not satisfied by a following segment which is unspecified for palatality (the 'C' would be specified only [+coronal]). Robinson's revised procedure (4 above) would also fail here, since comparison of the variable feature specifications inserted in undoing 'C Drop' with those in post would yield po[śtn]y, and structural description of palatal assimilation would still not be met. It thus may well be that UDP cannot relate these forms in Old Polish and that the theory fails solely on this count, with the recoverability principle, or the lack thereof, being quite beside the point. My knowledge of Old Polish is not sufficient to allow me to state that the forms in question cannot be related indirectly via a third form, although if the behavior of po[śn]y parallels that of rado[śn]y, it appears that they cannot, since in the latter case (p. 194) 'the noun stem within the adjective invariably appeared as radoś-' in Old Polish, and we would thus be faced with exactly the same situation with any member of the case paradigm.

²⁵ I refer here to what Gussmann calls (p. 301) the 'depalatalization rule which must depalatalize spirants before nonpalatal consonant clusters.' Gussmann may have intended that this function be performed by a generalization of a rule he calls (p. 291) 'anterior depalatalization', since he does not show a separate step corresponding to the former rule in his Old Polish derivations. If this is so, then the change in question requires a rule 'degeneralization' instead of a rule loss, as well as synchronic iterative application of (anterior) depalatalization. It is worth pointing out in this respect that if Old Polish had a rule of 'palatal assimilation' (Gussmann, p. 292) which not only cause 'dental spirants [to] assimilate to the place of articulation of the following palatal', but also performed the function of the depalatalization rule (i.e., if it made coronal spirants agree in palatality of a following consonant), then this additional rule (or generalization of one) is not necessary, as long as palatal assimilation can be applied to the output of anterior depalatalization. However, since the former is crucially ordered before the latter in Gussmann's account (although it is reordered so as to follow it in modern Polish), this would entail an ordering paradox. Whether or not Gussmann's analysis of Old Polish is the optimal one is thus certainly an open question, and it is unfortunate that Leben bases so much of his discussion of recoverability on it.

²⁶ This is not unique to UDP; it would also be the case for any rightside-up theory, such as that of Gussmann, which invokes a similarly vague recoverability principle.

²⁷ It should also be noted that, as I have argued elsewhere (Churma MS), a version of UDP which does not allow for 'rule inversion' (Vennemann 1972) as a mechanism of change--and Leben at least apparently does not want to (cf. Leben 1974, 1979)--cannot provide a reasonable account of certain changes in the Chadic language Kanakuru. This problem is not unique to UDP, however; it seems that the Kanakuru changes would be quite problematic for any theory which rejects rule inversion.

²⁸ I cannot see why Fromkin feels that it is 'possible' (let alone 'highly probable') that the latter slip involves 'a blend of concluding and conclusion': where did the -ment come from? If she only means by this that this slip does not provide terribly strong evidence for the productive application of a phonological rule converting d to z in rendering conclusion, however, I would have to agree.

²⁹ Such considerations would appear to indicate problems for any rightside-up theory of which I am aware, as well. What appears to be needed, at least as far as the nature of lexical representations is concerned, is some sort of combination of UDP (or some similar theory) and standard theories.

It should also be noted in this respect that the suggested 'line' between morphological processes need not be precisely as indicated here for the general point to hold; as long as there is some such division, there will be a corresponding problem for UDP (and standard theories, as long as no morphologically complex lexical entries are permitted by such theories).

³⁰Leben (personal communication) has questioned the relevance of these considerations on the grounds that 'it is hard to see what difference rule directionality would make to an account' of acquisition when what is acquired from the standpoint of UDP is taken to be 'the ability to abstract away from permissible allomorphy in perceiving morphological relationships'. I suppose that the considerations in question would indeed be irrelevant from such a point of view, but then it is hard to see what facts about acquisition could possibly be relevant to such a vague characterization of this process. That is, given only this characterization of language acquisition, UDP simply fails to attempt to solve 'the problem to which the linguist addresses himself....to account for the child's construction of a grammar...' (Chomsky and Halle 1968:331).

References

- Anderson, Stephen R. (1969). West Scandinavian Vowel Systems and the Ordering of Phonological Rules. Bloomington, IN: Indiana University Linguistics Club.
- Anderson, Stephen R. (1974). The Organization of Phonology. New York: Academic Press.
- Anderson, Stephen R. (1975). On the interaction of phonological rules of various types. *Journal of Linguistics* 11:39-63.
- Aronoff, Mark. (1976). Word Formation in Generative Grammar. Cambridge, MA: MIT Press.
- Bach, Emmon and Robert D. King. (1970). Umlaut in modern German. *Glossa* 4:3-21.
- Baker, C. L. and Michael K. Brame. (1972). 'Global rules': a rejoinder. *Language* 48:51-75.
- Berko [Gleason], Jean. (1958). The child's learning of English morphology. *Word* 14:150-177. [Reprinted in *Child Language: A Book of Readings*, ed. by A. Bar-Adon and W. F. Leopold, 153-167. Englewood Cliffs, N.J.: Prentice Hall.]
- Callow, John C. (1965). Kasem nominals--a study in analyses. *Journal of West African Languages* 2.1:29-36.
- Callow, John C. (1968). A hierarchical study of neutralization in Kasem. *Journal of Linguistics* 4:33-45.
- Carrier, Jill. (1975). Reduplication in Tagalog. MS., MIT.
- Chomsky, Noam and Morris Halle. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Churma, Donald G. (1977). On 'mutually non-affecting' rules. Winter LSA Meeting, Chicago, IL.
- Churma, Donald G. (1979). Arguments from External Evidence in Phonology. Ohio State University Ph.D. Dissertation.
- Churma, Donald G. (1980a). Diachronic evidence for synchronic analyses in phonology. This volume.
- Churma, Donald G. (1980b). A further remark on the 'Halleian syllogism'. This volume.
- Churma, Donald G. (1980c). Rule inversion in Chadic: a closer look. MS., Ohio State University
- Churma, Donald G. (1980d). We must argue about rule ordering. *Linguistic Analysis* 6:41-51.
- Dinnsen, Daniel A., ed. (1979). *Current Approaches to Phonological Theory*. Bloomington, IN: Indiana University Press.

- Donegan, Patricia Jane and David Stampe. (1979). The study of natural phonology. In Dinnsen 1979.
- Goyvaerts, Didier L. (1978). Aspects of post-SPE Phonology. Ghent: E. Story-Scientia.
- Gussmann, Edmund. (1976). Recoverable derivations and phonological change. *Lingua* 40:281-303.
- Halle, Morris. (1959). The Sound Pattern of Russian. The Hague: Mouton.
- Halle, Morris. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry* 4:3-16.
- Halle, Morris. (1978). Further thoughts on Kasem nominals. *Linguistic Analysis* 4:167-85.
- Hetzron, Robert. (1975). Where the grammar fails. *Language* 51:859-72.
- Hill, Jane H. (1970). A peeking rule in Cupeño. *Linguistic Inquiry* 1: 534-39.
- Herbert, Robert K. (1977). Reconstructing phonetic processes in Proto-Austronesian. *Austronesian Symposium*, Honolulu, HI.
- Hooper, Joan B. (1976). An Introduction to Natural Generative Phonology. New York: Academic Press.
- Howard, Irwin. (1969). Kasem nominals revisited. *Working Papers in Linguistics* 1.10:63-99, University of Hawaii.
- Howard, Irwin. (1970). Kasem nominals and the ordering of phonological rules. *Working Papers in Linguistics* 2.3:113-29, University of Hawaii.
- Hudson, Grover. (1974). The representation of non-productive alternations. *Historical Linguistics I*, ed. by John M. Anderson and Charles Jones, 203-29. Amsterdam: North-Holland.
- Hudson, Grover. (1980). Automatic alternations in non-transformational phonology. *Language* 56:94-125.
- Hyman, Larry M. (1970). How concrete is phonology? *Language* 46:58-76.
- Jackendoff, Ray. (1975). Morphological and semantic regularities in the lexicon. *Language* 51:639-71.
- Janda, Richard D. (1977). 'Upside-down' phonology: regenerative or degenerative? Winter LSA Meeting, Chicago, IL.
- Janda, Richard D. (1980). 'Upside-down' phonology: regenerative or degenerative? Bloomington, IN: Indiana University Linguistics Club.
- King, Robert D. (1973). Rule insertion. *Language* 49:551-78.
- Kiparsky, Paul. (1973). Phonological representations. *Three Dimensions of Linguistic Theory*, ed. by Osamu Fujimura, 1-136. Tokyo: TEC Company, Ltd.
- Koutsoudas, Andreas, Gerald Sanders, and Craig Noll. (1974). The application of phonological rules. *Language* 50:1-28.
- Kuhn, Thomas S. (1970). *The Structure of Scientific Revolutions*, second ed., Chicago, IL: University of Chicago Press.
- Lakoff, George (1973). Some thoughts on transderivational constraints. *Papers in Linguistics in Honor of Henry and Renée Kahane*, ed. by B. B. Kachru, 442-52. Urbana, IL: University of Illinois Press.
- Langendoen, D. Terence. (1975). Acceptable conclusions from unacceptable ambiguity. *Testing Linguistic Hypotheses*, ed. by D. Cohen and J. Wirth, 117-27. Washington, D.C.: Hemisphere Publishing Company.
- Leben, William R. (1977a). Doubling and reduplication in Hausa plurals. *Linguistic Studies Offered to Joseph Greenberg*, ed. by Alphonse Juillard, 419-39. Saratoga, CA: Anma Libri.

- Leben, William R. (1977b). On the interpretive function of phonological rules. *Phonologica* 1976, ed. by Wolfgang U. Dressler and Oskar G. Pfeiffer, 21-7. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- Leben, William R. (1977c). Parsing Hausa plurals. *Papers in Chadic linguistics*, ed. by Paul Newman and Roxanna Ma Newman, 89-106. Leiden: Afrika-Studiecentrum.
- Leben, William R. (1979). The phonological component as a parsing device. In Dinnsen 1979.
- Leben, William R. and Orrin W. Robinson. (1977). 'Upside-down' phonology. *Language* 53:1-20.
- Lightner, T. M. (1975). The role of derivational morphology in generative grammar. *Language* 51:617-38.
- Manaster-Ramer, Alexis. (1980). On psychological reality. *CLS* 16, Chicago, IL.
- McCawley, James D. (1979). Comment. In Dinnsen 1979.
- McNeill, David. (1970). *The Acquisition of Language*. New York: Harper and Row.
- Newton, Brian E. (1971). Ordering paradoxes in phonology. *Journal of Linguistics* 7:31-53.
- Phelps, Elaine. (1975). Simplicity criteria in generative phonology: Kasem nominals. *Linguistic Analysis* 1:297-332.
- Phelps, Elaine. (1979). Abstractness and rule ordering in Kasem: a refutation of Halle's maximizing principle. *Linguistic Analysis* 5:29-69.
- Pollack, Jay M. (1977a). Lexical features in phonology. Ohio State University Ph.D. Dissertation.
- Pollack, Jay M. (1977b). 'Upside-down' phonology and natural processes. *CLS Book of Squibs*, ed. by Samuel E. Fox, et al., 74-75. Chicago, IL: Chicago Linguistic Society.
- Robinson, Orrin W. (1977). Abstractness in a phonological parsing model. Winter LSA Meeting, Chicago, IL.
- Robinson, Orrin W. (1980). Rule ordering in a parsing model of phonology. *Phonology in the 1980's*, ed. Didier L. Goyvaerts, 119-39. Ghent: E. Story-Scientia.
- Schane, Sanford A. (1968). *French phonology and morphology*. Cambridge, MA: MIT Press.
- Sommerstein, Alan H. (1977). *Modern phonology*. Baltimore: University Park Press.
- Stampe, David L. (1973). *A Dissertation on Natural Phonology*. University of Chicago Dissertation.
- Sullivan, William J. (1975). Alternation, transformation, realization, and stratification revisited. *The First LACUS Forum*, ed. by Adam Makkai and Valerie Becker Makkai, 472-522. Columbia, SC: Hornbeam Press.
- Wilbur, Ronnie Bring. (1973). The phonology of reduplication. Bloomington, IN: Indiana University Linguistics Club.
- Wilbur, Ronnie Bring. (1975). Morphology and the rule ordering controversy. *Berkeley Linguistic Society* 1.460-74.
- Wilder, Raymond L. (1965). *Introduction to the Foundations of Mathematics*, second edition. New York: John Wiley & Sons.

OSU WPL # 25 (1981) 107-173.

On the Natural Phonology of Consonants*

Roderick D. Goman

0. Synopsis of natural phonology

A natural phonological process is an innate, freely applying constraint on what is normally pronounced. Given a potential phonological opposition, a process will eliminate that member of the opposition which presents the greatest difficulty in terms of pronunciation or perception, or both. In American English (and in some forms of British English), the flapping of /t/ and /d/ to [ɾ], as in Betty [b'ɛɾi], allows the speaker to expend less effort in producing the stop sound, i.e., the application of the process makes the stop easier to pronounce. The English aspiration of voiceless stops before stressed vowels affords the speaker a clearer distinction between these stops and the context in which they occur, i.e., the application of the process makes the stops easier to perceive.

Speakers (especially children), confronted with a new language, at first allow processes to apply that aren't, properly speaking, an inherent part of the language. This observation accounts in children for "baby talk" (child language) and in adults for foreign accents. Speakers confronted with the difficulty of learning a language must suppress or limit the application of the otherwise freely applying processes in order to match the native adult model. Suppressing a process completely negates its effect on the forms of a language, e.g., English speakers must suppress the process which devoices final obstruents in order to distinguish between such words as bet and bed. Limiting a process partially negates its effect, e.g., in English initial obstruent clusters are always voiceless: spat, stop, skit (cf. Russian zdanie 'building'). But medially and finally the devoicing process has been suppressed: Mazda, used [juzd], etc. So, in English the devoicing of obstruent clusters is limited to word-initial position.

Through a procedure of suppressing some processes, limiting others, and allowing still others to apply freely, the individual speaker arrives at the mature pronunciation of the language he is learning. In the case of children, they are normally eventually entirely successful at this. Adults have a harder time.

The suppression or limitation of a process, in terms of either historical change or language acquisition, proceeds in a hierarchical fashion. For instance, the sequence of [hj] in English, as in huge [hjudz], is sometimes simplified to [j], e.g., yuge [judz]. The sequence [hw] is, for most American English speakers, simplified to [w], e.g., what [wat], instead of [hwat]. The [h] in such sequences as [hn hl hr] has been lost historically in Middle English times. From these facts, we can set up this hierarchy of h-loss:

(0.1) n l r w j

Sounds at the left end of the hierarchy simplify h-clusters earliest and

most widely. Those at the right end simplify last and least widely. The loss of [h] in English has worked its way from left to right in this hierarchy, [n l r] having simplified first, [w] next, and last of all, [j] is now in the process of simplifying (cf. Stampe 1973).

0.1. Process vs. rule.

Processes regularly perform minimal substitutions in accordance with the principles outlined above, in addition to principles yet to be discussed. Processes represent an innate (intuitive or instinctive) restriction on what regularly passes for a linguistic message. Because they are freely applying, there is no linguistic "cost" involved in allowing them to apply. The "cost" comes when the speaker must exert an effort to keep them from applying, in terms of the suppression and limitation just discussed.

Opposed to processes are rules, which are learned and not innate. Rules are generally morphophonemic in character and do not represent limitations on pronunciation but are merely traditional or customary ways of handling phonological particulars inherited from former stages of the language. Processes, on the other hand, are by and large allophonic and represent real limitations on what can be pronounced.

The "velar-softening rule" in English, in which s is substituted for k in certain instances (see Chomsky and Halle 1968), is indeed a rule and not a process. It produces alternations like electric with final [k] vs. electricity with [s] before -ity. But notice that this rule is not a limitation on what one can say but merely a customary substitution of one sound for another (apparently borrowed from French). It is easy to fly in the face of custom and say electrickity, pronouncing [k] where we would normally expect [s].

The alternation in English of [k] and [k̠] is due to a process. Unlike the rule above, the average speaker does not consciously control the use of these two alternates, instead using [k̠] only in the neighborhood of palatal vowels, and [k] elsewhere. Thus the spoonerism of sitcom [sɪtkam] is not *[kɪtsam] but [k̠ɪtsam], revealing the palatal alternate at the phonetic level. So while rules do not limit pronunciation, processes are real limitations constraining what is normally pronounced and defining, language by language, what sorts and sequences of sounds are to be considered pronounceable.

0.2. Fortition vs. lenition.

There are two types of processes: fortitions and lenitions. Fortitions (or paradigmatic processes) are based on the requirement that linguistic messages must sound different to be understandable. Fortitions govern such things as whole sound systems and lexical representation. They optimize the perception of individual segments, i.e., they are based on the notion that individual segments must be distinct from their environment in order to be more easily perceived. They are dissimilative and most noticeable in slow, precise (hyperarticulate) speech. While they may apply in a context-free way, they are often catalyzed, i.e., their application may be triggered by the context in which they occur. The aspiration of voiceless stops in English, which was mentioned at the opening of this article, is an example of a fortition.

Lenitions (or syntagmatic processes) are based on the assumption that, in a given setting, only a few different linguistic messages are normally possible and that these need only differ in a relative way to be understandable. Lenitions are sequence-optimizing, i.e., they provide strategies for pronouncing sequences of segments. They are based on the principle of "least effort" (ease of pronunciation) and thus are assimilative and most apparent in casual, fast, sloppy, or lazy (hypoarticulate) speech. They are always context-sensitive. The flapping of /t/ and /d/ in American English, mentioned at the opening of this article, is an example of a lenition.

0.3. Explanation of terms.

The remainder of this article makes use of a number of terms which may be unfamiliar to the reader. It is best to introduce these terms in anticipation of their use in the text.

There are five basic positions (tonalities) referred to in the text: (1) chilic (= labial), (2) coronal, (3) dorsal (= velar), (4) radical (= pharyngeal), and (5) glottal. There are likewise five colors (chromaticities): (1) labial, (2) palatal, (3) velar, (4) pharyngeal, and (5) rhotacized (= r-colored). So, while labial refers to color, chilic refers to position--both being articulated with the lips.

There is a rather complicated set of terms referring to types of sonority and stricture; fortunately, most of these are not unfamiliar terms. There are five types of sonority. From least sonorant to most sonorant they are: (1) stop, (2) fricative, (3) nasal, (4) approximant, and (5) vowel. Stops and fricatives are obstruents; nasals, approximants, and vowels are sonorants. Nasals and approximants are sonorant consonants, or resonants.

There are four types, or degrees, of stricture: (1) closure (stops and nasals), (2) fricative, (3) approximant (approximants and high vowels), and (4) sonants (nonhigh vowels). The following table summarizes the various sonority and stricture types:

(0.2)	Obstruent	{ STOP }	Closure	} Consonant	} Sonorant
		{ FRIC }	Fricative		
	Resonant	{ NASAL }	Closure		
		{ APP }	Approximant		
	Vowel	{ HIGH V }	} Sonant		
		{ NONHIGH V }			

1. Sonority distinctions.

1.1. Introductory remarks.

Sonority can be defined as the ability of a speech sound to carry a syllable. So, typically, we would not expect that a stop, like t, would be a syllabic nucleus. On the other hand, we would expect a vowel, like a, to be one. We would thus expect vowels to occur centrally in the syllable and stops (and other consonants) to be peripheral to the syllable, as in the nonsense syllable [tat]. Vowels are thus demonstrably more sonorant than stops (and other consonants).

This is not to say that vowels, when they occur, always carry the syllable. Sometimes they are simply closely associated with the nucleus of the syllable, being part of the vocalic nucleus, but not the main part of it, as with the nonsyllabic vowel \dot{i} in the nonsense syllable [t \dot{i} it].

Further, this is not to say that consonants cannot carry a syllable. Resonants in English, like m n l, are capable of carrying an unstressed syllable, e.g. bottom, button, and bottle, which end in syllabic m n l. Also compare stressed syllabic trilled r in Serbo-Croatian, e.g. srce [s'rt \dot{s} e] 'heart' (personal comm., Ilse Lehiste; Kondrašov 1962).

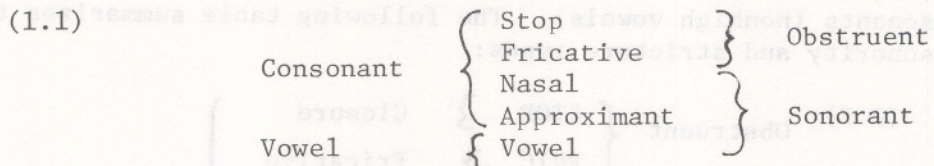
Obstruents can also carry syllables, e.g., English interjections like psst! and shh! In Itel'men, a Paleo-Siberian language of central U.S.S.R., there is at least one word with no vowels: ckpc 'spoon' (Skorik 1968:238).

But nevertheless, gradations of sonority exist among consonants. In English, obstruents and resonants, as we noticed above, are used in two different ways. Obstruents are used to carry syllables only in interjections, and as Jakobson (1968) points out, imitative or suggestive sounds are outside the purview of normal phonology. Resonants in English, on the hand, carry syllables in nonimitative lexical items, like those mentioned above for m n l. We would expect, given these considerations, that resonants are the more sonorant ones, and this is exactly right.

The four types of sonority for consonants are: stop, fricative, nasal, and approximant. These four are classified consonants and specified in features as C. They oppose sounds classified as vowels, specified V, or -C.

Stops and fricatives are obstruents; nasals and approximants are resonants. Obstruents are specified Obs and oppose sonorants, specified -Obs. Sonorants include both resonants and vowels.

So, we have the following categories of speech sounds:



In implicational terms, we can see that: (1) all vowels are sonorant, (2) all obstruents are consonants, and (3) resonants--the middle case--are sonorant consonants.

The four sonority types for consonants suggest that there is a linear gradation of sonority with stops as least sonorant and approximants (and vowels) as most sonorant. Grammont (1933:99) reports the following degrees of sonority:

- (1.2)
0. occlusives (stops)
 1. spirants (fricatives)
 2. nasals
 3. liquids (l, r)
 4. semivowels (j, w, y)
 5. high vowels (i, u, y; \dot{i} , \dot{u} , etc.)
 6. mid vowels (e, o, ϕ)
 7. low vowels (a).

Here the numbers 0 through 7 represent "degrés d'aperture." Saussure (1959:44ff.) reports a similar hierarchy which distinguishes only seven degrees of aperture (0-6) combining liquids (Grammont's degree 3) and semivowels (degree 4) into a single group "liquids" (47).

Such hierarchies are confirmed by Zwicky's more recent (1972) study of sonority. The study is based on comparisons of slow-speech and allegro forms of speech by speakers of American English. It depends on such rules (or processes) as Slur, which deletes [ə] following any consonant and preceding r, l, or n plus an unstressed vowel, e.g., happening becomes happ'ning. Zwicky delineates the exceptions to Slur which would not otherwise result in an unpronounceable sequence. These include those examples in which schwa occurs before stressed syllables, clusters, obstruents, the nasal m, and across word boundaries. Referring to the hierarchy given below in (1.3), Zwicky states:

I shall argue that the hierarchy...[in 1.3/RDG]...corresponds to a differential in the acceptability of the outputs of Slur, ranging from normally entirely acceptable before [r], to less acceptable before [n], to normally unacceptable before [m] and [ŋ], to entirely unacceptable before obstruents... (285)

Based on such rules as Slur which apply hierarchically, Zwicky presents the following sonority hierarchy:

(1.3) Stop Fric η m n l r Glide Vowel

In this hierarchy sonority decreases toward the left and increases toward the right. Fric represents fricatives. The "r" referred to is, of course, American English [ɹ], not the trill.

Restating Zwicky's hierarchy in terms of the classes discussed above, we get:

(1.4) Stop Fric Nasal App Vowel.

Here Fric is fricative and App is approximant. We are therefore left with the conclusion that nasals are less sonorant than approximants. Further, within the class of nasals, the velar (dorsal) nasal is least sonorant, the labial (chilic) in the middle, and the coronal most sonorant. Also "liquids" like l ɹ are less sonorant than "glides" like w j. And further, l is less sonorant than ɹ. We reach similar conclusions below, in sec. 1.2.3.

We turn now to a discussion of perceptual, articulatory, and acoustic considerations of consonants and vowels vis-à-vis the question of sonority.

1.1.1. Perceptual, acoustic, and articulatory properties of consonants.

In terms of perception, the presence of sonority is linguistically less salient than its lack. Consider, for instance, the fact that consonants (of low sonority) can be used as the only markers of words ("letters") in an orthographic system, as for example that of Hebrew. Vowels (of high sonority), on the other hand, would not offer unique orthographic representations, e.g., the consonant sequence ct could stand for cat or act, but the significance of the vowel letter a is less obvious.

In acoustic terms, sonority peaks are marked by a spectrum in which the various formants transit rapidly from their consonant values to their steady-state values for vowels and then back to their consonant values. The steady-state periods for vowels are longer in duration than those for consonants. Further, sonority valleys like stops are largely inaudible (spectrographically invisible), being defined instead in terms of the various transitions which they exhibit and not in terms of an identifiable steady-state.

It is clear from these considerations of the perceptual and acoustic properties of speech sounds that two separate phonological teleologies are involved. On one hand, it is clear that consonants are perceptually more important. This goes hand in hand with the idea that consonants are semantically more salient. On the other hand, it is clear that vowels are acoustically more important. They are the more audible "carrier waves" of speech, broken up by the intrusion of less audible consonants. In a word, consonants are semantically indicative; vowels are semantically ambiguous. Further, vowels are more clearly audible; consonants are less so.

In articulatory terms (after Catford 1977), with reference to stricture, closures like stops and nasals are clearly consonants and so are fricatives. Stops, which are largely inaudible and identified by their transitions, have already been discussed. Fricatives have a more audible spectrum involving diffuse areas of noise. They are identified partly by the frequency around which this noise gathers, partly by their transitions, like stops. Nasals, like vowels, may exhibit a steady-state but they are identified, in part, by their transitions; they can be distinguished from other consonants by resonant frequencies in their spectra corresponding to the resonancies of the nasal pharynx (cf. Lehiste 1970:156).

Sonants are clearly vowels, having steady-states in which the formants move toward the average frequency of the oral resonator. These consist of the nonhigh vowels.

But approximants, which represent the last type of stricture, may be either consonants or vowels and stand on either side of the line between these two classes. Acoustically, approximant consonants (or more simply "approximants") exhibit dynamic formant transitions with virtually no steady-state. Approximant vowels ("high vowels") exhibit steady-states but ones which are more extreme in comparison with those of sonant ("non-high") vowels. The formant values of high vowels are further from the norm, further from the average resonant frequency of the oral resonator. Approximants which correspond to high vowels, such as the correspondence between j and i, or w and u, show a similar displacement. But unlike the vowels with their steady-state, the dynamic spectra of approximant consonants reflect a rapid movement of the articulators to and from their place of articulation. For j and w, this dynamic mode serves to differentiate them from i and u, their corresponding vowels. For approximants lacking a corresponding vowel, such as l r ʝ, the movements, while more deliberate, apparently operate on the same dynamic basis. In the case of the flap ɾ, the tip of the tongue taps against the roof of the mouth; for the trill r̄, it taps repeatedly and intermittently.

In such a way, arguments for the various types of sonority may be based on: (1) perception, to give the distinction between sonority valleys and sonority peaks (consonant versus vowel), (2) acoustic information, to

verify this same distinction and further to indicate the four stricture types (closure, fricative, approximant, and sonant), and (3) articulatory considerations, to verify the stricture types and further to account for the difference between approximants and high vowels. As we shall see, more discerning arguments may also be based on processes affecting sonority (as in sec. 1.2) and sequences of segments which reflect sonority.

1.2. Sonority and process.

1.2.1. Fortition and lenition.

In arguing for the various classes of sonority and the features which express these same classes, it is sufficient to show how natural processes affect speech sounds in general. Of course, we are not totally in the dark about what these sonority classes should be and the features which should be used to express them, so natural processes appear to serve simply to confirm or disconfirm our initial impressions. But more than this, natural processes show us how individual sounds are perceived, in terms of mental representations, and how sequences of sounds are integrated into pronounceable units, in terms of ease of articulation.

There are two types of processes: fortitions, and lenitions. Fortitions, also called paradigmatic processes, affect mental representations of speech sounds, in terms of their being perceived as individual sounds. Fortitions not only affect phonological representation, they also exhibit the endeavor of the conscious mind to make understandable conversation with the external world. Such understandability is based on the succinct notion that words with different semantic content should have different phonological forms. Fortitions thus account for the concept of phonological differentiation.

Generally, the language succeeds in making different messages bear a different form. Nevertheless, entire sentences may be phonologically ambiguous, as in these two sentences given by Hockett (1958:15):

- (1.5) a. The sons raise meat.
b. The sun's rays meet.

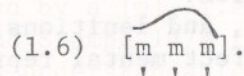
Although these are pronounced exactly alike--i.e., phonologically, they cannot be differentiated--such an arrangement is rare. If this were not so, we would be forced to comprehend the world and the speakers in that world on the basis of animalistic (and somewhat psychic) inferences. (And perhaps this is the way that the child originally perceives the problem involved in learning his native language.)

While the concept of the fortition and its application to paradigmatically defined speech acts is largely the creation of the subconscious mind, which views everything as relative to the present moment, it is the endeavor of the conscious mind to make these internal relationships tantamount to absolutes. So, when it comes to any phonological parameter--whether it is sonority, position, or color--the upper mind tries to look beyond the present moment in order to conclude that the difference between p and b is not one of voicing but is in fact a global difference. Further, an occurrence of p in the word pin in the present moment is perceived to be a repetition of the p in pen, pan, pun, etc., and further, it is concluded that the p in such words is an instance of the same sound. These are not unreasonable conclusions.

Language has essentially two facets: the external and the internal. In terms of the speech act, there is at once the observation of the external parameters of speech in terms of pronunciation and comprehension ("phonetics") and the inner motivation and perception of that self-same speech act ("phonology").

Fortitions account for phonological differentiation and also reinforce it. They not only refer to our perception of the speech act, they also serve as a model of it. Fortitions regulate what sort of thing can count as a mental representation, or mental intention, concerning speech. Lenitions, on the other hand, lead to phonetics. They regulate our notions about what is a suitable or favorable utterance.

Lenitions, also called syntagmatic processes, are based on syntagmatic concepts of speech and thus account for the need for ease of articulation. They are strategies for pronouncing sequences of segments and are based on the pragmatic notion that in a given context there are generally only a few messages that are likely to be given and these need differ only in a relative way from all the other likely messages. So, for instance, if a question has been asked, the response might be:



This message represents the sentence "I don't know." Only the "melody" or intonation of the utterance remains, as indicated by the curved line, the segmental part being reduced to three moras of a syllabic bilabial (labial) nasal.

1.2.2. Evidence for sonority types.

Processes themselves provide evidence for the categories of sounds which may properly be attributed to the phonological parameter of sonority. For instance, consider the common change of w to v which occurs in such Indo-European languages as Sanskrit and the Romance family, as well as the Slavic family, northern Germanic (Scandinavian), and western Germanic except for English. I have observed it in child language and it is also common in such non-Indo-European language groups as eastern and nuclear Polynesian, e.g., Tahitian (eastern) and Samoan (nuclear), and the related Tongic family, e.g., Tongan (Biggs 1971). In Hawaiian, an eastern Polynesian language, the w:v distinction appears allophonically, w occurring after labial vowels, v after palatal vowels, and either (apparently depending on the dialect or disposition of the speaker) after achromatic /a/ and initially, e.g. Hawai'i [həvʌʔi] or [həwʌʔi].

The change from w to v can be analyzed this way:



The sound w may be analyzed as a velar-labial approximant (as a dark velar "l" made with the lips rounded). The first step (w to ω) simply eliminates the weaker of the two colors, the velar one (w). The second change (ω to v) also eliminates the labial coloring (lip-rounding), leaving behind a bilabial (nonlabial) approximant pronounced with spread lips. The third step (v to β) narrows the stricture one notch, from approximant to fricative.

And the fourth and last step (β to v) changes the bilabial to a labiodental one (dentalizes it).

The third step (v to β) is the one we are interested in here. It establishes the existence of two types of sonority neither of which involves total closure of the oral cavity (as with stops and nasals). This change, involving the narrowing of the stricture, is similar to the change of j to z in Puerto-Rican Spanish (and elsewhere in the Spanish dialects) and also the change of r to z in the central and southeastern dialects of fifteenth-century French, as in the change of chaire 'chair' to chaise (Pope 1934: 157f.). The process producing such narrowing of stricture is a fortition called, appropriately enough, narrowing (see sec. 3.1.1. below).

It is also possible to find examples of widening, produced by a contradictory process of that name (see sec. 3.1.2.). An example is rhotacism (as in western and northern Germanic and Old Latin), first voicing s to z between vowels. The widening then occurs as z becomes r . Proto-Germanic final z also widened to r in northern Germanic; cf. Gothic (Eastern Germanic) dags, Old Icelandic dagr, and Runic Norse dagaR, all meaning 'day (nom. sg.)' (Moulton 1972).

Other facets of these same processes reveal other sonority types. For instance, the change of the sound θ to \dagger in languages lacking θ establishes the sonority type called stops. The English substitution of k for x , e.g. [bɔk] Bach, is a similar example.

Changes to and from nasals also occur, but less commonly, and establish this fourth sonority type. A student of mine from Hong Kong, whose native language was Cantonese, regularly substituted n for l in her English, e.g. nook for look. A child speech example relating nasals to stops is Joan Velten's denasalization of nasals to voiced stops, e.g., [bub] for broom and [bud] for spoon (Velten 1943).

Such an exposition of sonority types can easily be expanded, but at least it establishes the four main types. Other lines of argument are needed to determine the tensing or laxing of sounds, but this will be left to sec. 3.2 below.

1.2.3. Sonority within the syllable.

Another kind of evidence that can be used to establish sonority types and relative sonority, incident to the process, is the evidence of sonority patterning within the syllable. David Stampe (pers. comm.) has suggested that the patterning of types of segments within the syllable is in a fairly direct relation to their sonority. Basically, he suggests that the nearer to the nucleus of the syllable a segment is, the more sonorant it is. Conversely, the more removed or separated it is from the nucleus, the more obstruent it is. All this is subject to language-specific variations and exceptions, but the basic tenet stands.

For instance, in an English word like pit, the i is the nucleus and the p and the t are peripheral to it. This syllable may be analyzed as CVC, or more revealingly, OVO (where O = obstruent). In a word like print another part of the sonority hierarchy is revealed since r and n represent resonants. This syllable can be analyzed as ORVRO (where R = resonant). This example may bother some more discerning readers since print is phonetically [pɹɪnt̚] (no nasal remaining). Consider then quilt.

Consider the word ironed [aɪrənd], which is pronounced as one syllable. This syllable can be analyzed as V̄VANO (where V̄ = a non-syllabic vowel,

or "glide", A = approximant, and N = nasal). This effectively demonstrates that nonsyllabic vowels are less sonorant than syllabic ones, yet more sonorant than resonants. And the [ɹŋ] sequence divides those resonants into approximants (more sonorant) and nasals (more remote from the nucleus and thus less sonorant). (Cf. Saussure's 1959 distinction between explosive and implosive parts of a syllable. The sounds of ironed would all be implosive, except for [ɹ], which is neither (51ff.))

Also consider the word girl, phonetically something like [gɹl] and phonemically probably /gɹl/. Using the phonemic form as a model (or extrapolating from the phonetic form), we can see that the ɹ is more sonorant than the l. (Recall Zwicky's hierarchy at the opening of this section.) As further evidence, note that if the two approximants were reversed, a two-syllable utterance would result: [gɹlɹ].

Sonority gradation patterns within the syllable also suffer what might be termed language-specific exceptions. For instance, consider the English word scamps. The two instances of s are more sonorant than the two stops (represented by c and p) but nevertheless are more remote from the nucleus. They represent then sonority peaks subordinate to the main sonority peak represented by a [æ], i.e., the s's are satellite sonority peaks. (Cf. Donegan and Stampe 1978 for their discussion of the German word Stumpf.)

It is understandable that such satellite half-syllables then would cause learning problems for children. Children usually handle this problem by combining the position of the stop with the sonority of the s, e.g., spoon comes out foon. Only later do they acquire the adult pronunciation.

At any rate, sonority patterning within the syllable accounts for the various gradations, or types, of sonority of the kind we have discussed. They are based on strategies associated with connected speech such as are reflected in lenitions (see sec. 5 below).

2. Position and color distinctions.

2.1. Introductory remarks.

Position and color may best be defined in terms of their perceptual properties. The perceptual property of position is tonality. In a word, position is the tonality associated with a given articulation. The perceptual property of color is chromaticity. Color is chromaticity, as I will be using the word here, so that any particular color is the chromatic type associated with its particular articulation.

It remains then to specify the meaning of tonality and chromaticity. This will be done in the following two sections, beginning with tonality and ending with chromaticity.

2.1.1. Tonality.

Tonality may be briefly defined in terms of the tonal quality with which various sounds are associated. Tonal qualities associated with the various positions depend on three independent factors: (1) The tonal quality may be lower or higher (darker or lighter). Lower (darker) tonal qualities are expressed in features as grave (Grv); higher (lighter) ones, as acute (-Grv). (2) The tonal quality may be lingual or nonlingual. Lingual tonal qualities are, naturally enough, specified as lingual (Ling.); nonlingual ones, as nonlingual (-Ling). (3) The tonal quality may be implosive or explosive (somewhat in the Saussurean sense). Implosive tonal qualities are specified as retracted (Ret); explosive ones, as advanced (-Ret).

We can thus set up the following feature matrix. Here the four stops represent four positions: k represents dorsal, p bilabial, t coronal, and ? glottal.

(2.1)		k	p	t	?
	Grv	+	+	-	-
	Ling	+	-	+	-
	Ret	+	-	-	+

The typical order of voiceless consonants occurring in coarticulations (combinations of consonants having the same sonority but different positions) is as above in matrix (2.1). That is, k regularly precedes p t ? in coarticulations, e.g., kp kt k?. Then p regularly precedes both t and ?, e.g., pt p?. But t regularly precedes only ?, e.g., t? (combinations like ?t being analyzed as ?d). Much the same is true for nasals, i.e., their corresponding order would be η m n (there being no glottal nasal). So, we would expect the coarticulations ηm ηn mn, which typically do occur.

Generalizing on these two sets of examples, the stop order k p t ? and the nasal order η m n, we can make the following statement:

- (2.2) Within a given sonority type, the degrees of tonality function as a way of determining the order of sequence in combinations (specifically, coarticulations) of consonants. The order is as follows: (1) Grave precedes nongrave. (2) For the grave sounds, lingual precedes nonlingual. (3) For the nongrave sounds, the same thing applies: lingual precedes nonlingual.

Thus the function of tonality is to structure the syllable within the constraints of sonority. It is assumed that the above rule (2.2) applies within the limits of a single syllable; otherwise, no coarticulation would be involved.

2.1.2. Chromaticity.

Chromaticity is best examined in relationship to the articulations and interrelationships of the members which compose it. These members consist of the five colors, namely: labial (Lab), palatal (Pal), velar (Vel), pharyngeal (Phar), and rhotacized (Rho). Of these five, Donegan (1978) has identified the first two, labial and palatal, as the "primary colors". This is altogether fitting since these two colors are rarely lacking in a given language. They are certainly the most basic, with palatality taking an edge over labiality, e.g., Donegan (1978:47) reports that for vowel systems, there are languages with palatal vowels and no labial ones, but no cases of labial vowels without palatal ones.

Labiality, of course, is lip-rounding; and palatality is j- or i-coloring produced by approximating the blade of the tongue to the (hard) palate. This means that both labiality and palatality are advanced (nonretracted) sounds, wholly isolable to articulations made toward the front of the mouth. It is probably no accident that the two most basic colors are articulated toward the front. Front sounds are the first to be learned by children, and their existence is categorically implied by the existence of backer consonants (cf. Jakobson 1968:53).

The other three colors are all retracted, being articulated further back in the mouth. Velarity is ω -coloring produced by approximating the back of the tongue to the velum, or soft palate. Pharyngeality is α -coloring produced by approximating the root of the tongue toward the pharyngeal wall. And rhotacism is r -coloring produced by approximating the extreme lower end of the root of the tongue toward the lower pharyngeal wall.

If we were to classify the five colors according to classes of tonality, as addressed in the section above, the following would be the result:

(2.3)

	ω	J	υ	α	r
Grv	+	-	+	+	-
Ling	-	+	+	+	+
Ret	-	-	+	+	+

In the matrix above, labiality is represented by its approximant symbol ω , palatality by J, velarity by υ , pharyngeality by α , and rhotacism by r .

One thing is perfectly obvious from the above matrix and that is that the three features used to differentiate the four tonalities leave υ (velar) and α (pharyngeal) undifferentiated. They are both marked plus (+) for grave (Grv), lingual (Ling), and retracted (Ret). This can be remedied by adding either the feature high (High) or the feature low (Low). The colors J and υ are High; the colors α and r are Low. Three features are all we really need, as follows:

(2.4)

	ω	J	υ	α	r
Grv	+	-	+	+	-
Ret	-	-	+	+	+
High	-	+	+	-	-

In such a way, the colors can be distinguished in terms of tonality. The advanced colors are the basic ones; the retracted colors are of a more subtle nature.

2.1.3. Chromaticity and tonality in synesthesia.

Chromaticity (the property of colors) may produce visual images or evoke colors, e.g., a color like red or white may be associated with the vowel [a]. This is an example of a synesthesia, in which the stimulation of one sense evokes the sensation in another; in the case at hand, hearing evokes sight. Synesthetes, i.e., people with synesthesia, reported on by Jakobson (1968), reveal a constant and clear sound-color agreement which they have in most cases perceived since childhood. A typical case is SP, a speaker of Czech, who has the following sound-color agreements for vowels occurring in Czech:

(2.5)

i	canary yellow	u	dark blue
e	light green	o	blue-red
		a	red

Langenbeck, Deichmann, and Argelander (as reported by Jakobson 1968:82f.) all report similar photisms for vowels, e.g., Argelander reports that white is most often associated with i, black with u, yellow with e, brown

with o, and red, white, or a dark color with a. These color sensations are synesthetic responses to the chromaticity of the vowels, as I am using the term here.

SP reports that consonants evoke colors that are mainly various shades of gray with brighter overtones, some of the brightest overtones being connected with consonants with palatal coloring. Langenbeck, a synesthete himself, reports that for him consonants are all equally colorless. These color sensations, or relative lack of them, are synesthetic responses to the tonality of the consonants, as I am using the term here.

But the meaning is evident, and the analogy is particularly pertinent since it is based on perceptual data due to the psychology of synesthesia. Namely, chromaticity is associated with bright colors and gross distinctions of color. Vowels (and chromatic consonants) have bright colors. However, tonality is associated with various shades of gray or with various degrees of lightness and darkness, more subtle color distinctions. Consonants are various shades of light and dark.

Colors--i.e., palatal, labial, and so forth--represent gross distinctions of phonological coloration, e.g., they have great chromaticity. The vowels, which all exhibited bright colors, are highly chromatic. The approximants of color (ω j ɥ α ɾ) are also highly chromatic. Consonants without color exhibited low degrees of chromaticity and thus need to be distinguished on the basis of tonality.

2.1.4. Consonant-vowel agreement.

Various phonologists from Panini on have assumed that certain consonants correspond to certain vowels. For instance, consonants like c j correspond to the vowel i, agreeing with it in certain aspects of chromaticity and tonality. For the set c j i, all are of the same chromaticity, palatal, and of the same tonality, coronal--granted that i can be referred to as a coronal vowel.

Trubetzkoy (1969) proposed a linear arrangement of consonants and a triangular, more or less linear, arrangement of vowels. In such a system it was impossible to correlate consonant and vowel subsystems, and Trubetzkoy made no attempt to do so.

But the Indian phoneticians, starting with Panini (as far as we know), used a system of interrelationships among the consonants and vowels of Sanskrit. These sounds are given below in Table 1, as adapted from W. S. Allen's Phonetics in Ancient India (1953:20). The chart itself reads in a reverse order to what we would normally expect, because the Indian grammarians started with those sounds "nearest to the origin of the airstream" (Allen 1953:48) and from there they "work upwards and forwards towards the lips" (ibid).

Of the five orders of vowels, a is considered to be glottal or pulmonic, l is palatal, ṛ is retroflex, ḷ dental, and u labial. Referring to the stop consonant in each file of the chart, we can establish the following basic consonant-vowel correlations:

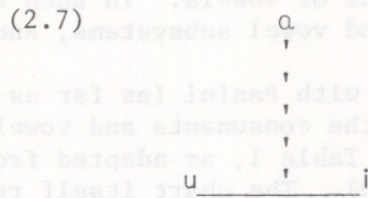
(2.6)	h	k	ç	ṭ	ṭ	p
	a		i	ṛ	ḷ	u

Table 1
The sounds of Sanskrit

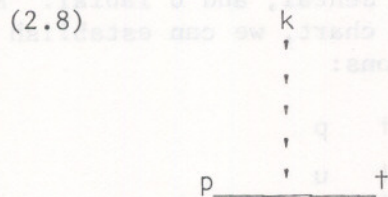
	glottal or pulmonic	velar	palatal	retroflex	dental	labial
Closures		k	č	ṭ	t	p
		kh	čh	ṭh	th	ph
		ḡ	Ḷ	ḍ	d	b
		gh	Ḷh	ḍh	dh	bh
Non-closures		ṅ	ñ	ṇ	n	m
		(x)	ś	s	s	(ϕ)
Vowels	h		i	ṛ	ṛ	u
	ḥ		ī	ṝ	ṝ	ū
	a		e			o
	ā		ai			au

So a is correlated with h, i with č, u with p, etc.--and there is no vowel corresponding to k.

Jakobson (1968:73-81) proposes an ontogenic view of consonant and vowel correlation with the following chart (74).



He notes that the u/i process (the base line) expresses lightness and darkness distinctions, where i is considered light and u dark. The a process, on the other hand, expresses the degree of chromatism. A similar chart is given for consonants (ibid.):



So, k corresponds to a, p to u, and t to i. Cf. the Indian phoneticians' h (not k) to a, c (not t) to i, and both agree on p to u.

Even Jakobson's feature system responds to this relationship (as given in Jakobson and Halle (1956:29ff.)). It is summarized in the following feature matrix:

(2.9)

	k	a	p	u	t	i
consonantal	+	-	+	-	+	-
compact	+	+	-	-	-	-
grave	(+)	(+)	+	+	-	-

Chomsky and Halle (1968) added the features high, low, and back not only for vowels but also for consonants. This produced the agreements given in Table 2. Vowels in the table are given above, and consonants and glides below. Thus č, c (k), y (j) correspond to i, k and w to u, p and t to e, q to o, h and ? to ə, and finally to a. All of these agreements are drawn from the chart on p. 307 of Chomsky and Halle (1968). The consonant features anterior and coronal, for which vowels are all specified negatively, have been omitted.

Table 2

Chomsky and Halle's system of agreements

Vowels	i	u	e	o	ə	a
high	+	+	-	-	-	-
low	-	-	-	-	+	+
back	-	+	-	+	-	+
Labials	p ^y	p ^w	p	p ^o	p ^ə	p ^a
Dentals	t ^y	t ^w	t	t ^o	t ^ə	t ^a
Palato-alveolars	č	c ^w	c	c ^o	c ^ə	c ^a
Palatals	c					
Velars		k				
Uvulars				q		
Pharyngeals						ħ
Glides	y	w				h, ?

While the columns of the chart under i, u, o, and a express plausible enough sound relationships, it is hard to see how p and t are related to e, and how h and ? are related to ə.

Based on what I have learned about consonants and their interrelationships, and their relationships to vowels, I would suggest the following agreements:

many speakers of American English: it is the dark - | sound for some speakers in little [ɹɪ] or, for some, even [ɹɪ]. And the radical (pharyngeal)

(2.10)	p	†	†̣	†̣	k	q	ḳ	?
	ω	-	ɹ	ɹ̣	ɥ	ɥ	α	-
	-	-	i	ị	u	ɤ	ọ	-

In the first row are stops, representing the individual positions: chilic (p), coronal († † †̣), dorsal (k q), radical (ḳ), and glottal (?). In the second row is the corresponding color, if any. And in the third is the corresponding vowel.

We should first note that the vowels e æ are lacking. These are internal vowels, being remote from the external areas where consonants are articulated.

Chilic p has no corresponding vowel; however, it does have the color ω, which applies to rounded consonants and vowels. The dorsal labial consonant k^ω would correspond to w (= ω^ω).

It is probably significant that neither plain coronals (†) or glottals (?) have any corresponding colors or vowels. These two have the highest tonality of all the stops.

In the column headed by †̣ (retroflex), the corresponding color is ɹ (rhotacized) and the vowel ị (high central unrounded). These three would all be specified as -Grv Ret Ling (see sec. 2.2. just below).

The q (uvular) column corresponds to uvular coloring (if there be such), represented by the symbol ɥ. The corresponding vowel is the mid (-High -Low) back vowel ɤ.

The argument for these consonant-vowel agreements carries over into the relationships among position, color, and the application of processes, to which we now turn.

2.2. Position, color, and process.

Evidence for the various types of position and the various types of color is revealed in the operations of processes. We start with a consideration of position (primary and secondary) and continue with a discussion of color, which creates various secondary positions. A consideration of the ontogeny (individual development) of position and color concludes our discussion.

2.2.1. Position and process.

The primary positions are easily established according to the workings of processes. The primary positions are the five given for the primary consonants in sec. 0 above: chilic through glottal.

The fortitions affecting position, appropriately called fronting and backing, normally front or back a consonant one "notch" on the position continuum. For instance, chilic p can be backed to coronal †, one position further back, as in the speech of Elizabeth Stampe at age 1:6 (pers. comm., David Stampe). Coronal † in turn was backed historically in Hawaiian, and † and k alternate in various styles of Samoan (Biggs 1971). I know of no language in which k backs to ḳ; however, the velar approximant ɥ backs (or alternates with) pharyngeal α in the individual speech patterns of many speakers of American English: it is the dark -l sound for some speakers in little [ɥiɹ̣ɥ] or, for some, even [αiɹ̣α]. And the radical (pharyngeal)

fricatives $\text{ħ} \text{ʕ}$ in proto-Semitic reportedly became ʔ in Akkadian, along with several other sounds, e.g., uvular q and the glides $\text{j} \text{y}$ (Gray 1934:10-20).

In such a way, the primary positions may be established. I have used here the process backing; its opposite, fronting, would have given less complete results, e.g., I know of no change which fronts ʔ to k̟ or k̠ to k . However, Walter, a first-grade pupil, when asked what the sound of the letter h was, replied "Huck". Apparently, huh [$\text{h}\text{ʌ}$], the expected answer, took on a final glottal stop (became [$\text{h}\text{ʌ}\text{ʔ}$]), which in turn fronted to k (to give [$\text{h}\text{ʌ}\text{k}$]).

Chilic has two subpositions, bilabial and labiodental, that we need to consider. Similarly, coronal and dorsal have likewise two subpositions, neither of which involve color. The two coronal subpositions are (inter)-dental and (post)alveolar; those for dorsal are velar and uvular.

There is a process that dentalizes chilic fricatives and coronal smooth fricatives; namely, $\text{ɸ} \text{β}$ become $\text{f} \text{v}$, and $\text{ɸ} \text{β}$ become $\text{θ} \text{ð}$. Speakers of languages with $\text{f} \text{v}$, e.g., English, hear instances of $\text{ɸ} \text{β}$ as their own $\text{f} \text{v}$. For instance, the Spanish word saber [$\text{s}\text{a}\text{β}'\text{e}\text{r}$] 'to know' is heard as [$\text{s}\text{a}\text{v}'\text{e}\text{r}$] by beginning students of Spanish, sometimes also as [$\text{s}\text{a}\text{b}'\text{e}\text{r}$]. Notice that both these changes, β to b or v , involve a nontangent sound (β) becoming a tangent one (b or v), i.e., for the articulation of b and v , both the articulators actually touch. In the case of b , the two lips come together; in the case of v , the upper teeth articulate directly against the lower lip. The dentalizing of $\text{ɸ} \text{β} \text{ɸ} \text{ð}$ all involve an accompanying introduction of tangency.

For dorsals, the distinctions velar and uvular commonly show up for stops, k/q , and for fricatives, $\text{x}/\text{χ}$. The backing process has changed x to χ in Castillian Spanish (cf. Harris 1969:196f.).

For coronals, the distinctions dental and alveolar commonly show up for stops t/t (as in Bengali), and fricatives $\text{θ}/\text{s}$ (as in English).

On the basis of tonality properties (without reference to color), we can distinguish the following set of positions and subpositions. I have used fricatives here to symbolize the variations just discussed.

(2.11)	ɸ	f	θ	ɸ	x	χ	ħ	h
Grv	+	+	-	-	+	+	+	-
Ret	-	-	-	-	+	+	+	+
Low	-	-	-	-	-	-	+	+
High	-	-	-	-	+	-	-	-
Dent	-	+	+	-	-	-	-	-

The first three features delineate the five primary positions. Only two more features are needed to distinguish the three subposition variations which we considered above.

2.2.2. Color and process.

The interrelationships of colors themselves are best considered with reference to their ontogeny; this is done in the next subsection (sec. 2.2.3.). But the effect of color on consonant position and the relation of chromaticity to tonality are other matters. Let us start with a consideration of tonality, chromaticity, and sonority.

In languages with simple consonant systems, we can observe the variations of the development of positions, expressed in terms of tonality.

Every language must have at least two distinctions for tonality (otherwise there is no distinction). Yet even languages with simple consonant systems exhibit at least three tonality distinctions. For stop systems, we have /p t k/ in Iwam (a language of northern New Guinea), /p t ʔ/ in (Classical) Samoan, /p k ʔ/ in Hawaiian, and /t k ʔ/ in Oneida (cf. Biggs 1971; Ruhlen 1976). Yet no language has a stop system based purely on chromaticity, e.g. */p p^ω p^j/. It is evident that processes apply that eliminate this possibility. Tonality is thus a more salient parameter for consonants than chromaticity.

Moreover, sonority for consonants seems to measure up to tonality. There are always at least three sonorities, typically stop, nasal, and approximant. So, sonority is also more salient for consonants than chromaticity.

But as Donegan (1978) has pointed out, sonority is more salient for vowel systems than either chromaticity or tonality. Vowel systems like /ī ə a/ are possible, which exhibit height (sonority) distinctions but not timbre (tonality) distinctions. Yet when timbre distinctions are involved, they typically also exhibit chromatic oppositions, generally palatal versus labial. These colors act to maximally differentiate front (palatal) and back (labial) vowels from one another.

The addition of color to an otherwise achromatic (noncolored) consonant generally results either (1) from occurrence adjacent to a vowel bearing a color, or (2) from occurrence adjacent to a color approximant. In the first case, that of the vowel, a consonant becomes palatalized next to a palatal vowel, labialized next to a labial vowel, velarized next to a velar vowel, or pharyngealized next to a pharyngeal vowel. So, tɛ becomes t'ɛ, tɔ becomes t^ωɔ, t^w becomes t^w and t^a becomes t^a. The process which performs this color addition is a lenition called color matching (see sec. 5. below).

In the second case, that of occurrence with a color approximant, the approximant loses its identity as a separate segment and becomes attached to the consonant. In so doing, the consonant takes on a color, e.g., t + j becomes t̄, t + w becomes t^ω, t + r becomes t̄ (retroflex), etc. For t̄j becoming t̄, compare the change of t̄j to t̄š in gotcha (from got you), etc.

In Table 3, an attempt is made at distinguishing eleven subpositions. The features grave (Grv), lingual (Ling), and retracted (Ret) delineate the basic positional distinctions. Then high (High) and low (Low) are used to differentiate some of the backer sounds. Finally, dental (Dent) is used to distinguish labiodentals and (inter)dentals from everything else.

The letters across the top represent the subposition categories: A = bilabial, B = labiodental, C = dental, D = (post)alveolar, E = (alveo)-palatal, F = retroflex, G = (velo)palatal, H = velar, I = uvular, J = pharyngeal, and K = glottal.

Below that, four rows of fricative symbols appear. Fricatives have been chosen as representatives here since they are the commonest occurrences of these various subpositions. These rows are numbered at left. In row (1) are the laterals. They are the least sonorant of the fricatives and are accordingly placed at the top. They are differentiated from other fricative sonority types by being specified tangent (Tan). In rows (2) and (3) are given the grooved (Gru) fricatives. Those in (2) are laminal (-Api); those in (3) are apical (Api). In row (4) the smooth (Smu)

fricatives are given. Any other distinctions of subposition would require the use of the individual color features, e.g., palatal (Pal), labial (Lab), etc.

Table 3

Positions and color distinctions

	A	B	C	D	E	F	G	H	I	J	K
(1)				ʧ	ʧ	ʧ					
(2)				ʃ	ʃ	ʃ					
(3)				ʂ	ʂ	ʂ					
(4)	Φ	f	θ	β	ɸ	χ	χ	χ	χ	h	h
Grv	+	+	-	-	-	-	+	+	+	+	-
Ling	-	-	+	+	+	+	+	+	+	+	-
Ret	-	-	-	-	-	-	-	-	-	-	+
High	-	-	-	-	+	+	+	+	-	-	-
Low	-	-	-	-	-	-	-	-	-	+	+
Dent	-	+	+	-	-	-	-	-	-	-	-

The main additions to Table 3, over the eight distinctions given in (2.8) above, are the inclusion of the alveopalatals (E), retroflexes (F), and the velopalatals (G). They agree for all features except grave (Grv) and retracted (Ret). So, with the addition of velars (H), the following matrix is needed to distinguish them:

(2.12)	ʃ	ʂ	χ	χ
Grv	-	-	+	+
Ret	-	+	-	+

Above, in sec. 2.1.4., it was pointed out that ʃ (represented in (2.7) above by ʧ in the list of consonant-vowel agreements) corresponds to the front (apalatal) vowel i, which is, like ʃ, -Grv -Ret. The sound ʂ, here representing the reflex ʧ of (2.7), corresponds to the central vowel ɨ, which is -Grv Ret. No vowel is like χ, which is Grv -Ret, and so it corresponds to no vowel. The sound χ corresponds to u; both are Grv Ret.

It is clear that front vowels are neither grave nor retracted and that back vowels are both. It is fairly clear that central vowels are retracted but nevertheless nongrave. The color rhotacized (Rho), symbol r̥, is the approximant related to retroflexes like ʂ and accordingly this color appears almost exclusively with retroflex coronal consonants and central vowels. It also appears with nonretroflex coronal approximants like r̥. I know of no examples of rhotacized occurring with noncentral vowels. When it does occur with central vowels, it occurs with lower ones, e.g. ə a to give ər̥, (= ɛ) and ar̥. This is due probably to the fact that rhotacized is specified Low, being produced in the lower pharynx.

The sound ʂ is typically to be identified with the retracted consonant χ, which often becomes substituted for it (e.g., Slavic ʂ produced by the ruki rule, becomes backed to χ). Accordingly, both ʂ and χ are specified Ret, but because of its lighter tonality, ʂ is specified -Grv.

The sound \check{x} is typically related to the sound \check{s} ; accordingly they are both specified -Ret. The fortition fronting changes \check{x} to \check{s} (more typically to \check{c}) and the backing of type \check{s} to \check{x} is not without example (e.g., Macedonian \check{c} became \check{k} ; Kondrasov 1962). Nevertheless, \check{x} still retains its darker tonality and is accordingly also specified as Grv.

2.2.3. The ontogeny of position and color.

Here I would like to make a few general remarks about the ontogeny (development within the individual) of position and color vis-à-vis whole systems of consonants. The material which inspired this section is drawn (either explicitly or implicitly) from observations by Jakobson (1968), Velten (1943), Leopold (1947), Edwards (1973), Stampe (1973), and Major (1977).

Of all the implicational relationships concerned with whole consonant systems, the only one that holds for sure is that back (retracted) consonants imply front (advanced) ones (Jakobson 1968:51-58). Thus in the ontogeny of position there is a marked preference for advanced consonants. (Jakobson also lists two sonority implications: fricatives imply stops and affricates imply fricatives.)

For stops, then p and t are naturally acquired first, first as variations of one sound, later as separate phonemes. Often during the variation period p comes to be associated with o , and t with i .

In languages which do not have p , p comes to be replaced by a retracted sound, either k (preserving gravity) or $?$ (preserving non-linguality). In languages with $p + k$, there is generally a period when t is identified with k , or if not identical, t becomes k when there is another velar in the word. In languages with $?$, this may be identified with any buccal stop, but especially k , with which it shares retractedness. Michou Landon at age 2 identified $?$ with p , e.g., button [bʌpʌ] (pers. comm., Bill Landon).

Voicing for stops starts out on a contextual basis, only later becoming nonautomatic. At first, initial stops may be voiced and final ones devoiced.

The first fricative is usually s ; if not, then f --advanced fricatives being preferred. At first processes relate f and s ; later, they become differentiated. Palatal coloring first enters the system through the sound \check{s} , which at first alternates with s (and sometimes with f). At this stage retracted fricatives are usually lost, especially in association with stops (cf. Major 1977).

Voicing for fricatives is contextual at first. Compare the situation for stops.

Nasals are similar in their ontogeny to stops, appearing very early. At first, m and n are related by processes. At this stage, retracted nasals become advanced, e.g., η typically becomes n .

Approximants are at first typically narrowed to fricatives or lost. So, j may become \check{z} or z or be lost. Typically, w is the first approximant to widen, generally having started out as f (v) or having been lost. At this stage, l becomes j or w according to context. Either j or w may replace l initially. The sound \downarrow , as in English, is typically replaced by w or lost. The development of w and j is largely independent. The general loss of approximants is probably due to their being identified with the retracted consonants.

Voicing for nasals and approximants is typically natural (sonorant voicing).

3. Sonority fortitions.

3.1. Narrowing and widening.

There are two contradictory fortitions (paradigmatic processes) which affect stricture: (1) narrowing, and (2) widening. As their names would indicate, the process called narrowing reduces the opening between the articulators, and the process called widening increases it. Typically, narrowing changes approximants to fricatives and fricatives in turn to stops. Widening does the exact opposite; it typically changes stops to fricatives and fricatives in turn to approximants.

Instances of both narrowing and widening follow the guidelines for what constitutes a fortition. These criteria were established above in sec. 1, and include the following principles: (1) Fortitions are segment-optimizing and dissimilative. (2) They are typically context-free, occurring regardless of the specific phonetic environment in which they are found. (3) They are strengthening, i.e., they reinforce some characteristic of the sound in question. (4) They are typically to be associated with hyperarticulate forms of speech. Both narrowing and widening answer to these criteria, as we shall see below.

It is reasonable to conclude, given that both narrowing and widening are fortitions yet have opposite effects, that they must in fact also have opposing teleologies, or reasons for being. In this section and the next I present four pairs of processes, two in this section and two in the next. In sec. 3, the two pairs are narrowing/widening and tensing/laxing. In sec. 4, they are fronting/backing and coloring/bleaching.

The first (left-most members) of these pairs are chromatizations. The second (right-most members) of these pairs are sonorizations. Chromatizations are fortitions which either increase chromaticity (color) or increase its favorability. Chromatizations do so at the expense of sonority. Sonorizations, on the other hand, increase sonority, at the expense of chromaticity.

Narrowing is a chromatization. Typically, the narrower a consonant is, the more likely it is to show the greatest variety of chromaticity, or color. It is not surprising then that, conversely, the wider (more sonorant) consonants should be the first to lose chromaticity. An example is the loss of palatality for every instance of r in Belo-Russian. Cf. Russian pered [p'ɛr'ɪt] 'before' (with palatalized r) with Belo-Russian perad [p'ɛrɔt] 'before' (with non-palatalized r) (examples from Kondrašov 1962:113-4).

Adjunct to the notion of chromaticity is the concept of tonality, or position. Typically, the narrower a consonant is, the more likely it is to show the greatest variety of tonality, or position. English is a good example here. For the voiceless obstruents, English has a p and a corresponding f , a t and a corresponding s (also θ), but k has no corresponding fricative ($*x$).

While narrowing is a chromatization, widening is a sonorization. It increases the opening between the articulators and makes them more vowel-like, i.e., more sonorant. In the following discussion of the two processes, we will find that narrowing is particularly applicable to lighter (more acute) colors (chromaticities); widening, on the other hand, especially affects

darker (more grave) positions (tonalities). So, colors like palatal and labial and consonants with these colors become narrowed. And consonants with darker tonalities, like dorsal and radical, become widened. A specification in terms of the two major sonority features, closure/nonclosure and obstruent/sonorant makes clear the relationships of the four sonority types. These relationships may be schematized as follows:

(3.1)		Obstruent	/	Sonorant
	Closure	STOPS		NASALS
	Nonclosure	FRICATIVES		APPROXIMANTS

Changes usually move horizontally or vertically, but only rarely diagonally. So, while approximants frequently narrow to nasals (upwards) or fricatives (to the left), they seldom make a one-step change to stops (across, diagonally). As we have seen, approximants do make two-step changes to become stops, typically with fricatives as the middle step. Other relationships, not just those of narrowing, can be gleaned from this little chart. For instance, stops rarely widen directly to approximants but rather become nasals or fricatives first.

The process of narrowing for consonants corresponds to Donegan's (1978) vowel process called raising; further, the consonant process widening corresponds to the vowel process lowering. The function of each pair is therefore similar. Narrowing and raising increase chromaticity at the expense of sonority; opposing these two are widening and lowering, which increase sonority at the expense of chromaticity. Thus both narrowing and raising are chromatizations; both widening and lowering are sonorizations.

We now turn to the discussion of these two consonant fortitions, first the chromatization narrowing (in sec. 3.2.1.) and then the sonorization widening (in sec. 3.2.2).

3.1.1.1. Narrowing.

The fortition narrowing may be stated as follows:

$$\left[\begin{array}{l} C \\ n \text{ narrow} \\ ! \text{ lighter} \\ ! \text{ Tns} \\ ! \text{ Gloz} \\ ! / \$ \text{ ____} \end{array} \right] \rightarrow [n + 1 \text{ narrow}].$$

This statement is to be read as follows: "A consonant of n narrowness, especially when it is of a lighter chromaticity or tonality, especially when it is tense, especially when it is glottalized, and especially when it occurs syllable-initially, becomes the corresponding consonant of n + 1 narrowness, i.e., becomes one notch narrower."

In the statement of the process, the exclamation point (!) signifies "especially when." It tells the specific conditions which catalyze (or help bring about) the operation of the process.

In the statement, the notation "n narrow \rightarrow n + 1 narrow" means that an approximant and a fricative become a fricative and a closure, respectively.

It is also possible for a (nonsyllabic) high vowel to become an approximant.

Here are some general examples in which the fortition narrowing operates:

- (3.2) a. Old Armenian \dagger , phonetically probably [u], has become [ɣ] in the modern dialects (Vogt 1974).
- b. In the central and southeastern French of the fifteenth century, intervocalic [r] lost its trill and assibilated to z. In the fifteenth and sixteenth centuries, this pronunciation became popular in Paris and produced changes like: chaire 'chair' > chaise, bericles 'spectacles' > besicles, and Oroir 'place-name' > Ozoir. (Pope 1934:157f.).

In each of the above two cases, an approximant was changed to its corresponding fricative. In (3.2a) the dorsal (velar) approximant u became the dorsal fricative ɣ, and in (3.2b) the (grooved) coronal approximant r became the (grooved) coronal fricative z. In the following example, a fricative is changed to the corresponding stop:

- (3.2) c. Speakers of English typically substitute [k] for [x] when [x] occurs in a word of foreign origin, e.g., Bach [bak] and Khrushchev [kɹuʃtʃɛf].

We now turn to an explanation of the specific "especially-when" cases which serve to catalyze this process.

3.1.1.1. ! Lighter.

In reference to chromaticity, the palatal color j can be said to have a lighter color than the colors u (velar) or α (pharyngeal). The labial color ω is in between these two groups (and rhotacism, or r-coloring, as symbolized by r, could conceivably be in this middle group as well). So, we have the following hierarchy of colors:

- (3.3) a. j ω u α

Leftwards in this hierarchy is lighter; rightwards is darker. Such terms as bright, clear, and slender have been used to indicate the tonal aspects of the term light. For dark, the terms obscure and broad have also been used.

In reference to tonality, a similar hierarchy exists, as follows:

- (3.3) b. † p k ḳ

Here stops have been used to signify positions: † for glottal, † for coronal, p for chilic, k for dorsal, ḳ for radical. There is an exact correspondence between the colors given above in (3.3a) and the positions given here in (3.3b): j corresponds to †, ω to p, u to k, and α to ḳ.

Glottal position has no corresponding color. Again, leftwards in the hierarchy is lighter; rightwards is darker. In features expressing tonality (position), glottal and coronal are acute, and the other (rightmost) three are grave.

Given this information, we see that lighter colors and positions exhibit narrowing most often:

- (3.4) a. Common Slavic palatalized \check{r} has become narrowed to \check{r} [žr] in Czech (Kondrašov 1972:136). The symbol [žr] refers to a ž-sound having r-coloring.
- b. In early Gallo-Roman, j and w were narrowed to dž and g^ω when initial, e.g., Latin jam 'already' became [džə], and Germanic waddju 'gaḡe, pledge' was borrowed as [g^ω adže] (Pope 1934:96f.). Ibero- and Italo-Roman also exhibit the same change.

In the above two examples, approximants with light color (palatality and labiality) have undergone narrowing. The following change shows us how the Gallo-Roman change in (3.4b) must have come about:

- (3.4) c. In Puerto Rican Spanish and other dialects of Spanish, j has become z and w has become y^ω before a vowel. When initial, z in turn becomes dž and y^ω becomes g^ω , e.g., yo [džɔ] 'I' and huevo [g^ω εβɔ] 'egg' (pers. comm., Barry Nobel; Harris 1969:20ff., esp. 25).

The narrowing of j and w thus proceeds this way:

- (3.4) d. $j \rightarrow z \rightarrow dž \rightarrow d\check{z}$
 $w \rightarrow y^\omega \rightarrow g^\omega$

Here is an example of the operation of narrowing on consonants of lighter tonality:

- (3.4) e. In American English, θ (after s and f, or before s) dissimilates to t, e.g., sixth > sixt, fifth > fift, and months > munts. Also in slightly sub-standard speech, f dissimilates to p (after s, or before θ), e.g., sphere > spere (pronounced like spear) and diphthong > dipthong.

This change differs from the change of x to k in that it is dissimilative in nature and not generally applicable.

3.1.1.2. ! Tense.

A full discussion of tenseness and laxness will be given below in sec. 3.2. For our purposes here, it is sufficient to note that voiceless obstruents are usually tense. Tense (and voiceless) fricatives become stops in the following examples.

- (3.5) a. In the English of speakers of various Indian languages (e.g., I have observed it for speakers of Hindi, Bengali, and Gujarati), f and θ become ph and tʰ (strongly aspirated) in careful or emphatic speech.
- b. Prakrit retroflex ṣ became x; this x in turn became kh (pers. comm., David Stampe).
- c. Indo-European *sw became hw in Armenian (by the general substitution of h for s). This hw then became k^h (aspirated), as in Armenian k^hoyr 'sister' (< IE. *swesor-) (Meillet 1936:50). Possibly hw became x before being narrowed to k^h.

3.1.1.3. ! Glottalized.

When a consonant is coarticulated with glottal stop ʔ, i.e., when it is glottalized, that consonant becomes narrowed. I know of no examples in which a fʔ, for example, has become pʔ, but the point is that the languages of the world rarely exhibit glottalized fricatives where there is no corresponding glottalized stop or affricate (as evidenced in Ruhlen 1976). Glottalization then is clearly a catalyst for narrowing. Conversely, aspiration (coarticulation with the glottal fricative h) is a catalyst for widening, as we see in 3.1.2 below. Other glottal phenomena, like voicing, will be mentioned in the discussion on tensing and laxing (see sec. 3.2).

3.1.1.4. ! Initially.

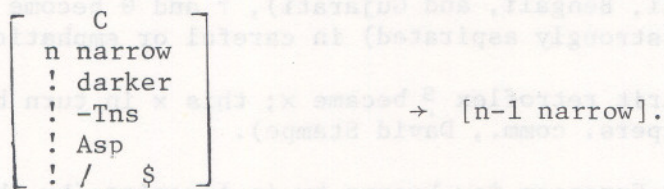
This section refers primarily to consonants occurring initially in the syllable, although most of the examples actually refer to consonants occurring initially in the word. The following examples demonstrate the operation of narrowing on consonants which appear in initial position:

- (3.6) a. The Puerto Rican Spanish example in (3.4.c) above shows narrowing of j and w in both syllable-initial and word-initial position. Initially in syllables, j and w become ʒ and ɣ^w respectively, e.g., calle [kaʒe] 'street', mayo [maʒo] 'May', and Chihuahua [tʃiɣ^waɣ^wa] 'state of Mexico'. Initially in words, j > ʒ (> dʒ) > dʒ and w > ɣ^w > g^w, e.g., yerba [dʒerβa] 'grass' and huerta [g^werta] 'garden' (Harris 1969:20ff.).
- b. The change of w to v is common, probably w > ω > v > β > v, where the change of v to β is a narrowing. It occurs in Sanskrit, Slavic, Romance, Germanic except for English, and various Polynesian languages, and child language (see sec. 1.2.2).
- c. In Kekchi, a Mayan language, j and w have become tʃ and kw, respectively, in initial position (Anttila 1972:69).

This completes the discussion of narrowing.

3.1.2. Widening.

Widening is a fortition and sonorization that can be stated as follows:



This may be read as "A consonant is widened one notch (becomes the corresponding consonant of the next wider stricture), especially when it is of a darker chromaticity or tonality, especially when it is lax (nontense), especially when it is aspirated, and especially when it occurs syllable-finally." Again, the exclamation point (!) signifies "especially when."

In the statement, the notation "n narrow → n - 1 narrow" means that closures would become fricatives, and fricatives would become approximants.

Here are some general examples of the operation of this process:

- (3.7) a. In Ukrainian and Slovak, two languages of the Slavic family, g widened to γ, which in turn gave ħ (voiced). Compare Russian golova 'head', with initial [g], and its cognates in Ukrainian and Slovak, golova and hlava, respectively, both with initial [ħ]. (Examples from Konrašov 1962.)
- b. In morphophonemic alterations, Finnish p + k voiced and widened to β ð γ. β then dentalized to v. γ then fronted to ʋ before a labial vowel, to j before a palatal vowel (after r or l), or was lost. ð appears as d in the standard language. (pers. comm., Ilse Lehiste; cf. Anttila 1972:219-222.) Examples:

nom. sg.	gen. sg.	gloss
papu	pavu-n	'bean'
kato	kado-n	'loss'
suku	suvu-n	'family'
järki	järje-n	'intelligence'
joki	joe-n	'river'

- c. The Russian adjective endings -ego and -ogo were also affected by widening. Here g became γ (and later fronted to ʋ under the influence of the labial vowel following). E.g., Russian xorošogo [xʌ'ʊʂəvə] 'good (gen. sg.)'.

3.1.2.1. ! Darker.

In accordance with the hierarchy given in sec. 3.1.1.1. above, narrowing affects lighter chromaticities and tonalities and now widening affects darker ones. I repeat the hierarchy here, using colors to symbolize themselves and stops to symbolize positions:

- (3.8)
- | | | | | |
|---|---|---|---|----|
| | j | ω | ɥ | α |
| ? | t | p | k | k̥ |

For colors, leftwards indicates lighter chromaticity and rightwards, darker chromaticity. For positions (the stops), leftwards indicates lighter tonality and rightwards, darker tonality. Widening works most often on positions of darker tonality, and on darker colors.

The following examples illustrate the "darker" condition:

- (3.9) a. The absence of radical (pharyngeal) closures (like $k̠$ $g̠$ $ŋ̠$) in the languages of the world is a good indication that the darkest closures must widen. (It is however, possible for some speakers to produce a voiced, imploded radical stop.)
- b. After i; Old English x t became Modern English jt (spelled ght, as in night), while ft and st remained unaffected. (Cf. Wright 1928:21.)
- c. Latin kt between vowels became early Portuguese ft or xt, while pt and tt fell together as the geminate tt (p and t both being lighter than k). Latin aktum 'done' became afto (whence Portuguese auto, as in auto da fe); Latin noktem 'night' became noخته (whence Portuguese noite); Latin kattum 'cat' became katto (whence Portuguese gato). Latin ps ks similarly became fs xs. (Grammont 1933:203ff.; Williams 1962:84f.)

We can thus see the widening hierarchy affects dorsals first, then chilies, and then coronals. Cf. the tensing hierarchy for the Old High German example given below in sec. 3.2.1.1, which works in the reverse order.

3.1.2.2. ! Lax.

Lax (voiced) stops become fricatives in the examples which follow. A full discussion of the tense/lax distinction is given in sec. 3.2 below.

- (3.10) a. Spanish b d g became β δ γ (allophonically), typically in allegretto speech. Harris (1969:38) states that the process applies when these environments occur:

- (1) V ___
- (2) Fric ___
- (3) App ___ except for ld
- (4) ___ App
- (5) ___ N

(All this happens while p t k remain unaffected.
Examples for b:

- (1) V__ (#): haba [aβa]; club [kluβ]
- (2) Fric__ : adverso [aδβersɔ]
- (3) App__ : calvo [ka|βɔ]
- (4) ___App: habla [aβla]
- (5) ___N: submarino [suβmarino].

- (3.10) b. Ancient Greek b d g have become Modern Greek v ð γ, while p t k remain unaffected. (At the time that this change occurred, Ancient Greek ph th kh (aspirated) had already become f θ x.)

The Ancient Greek aspirates, although tense, were also aspirated and thus fall under a slightly different teleology. (See example 3.12d) in the next section.) It is necessary, and sufficient, to realize for our purposes here that because the Spanish and Greek voiced stops were lax, they underwent widening, the voiceless (tense) stops being unaffected.

3.1.2.3. ! Aspirated.

Aspirated stops become fricatives under the provisions of the process of widening, as in the following examples:

- (3.11) a. For speakers of Indian languages who have served as my informants, the strong aspirates ph th kh become f θ x in casual speech, e.g., the Gujarati word phɔdz 'army' (as cited by Ladefoged 1971:13) is pronounced [fɔdʒ] by JL, a Gujarati speaker.

I might point out that BNP, a Bengali student of mine, could pronounce fricatives in English (voiceless ones, at least) only in fast or careless speech, i.e., when he wasn't paying attention. The full aspirates for him were connected with emphatic speech.

- (3.11) b. Early Germanic p^h t^h k^h k^{wh} from Indo-European p t k k^w, widened to f θ x x^w. (x x^w later became h h^w.) (cf. Prokosch 1939:53; Priebisch and Collinson 1934:67.)

- c. Old High German p^h t^h k^h, from earlier p t k (Indo-European b d g), underwent two different but related changes, both widenings. Initially p^h t^h k^h became pf ts kx (kx later reverted to k^h). Medially they became ff ss xx. These geminates later simplified to Modern German f s x, spelled ff, ss, and ch. As for chronology and dialectal variation, t^h changed first and most widely, p^h next, and k^h last and least widely (Prokosch 1939:54; Priebisch and Collinson 1934:111ff.).

- d. Ancient Greek ph th kh (aspirates), from Indo-European bh dh gh (and variously from g^{wh}), have become widened to f θ x in the modern language (pers. comm., Malikouti Drachman).

3.1.2.4. ! Finally.

This section of particulars refers primarily to widening changes occurring in syllable-final position, but also refers secondarily to those in word-final position. All three examples, in fact, refer to word-final phenomena.

- (3.12) a. Gallo-Roman \dagger d k widened to θ $\check{\delta}$ x in final position after a vowel, i.e., /V___# (Pope 1934:142).
- b. In Danish, \dagger d become widened (and in the case of \dagger also voiced) to $\check{\delta}$ in final position after a vowel; g becomes widened to γ in final position after a vowel (Spore 1965:41-43). Examples: skibet [sqi $\check{\delta}$ b $\check{\delta}$] 'the ship', tog [to γ] (also [to γ w] 'was taking'. (\dagger after d = [$\check{\delta}$] simply voices to d, e.g. badet [ba $\check{\delta}$ $\check{\delta}$ ed] 'the bath'.) (I am assuming here that what Spore writes as voiced lax b g are really voiceless lax b $\check{\delta}$ g.
- c. Spanish b d widen to β $\check{\delta}$ finally after a vowel, e.g., club [klu β] 'nightclub', usted [uste $\check{\delta}$] 'you (polite)' (my observation based on Harris 1969). Spanish b d g also widen to β $\check{\delta}$ γ syllable-finally before resonants (nasals and approximants), e.g., for b: submarino [su β marino] 'submarine' and habla [a β la] 'speaks'. (Cf. example (3.10a) above.)

As for other examples of syllable-final position, American English flapping occurs in that position as a rapid-speech pronunciation of the word hothouse [h'o \check{r} h,ə \check{s}]. But this is probably due to a lenition. Compare the Tagalog alternation of d and r. r becomes d initially: dusaryo 'rosary' from Spanish rosario. d becomes r between two vowels: tukod or tukuran, both meaning 'prop' (Tablan and Mallari 1961:xi). Also compare this Hindi example:

- (3.12) d. Hindi retroflexed \check{d} $\check{d}h$ \check{n} widen to \check{r} $\check{r}h$ \check{r} in syllable-final position, as in reduplications of VC syllables (Davison 1971).

This completes the discussion of widening.

3.2. Tensing and laxing.

Tensing is a chromaticization, i.e., it favors chromaticity at the expense of sonority. Laxing, on the other hand, is a sonorization: it increases sonority while simultaneously decreasing the favorability of chromaticity. So, like the processes of narrowing and widening just discussed, tensing and laxing respond to opposing teleologies. But it should be pointed out that, unlike narrowing and widening, the processes of tensing and laxing do not change stricture. Instead, they merely serve to change the "orientation," as it were, of a sound to its stricture, tensing (within a given stricture) favoring chromaticity and laxing (within a given stricture) favoring sonority.

To accomplish this orientation, or perhaps to act as its signature, tensing and laxing are closely related to other phenomena which act to serve notice that these processes have in fact applied. These "signatures" fall into three categories: (1) glottal phenomena, like voicing and aspiration, (2) timing phenomena, i.e., longness or shortness, and (3) associations with accentual phenomena, e.g., tensing is connected with position before the syllable peak and laxing is connected with position before the measure peak.

Tensing, then, is related to (and often expresses itself as): (1) voicelessness, glottalization, and aspiration (even though these last two give separate results for narrowing and widening), and (2) length (longness). It is also associated with (3) initial position in a stressed syllable (position before the syllabic peak).

Laxing, however, is related to (often expresses itself as): (1) voicedness, and (2) lack of length (shortness). It is associated in turn with (3) unstressed syllables occurring before stressed ones (before the measure peak), generally in syllable-final position.

It is necessary at this point to question the meaning of the phrase "often expresses itself as." Voiceless consonants are typically tense; aspirated, glottalized, and long (geminate) consonants are, as it were, "overtense." Compare Jakobson and Halle's (1956) provision for their feature tense/lax to cover the phenomenon of aspiration. Voiced consonants are typically lax; short consonants are also typically lax.

But "overtenseness" is not real and the basic binary distinction of tense/lax holds. Let us assume that there are but three combinations of tense/lax and voiceless/voiced. These three are then possible: tense voiceless, lax voiceless, and lax voiced. The fourth, tense voiced, is impossible (for obstruents, at least). Then given this one limitation, all other features are independent. Tense voiceless t can be aspirated, glottalized, long, or short. The same goes for lax voiceless $ɸ$ and lax voiced d . (Incidentally, we must distinguish $ʔd$ and $dʔ$, the first simply glottalized and the second glottalized and also imploded, i.e., having the feature suction.) Further, glottalized $tʔ$ (tense) counts as an "ejective"; glottalized $ɸʔ$ (lax), as an "injective." As a result, we must distinguish between "typical" cases, e.g., t^h being tense, and "independent" cases, e.g., d^h being lax.

In this same general vein, all these phenomena--voicing, devoicing, aspiration, glottalization, lengthening, and shortening--may at one time or another be the result of lenitions. For instance, if we find a proto-form $[+t.t.c]$ and a reflex form $[+t.d.c]$, we would be forced by the weight of the argument advanced here to posit an intermediate form with $ɸ$: $[+t.ɸ.c]$. The significance of this is that the change of t to $ɸ$ is a fortition, merely catalyzed by the more sonorant, voiced environment in which it occurs. But the change of $ɸ$ to d is a lenition, an assimilation to voicing. If it were not otherwise, we would be forced to conclude that laxing, one of the fortitions to be discussed in this section (along with tensing), was in reality a lenition.

But both tensing and laxing are fortitions, and we now turn to a discussion of them, starting with tensing.

3.2.1. Tensing.

The fortition tensing may be stated as follows:

$\left[\begin{array}{l} C \\ ! \text{ narrower} \\ ! \text{ Long} \\ ! \text{ -Voi} \\ ! / \text{ ____ Syl Peak} \end{array} \right]$	→ [Tns].
--	----------

This statement should be read as "A consonant becomes tense, especially when it is narrower, especially when it is long, especially when it is voiceless, and especially when it occurs before a syllable peak (i.e., before a stressed vowel in the same syllable)." Again, the exclamation point (!) signifies "especially when."

Here are some general examples of tensing:

- (3.13) a. Early Germanic b d g g^w (from Indo-European) became p t k k^w, even in voiced environments (Priebsch and Collinson 1934:64ff.) Compare Latin pedem 'foot', Gothic fōtus, and English foot (d to t); Latin jugum 'yoke' and English yoke (g to k). Examples of g^w to k^w are rare; b to p is relatively unattested.
- b. The same Indo-European series, b d g g^w, became Armenian p t ts k (Meillet 1936:28-29, chart on p. 37). The k results from delabialization of g^w (and subsequent tensing); the ts involves the palatalization of g: g > ġ > dź > dz > ts (the last step representing tensing).

We turn now to the specific contexts in which tensing is catalyzed.

3.2.1.1.1. ! Narrower.

- (3.14) a. Old High German b d g became p t k while β ð γ remained unchanged. d underwent tensing first and most widely, then b, and then g last and least widely (Priebsch and Collinson 1934:115ff.; cf. Prokosch 1939:54).

Many of the b to p and especially the g to k changes never made it into the dialect upon which modern standard German is based. But the d to t change was common; compare English dead, deer, door, and do with German tot, Tier 'animal', Tür, and tun.

- (3.14) b. Languages of the following families typically have no voiced obstruents: Sino-Tibetan, Salishan, and to a lesser extent Penutian. Salishan languages like Bella Coola, Clallam, Columbian, Comox, Cowlitz, and Squamish have extensive systems of obstruents, yet all are voiceless (my observation based on Ruhlen 1976).

We can conclude from the above that narrower consonants, e.g., obstruents, devoice readily (and become tense).

3.2.1.2. ! Long.

Here I will just make the general observation that long (geminate) consonants which are already voiceless do not voice (remain tense) even in cases where their nongeminate counterparts do, e.g., Old English f s voiced to v z between vowels while geminate ff ss remained tense, and early Latin s voiced to z between vowels and became rhotacized to r but ss remained unchanged, later simplifying to s after long vowels and diphthongs, e.g., causa 'cause' from earlier caussa (pers. comm., David Griffin).

3.2.1.3. ! Voiceless.

This can be demonstrated by lenitive final devoicing of stops and their subsequent fortitive tensing, e.g., American English dialectal variations of final b d g, first leniting in final position to b̥ d̥ g̥. The fortition tensing then changes these to p t k, e.g., cold [k^ho:l̥t̥], world [wɜ:l̥t̥], dog [dɔ:k̥], and big [bɪ:k̥]--all preserving the long or diphthongized form of the vowel typically found before voiced consonants. It is important to remember that all of the tensing changes involving voicing that are described here start with devoicing, either due to a fortition or a lenition, and then undergo tensing.

David Stampe (pers. comm.) reports that the type of devoicing referred to in the preceding paragraph is also typical of the English of speakers from Australia and New Zealand.

3.2.1.4. ! Before a syllable peak.

In the following two examples, aspiration and glottalization act as a way of "tensing" consonants which are already tense.

- (3.15) a. In German, English, and other Germanic languages, p t k (and ċ) become aspirated before a stressed vowel.
- b. In French and Spanish p t k (č) optionally become glottalized (slightly) before a stressed vowel. This is typical of emphatic or formal speech. (This statement is based on my own observations of native speakers of these languages.)

This completes the discussion of tensing.

3.2.2. Laxing.

The fortition laxing may be stated as follows:

$$\left[\begin{array}{l} C \\ ! \text{ wider} \\ ! \text{ Short} \\ ! \text{ /wider} \\ ! \text{ /___ Meas Peak} \end{array} \right] \rightarrow [-Tns].$$

This statement may be read as: "A consonant becomes lax, especially when it is wider, especially when it is short, especially when it occurs in the neighborhood of a wider sound (e.g., a vowel), and especially when it occurs before a measure peak (i.e., in the syllable before a stressed syllable)."

Here are some general examples of laxing, which also show lenitive voicing:

- (3.16) a. In Joan Velten's speech, p + s become b d z before a vowel, e.g., peach [butʃ], toe [du], and sauce [zɔs] (Velten 1943).
- b. A number of northern Australian languages, such as Dyirbal, Gudandji, Ngarndji, and Yanyula, have a series of voiced stops (constituting the only obstruents) with no voiceless series (Ruhlen 1976).

In both these cases, we can assume that the stops were laxed to lax voiceless stops, and then these lax stops assimilated to the voicing of the surrounding sounds. This last case establishes that laxing is in fact a fortition, since otherwise it could not affect a whole system in a context-free way.

- (3.16) c. In Cherokee, t k and d g contrast only before the vowel a; before the other vowels, i e ə o u, only d and g occur. (Walker, 1975).

This last is again an example of laxing followed by voicing. We turn now to specific catalyzing elements.

3.2.2.1. ! Wider.

- (3.17) a. Resonants (nasals and approximants), which are typically voiced, may generally be assumed to be lax and to have tense counterparts (in the strict sense of the word) only allophonically, e.g., the n of French words like bonne [bɔn⁺] 'good' (fem.) is pronounced rather tensely in the standard dialect (my own observation).
- b. In the Kentish dialect of late Old English, f s θ become laxed and voiced to v z ð initially, the corresponding stops p t (k) remaining unchanged in this same position (Wright 1928:107f.). Examples (from Wright): vrend 'friend', zinne 'sin', and ðing 'thing'.

This last case can be connected to the occurrence of two series of fricatives (voiceless and voiced) in languages which have only one (voiceless) series of stops, e.g., Modern Greek. This establishes the pattern but is not meant to preclude languages like Spanish that have both series of stops but one (phonemic) series of fricatives (voiceless).

3.2.2.2. ! Short.

This is the converse of the examples given for ! Long in sec. 3.2.1.2. for the process tensing. There it was pointed out that single (nongeminate) consonants lax and become voiced in the same environments where long (geminate) consonants remain tense, e.g., the Old English voicing of

f s to v z between vowels, geminate ff ss remaining tense in the same position. Further, short (i.e., flapped) resonants such as ñ or ʀ devoice only lenitively, e.g., from my own observations of the Spanish of Mexico City, speakers optionally devoice the last half of final ʀ to produce an apical fricative effect: señor [sɛnɔʝɾ̥s̺]. In all other (nonlenitive) cases, laxing allows the voicing of flapped sounds, e.g., the voiceless flapped ʀ̥, resulting from the American English flapping of t, is generally voiced to ʀ.

3.2.2.3. ! Around wider sounds.

- (3.18) a. In late Latin (Vulgar Latin) p t k laxed and voiced to b d g between vowels (Pope 1934:137).
- b. Danish p t k become lax b̥ d̥ g̥ in combination with a fricative even when in final position, e.g., spille [sɸelə] 'to play,' lapset [lɔpsɛʔ] 'elegant,' and gisp [gisɸ] 'groan' (Spore 1965: 35-39). Similar to sp, etc., st and sk are realized as [sɸ] and [gɸ], etc. (I am assuming that what Spore writes as b d g are b̥ d̥ g̥ in these positions.)
- c. In Pre-Welsh (Brythonic Celtic) p t k laxed to b d g between vowels (or after a vowel and before n l r plus a vowel). This was incidentally accompanied by a previous widening of b d g m to v ɸ γ ʋ̃ (this last sound remaining in some dialects of Breton). (Morris Jones 1913:161-67).

3.2.2.4. ! Before a measure peak.

This environment refers specifically to consonants appearing in a (usually initial) syllable preceding a stressed syllable:

- (3.19) a. My childhood pronunciation of potato was [ɸət^h'eɪɾə], or even [bɛ...], this pronunciation recurring in my stepson Arthur's speech at age 6 (cf. my own lazy speech pronunciations, e.g., tomorrow [dm'ɑʊ]).
- b. ks is laxed and voiced to gz before stress, e.g., e[gz]ist, e[gz]ert, e[gz]haust, etc.; otherwise ks remains: e[ks]it, a[ks]is, Me[ks]ico, etc.

Verner's Law changed voiceless f θ s x to voiced v ɸ z γ after an unstressed vowel (cf. Prokosch 1939:60). Generally, the vowel following was stressed. As a result, Verner's Law is probably related to the two cases above, particularly the second.

This completes the discussion of laxing.

4. Position and color fortitions.

4.1. Fronting and backing.

The two fortitions which affect position are fronting and backing. The first of these, fronting, is a chromatization; the second, backing, is a sonorization.

As a chromatization, the fronting process favors chromaticity or increases it, often at the expense of sonority. Like the chromatization process of narrowing, fronting applies to tense sounds. It typically affects narrower sounds and sounds bearing a color, or colors themselves, especially when the color involved is one of the fronter colors, either labiality or palatality. Fronting also affects dental sounds; specifically, it fronts coronal sounds like $\theta \delta$ to chilic $f v$.

As a sonorization, backing increases sonority, often at the expense of chromaticity. It most typically applies to achromatic (non-colored) consonants, lax consonants (cf. the sonorization process of widening), and wider consonants, those consonants with some degree of sonority already. It is also very common in word-final position.

The statement of and evidence for each of these processes will now be presented, starting with the discussion of fronting, and ending with backing.

4.1.1. Fronting.

The fortition fronting may be stated as follows:

$$\left[\begin{array}{l} C \\ n \text{ front} \\ ! \text{ Color} \\ ! \text{ Tns} \\ ! \text{ narrower} \\ ! \text{ Dent} \\ ! \text{ /Max. Diff.} \end{array} \right] \rightarrow [n + 1 \text{ front}]$$

This statement may be read as: "A consonant of a given position (n front) becomes the next position forward (n + 1 front), especially when it bears a color, especially when it is tense, especially when it is narrower, especially when it is dental, and in cases involving the principle of maximal differentiation." Again, the exclamation point (!) is to be read "especially when" (see sec. 3.1.1.).

A typical change which fronts dorsals to coronals occurs in the speech of children. Specifically, $t d$ are substituted for $k g$, e.g., my cousin Lori at about age 4 pronounced her last name Kisling as "Tisling." Such fronting occurs not only in child language but also as part of the natural history of languages of the world, to which the examples which follow testify.

We now turn to a discussion of the specific environments which catalyze (help bring about) instances of fronting.

4.1.1.1. ! Color.

Frontings involving color always front to the position associated with that color. So, labial(ized) consonants become chilic, and palatal(ized) consonants become coronal. Here are some examples of the fronting of labial consonants:

- (4.1) a. Latin $k^{(w)} g^{(w)}$ (qu, gu) has become $p b$ in Rumanian, e.g., Latin aqua > Rum. apă 'water', Latin quattuor > Rum. patru 'four', and Latin lingua > Rum. limbă 'language' (Nandris 1963:263).
- b. Indo-European $*k^{(w)} *g^{(w)}$ have become $p b$ in the Brythonic branch of Celtic (consisting of Welsh, Cornish, and Breton) and $g^{(w)}$ has become b in Irish. Cf. Latin quinque, Old Irish cóic, Old Welsh pimp 'five'; Latin vivus (Indo-European $*g^{(w)}iwo-$), Irish beo, Welsh byw 'living' (Lewis and Pedersen 1937: 43f., 34f.).

There is also the Latin change of $*g^{(w)h}$ (Indo-European) apparently to $*x^{(w)}$, fronting then to f , e.g., Latin formus 'warm' (from $*g^{(w)}hermos$, cf. Greek thermos 'warm'; cf. Pokorny 1959). There is also the fronting of $t^{(w)}$ to p in the Yiddish word epas 'something' (cf. standard German etwas). This is apparently an isolated example.

Palatal consonants similarly front to the position corresponding to palatal coloring, i.e., coronal. Here are some examples:

- (4.1) c. Latin $k' g'$ (from $k g$ before a palatal vowel) fronted to $tš dž$ (= $č ĵ$) in virtually all the dialects (probably through an intermediate stage of $t' d'$). Italian and Rumanian still preserve this pronunciation, e.g., Latin pacem 'peace', Italian and Rumanian pace with $č$; Latin legit 'reads', Ital. and Rum. lege with $ĵ$ (cf. Nandris 1963).
- d. Common Slavic $k' g' x'$ (from $k g x$ before original palatal vowels or j) became $tš dž š$ in the dialects ($dž$ most usually simplifying to $ž$). In Russian, this sound change produced the following alternations: oko 'eye', oči 'eyes'; sluga 'tear', služat' 'to cry'; uxo 'ear', uši 'ears' (Kondrašov 1962:39-41).

Similar instances of color fronting occur in Chinese, Old English, and Tumbuka, a Bantu language.

The backing of a sound to match the position of a color is, however, usually assimilative in nature, i.e., a lenition. For instance, a change in eastern and southern Slavic, similar to the one described above, shifted $čiljics$ plus j (i.e., $pj bj vj mj$) to $čiljics$ plus l plus j ($plj blj vlj mlj$). The wide j has assimilated to the narrowness of the $čiljics$ involved by becoming lj [l']. One example is Russian zemlja 'land', cf. Polish ziemia (examples from Kondrašov 1962).

Also, Nandris (1973:111-12, 240ff.) notes cases of dialects of Rumanian where palatalized f' and b' have become palatals of various types, some showing intermediate stages of assimilation, some showing complete stages, e.g., f' becomes intermediate $f'k$, $f'h$ and complete $š$, $ś$, $č$, h' ; and b' becomes intermediate bd' $bdž$ $b'g$ and complete d' $dž$

ġ. The Old French change of $\acute{p} \acute{b}$ to $pt\check{s} \acute{b}d\check{z}$ and then to $(t)\check{s} (d)\check{z}$ (and of $\acute{m}\acute{v}$ to $n(d)\check{z}$) appears everywhere complete in the modern language, e.g., sache 'may know (subjunctive)' from Latin sapiam, and rouge 'red' from Latin rubeum (Pope 1934:129).

4.1.1.2. ! Tense.

I have but one example here, but it is, I believe, a rather convincing one:

- (4.2) A two-year old Russian child fronted k (tense) to t but g (lax) was merely devoiced to k , e.g., ruki 'hands' became $[j'ut'i]$ but knigi 'books' and ogon 'fire' became $[h'iki]$ and $[ak'a\check{n}]$, respectively (Jakobson 1968:15, based on observations by Alexandrov).

Jakobson goes on to point out (ibid.) that the boy probably first fronted both k and g , only later limiting the fronting process to the tense counterpart.

4.1.1.3. ! Narrower.

The child language frontings of $k \ g$ in the speech of Hildegard Leopold and Joan Velten, mentioned above in the opening of this section (4.1.1), are examples here. Other examples include:

- (4.3) a. Proto-Polynesian $*\eta$ was fronted in all positions to n in Hawaiian and dialects of Marquesan and Maori (Biggs 1971). Accordingly, North Auckland Maori η corresponds to Bay of Plenty Maori n .

An interesting case of fronting which preserves the gravity (Grv) of the consonants fronted occurs in Rumanian:

- (4.3) b. Latin $kt \ ks \ gn$ have become Rumanian $pt \ ps \ mn$ (Nandris 1963:108, 117). An example for kt is Rum. opt from Latin octo 'eight'; for ks , Rum. coapsă from Latin coxa $[k\check{o}ks\check{o}]$ 'thigh'. Vulgar Latin must have nasalized gn to ηn ; the η then fronted to m , e.g., Rum. lemn from Latin lignu(m) 'wood'.

4.1.1.4. ! Dental.

In these examples the dental coronal fricatives $\theta \ \delta$ become labio-dental (chilic) $f \ v$. I know of no examples in which dental $t \ d$ become fronted.

- (4.4) a. In the English of London, standard $\theta \ \delta$ are replaced by $f \ v$, e.g., three $[fre\check{i}]$, father $[fa:v\check{e}]$ (Jones 1914:29).
- b. Greek θ was often replaced in Russian by f in names borrowed from Greek, e.g., Russian Fëdor, cf.

Theodore in English; Russian Marfa, cf. Martha (cf. Passy 1890:151).

- c. In the development of Indo-European into Latin, IE. *dh must have been pre-Latin θ. This θ fronted to f in initial position (falling together with f developed from IE. *bh). (Similarly, IE. *gh became x in initial position in Latin, eventually backing to h.) For θ to f, compare English deer and Latin fera 'wild beast' both reflecting IE. *dh (cf. Passy 1890:151).

4.1.1.5. Fronting changes due to maximal differentiation.

There is one type of fronting that affects dorsals like k especially when there is a backer dorsal like q in the same sound system. For instance, in Greenlandic Eskimo, a k/q distinction exists in which k has become slightly fronted to k<; somewhere in between k and k̂. The result sounds "slightly palatal" (pers. comm., Jerrold Sadock). This change is a result of the principle of maximal differentiation.

One might also speculate that maximal differentiation was the key factor in the fronting of Indo-European *k *g *gh to k̂ ĝ ħ in the satem dialects. This was apparently done to widen the contrast for the distinctions between plain dorsal and labial dorsal, i.e., k/k^ω, g/g^ω, gh/g^ωh.

This completes the discussion of fronting.

4.1.1.2. Backing.

The process backing may be stated this way:

$$\left[\begin{array}{l} C \\ n \text{ front} \\ ! -\text{Color} \\ ! -\text{Tns} \\ ! \text{ wider} \\ ! / ___\# \\ ! / \text{Max. Diff.} \end{array} \right] \rightarrow [n - 1 \text{ front}].$$

This statement may be read as: "A consonant of a given frontness gets backed one position, i.e., one notch farther back, especially when the consonant bears no color, especially when it is lax, especially when it is wider (i.e., more sonorant), especially when it occurs (word-) finally, and in cases involving the principle of maximal differentiation." Again the exclamation point (!) is to be read as "especially when."

Here are some general examples of backing:

- (4.5) a. The Polynesian languages exhibit a number of context-free instances of backing: (1) k becomes ? in Hawaiian, Samoan, Tahitian, Rurutu, South Marquesan, etc. (2) ŋ becomes ? in Tahitian and Rurutu. (3) t becomes k in Hawaiian and Colloquial Samoan (Biggs 1971). In each case, the

reflex corresponds to the prototype in some dialect, which is assumed to be more original, e.g., for (3) $t > k$, compare Tahitian tabu and Hawaiian kapu, both meaning 'taboo'.

Here is an example from child language:

- (4.5) b. Elizabeth Stampe at age 1:6 backed all chilics to coronals, i.e., $p b f m$ became $t d s n$, as in powder [tada], Baba (from baby, a name for herself) [dada], fox and fish [sa] and [s+s], and mama [nana] (pers. comm., David Stampe).

We now turn to specific instances of catalyzing environments.

4.1.2.1. ! Achromatic (noncolored).

All the examples of backing illustrate this principle, except for the cases involving maximal differentiation, which all affect color in what appears to be a special catalyzing environment. Particularly illustrative of this achromatic environment are examples like the Polynesian and child language data given above. For the first two Polynesian changes, dorsal $k \eta$ are backed to $ʔ$, and for Elizabeth chilics become coronals. Glottals and coronals are connected with higher tonalities and so are more sonorant; backing, as a sonorization, thus moves toward sonority. (Cf. sec. 4.1.1.1.)

4.1.2.2. ! Lax and ! Wider.

These two cases will be considered together since they constitute, for backing, similar cases; i.e., wider (more sonorant) consonants, as long as they bear no color, are typically lax.

- (4.6) a. Ukrainian and Slovak g first widens to γ , which in turn backs to h (the voiced glottal fricative) while x remains unaffected (Kondrašov 1962).

The above example and those to follow have occurred for lax consonants; those which follow especially concern wider consonants.

- (4.6) b. American English l becomes ω and even α in the idiolectal speech of individuals (cf. sec. 2.2.1 above).
- c. Old French trilled apical r became (trilled) uvular R and remains so in the modern language (Pope 1934: 156ff.)
- d. In some forms of Puerto Rican Spanish trilled apical \bar{r} (long) becomes (trilled) uvular R , then devoicing and narrowing to the uvular fricative x e.g., ron rico [$x'ɔŋ x'ikɔ$] 'rich rum (a brand name)' (pers. comm., Barry Nobel).

Here is one more example, whose teleology is not clear to me:

- (4.6) e. Dutch *f* becomes *x*, apparently a dissimilative backing after a labial vowel in what may be an isolated example, namely: Dutch *lucht* [lœxt] 'air'; cf. German *Luft*, same meaning (pers. comm., David Stampe).

4.1.2.3. ! Finally.

All of the following examples are instances of backing which occur in word-final position:

- (4.7) a. Puerto Rican Spanish *n* backs to *ŋ* word-finally, e.g., *ron* [xɔŋ] 'rum' (pers. comm., Barry Nobel). I have also noticed this change in the Spanish and English of my students from Venezuela, Colombia, and Ecuador.

- b. Similarly, Old French *n* backed to *ŋ* in final position before being lost (cf. Pope 1934:169-70).

The above examples concerned the backing of *n*; the next two, the backing of *m*:

- c. Pre-Greek *m* became Ancient Greek *n* in word-final position; cf. Latin *kentum* and Greek *-katon* 'hundred'.

- d. Ancient Chinese *m* became northern (Mandarin) Chinese *n* in word-final position, e.g., An Chin. *[-s^həm] and Mandarin [-ʂən], cf. Cantonese [ʰsəm], all meaning 'deep' (lines represent tones) (data from Forrest 1948:179 and Appendix II).

In the following example, both stops and nasals are affected:

- e. South Vietnamese (Hanoi) *t* *n* back to *k* *ŋ* word-finally except after tense palatal vowels, i.e., [i] and [e]. Conversely, *k* fronts to *t* word-finally after [i] or [e] (using data from Nguyen 1970: 236-240; Thompson 1965:94-103).

One last example of backing follows, which due to its nature is probably an example of stops becoming lax in word-final position:

- f. Ancient Chinese *p* *t* *k* in final position (still extant in some dialects) all become backed to ? (which remains in some dialects). In northern (Mandarin) Chinese, this ? is then lost (cf. Forrest 1948).

Some examples from Forrest (1948: Appendix II, 307ff.) include:

Anc. Chin.	Canton	Suchow	Peking	Gloss
-jɛp	ɟi:p	-jɛ?	\jɛ	'leaf'
_bjɛt	_pi:t	-bi?	/pjɛ	'different'
-tɛk	-tak	-tɛ?	/tɛ	'get'

See also 5.1.3.1 below on aspiration and glottalization phenomena.

4.1.2.4. Backing changes due to maximal differentiation.

High is the feature connected with high vowels like i ɨ u and high consonants like ʃ ʂ x, etc. Consonants which are high, sometimes in spite of their color and sometimes because of their retraction, occasionally become backed. This is in some cases due to the principle of maximal differentiation. Here is an example:

- (4.8) a. In dialects of American English, palatal \acute{n} gets backed to $\acute{\eta}$ in such words as onion [' $\lambda\acute{\eta}$ ən], bunion [b' $\lambda\acute{\eta}$ ən], apparently after achromatic [λ]. Cf. canyon with [æ], which is never *[k^hæ $\acute{\eta}$ ən].

In this example and some of those which follow, this backing proceeds in spite of the color present, which would normally act as a fronting agent, e.g., $\acute{\eta}$ often fronts to \acute{n} .

In the following example, the change occurs because of the retractedness of the sound:

- b. In Common Slavic retroflex ʂ produced by the ruki rule (which changes s to ʂ after r, u, k, i) is subsequently backed to x (my own conclusion based on my studies of the history of Slavic).

In the next two examples, both palatality and retroflexion (probably involving velarity) are involved:

- c. Old Spanish ʃ backed to (became retroflexed to) *ʂ which in turn backs to x, remaining in the modern language (cf. Harris 1969:193-98).
- d. In the standard (Moscow) dialect of Russian, ʃ' ʂ' get backed to ʂ̣ ʂ̣̣, but č' remains unaffected, e.g., sibko [ʃ'ipkə] 'quickly', židko [ʂ'itkə] 'weakly', but čisto [č'istə] 'clearly' (my own observations; cf. Kondrašov 1962).

In each of these two changes, it is important to note the sound system in which they occur. The reason for this is that these are both examples, I believe, due to the principle of maximal differentiation. In Old Spanish, ʃ (from original ʃ̣, and ʃ̣ from ʂ̣) was opposed to apical ʂ̣ (from original ʂ̣, and ʂ̣ from ʂ̣̣). Harris (1969:192) remarks on the palatallike sound of apical ʂ̣. The need for a more optimal contrast is apparent; this is accomplished by backing the ʃ̣ to ʂ̣̣ (in spite of its palatality, normally a fronting agent). Then the retractedness of ʂ̣̣ backs it still farther to x (x̣ in Castillian).

These changes may be expressed by the following two statements:

- (4.9) a. Fric
 -Grv → Ret
 High
- b. (Fric)
 Ret → Grv
 High

The first of these processes backs nongrave high fricatives to retracted (retroflex) ones ($\check{s} > \check{\check{s}}$). The second one backs retracted (retroflex) high fricatives to grave ones ($\check{\check{s}} > x$). (Cf. the distinctions in Table 3 and the discussion in sec. 2.2.2).

In the Russian example, a nonoptimal contrast is also involved. Here the original laminal fricatives \check{s} \check{z} contrast not with plain apical sounds but with palatalized ones, namely $\check{\check{s}}$ $\check{\check{z}}$. Here it is important to note that \check{s} \check{z} and $\check{\check{s}}$ $\check{\check{z}}$ differ only by the property of apicality, the former being nonapical (laminal = tongue-blade) and the latter being apical (tongue-tip). The phonological space is thus too close; it is widened by backing the laminals \check{s} \check{z} to retroflex laminals $\check{\check{s}}$ $\check{\check{z}}$. Note also that \check{c} is not similarly backed: there is no $\check{\check{c}}$ [$t^{\check{h}}\check{s}$], this sound having already become nonpalatal c [t_s] in the standard language. However, in Belo-Russian, $t^{\check{h}}$ has become the palatal affricate \check{c} ; \check{c} is accordingly backed to $\check{\check{c}}$ in that language (cf. Kondrašov 1962:112ff.).

This completes the discussion of backing.

4.2. Coloring and bleaching.

Of the two fortitions which affect the chromaticity of consonants, coloring is the chromatization and bleaching the sonorization.

Coloring adds color to an otherwise achromatic consonant. Generally, it adds the color associated with the position of the consonant, e.g., chilics become labial and coronals become palatal. Dorsals frequently also become palatal even though the color associated with that position is velar.

Consonants also take on a given color when they occur in the environment of that color, e.g., chilics become palatal around palatal vowels. This is, properly speaking, a lenition, called color matching and will be discussed with the lenitions affecting color in sec. 5 below.

Bleaching subtracts color from a consonant that already has color. Generally, this color is totally lost, leaving behind no trace of its former existence, e.g., the English word new, still pronounced [nju] in some southern American English dialects, has become depalatalized (bleached) to [nu] in midlands and northern dialects. Occasionally, another process may intervene to prevent bleaching, e.g., the [\check{s}] of American English words like sure and sugar reflects an original [$\check{\check{s}}$]. This $\check{\check{s}}$ was laminalized (backed) to \check{s} ; cf. American English tissue [$t^{\check{h}}i\check{\check{s}}u$] and some British dialects [$t^{\check{h}}i_sju$].

Coloring and bleaching respond to opposing teleologies. Coloring affects especially lighter consonants and long ones, and bleaching especially darker and wider ones. Coloring and bleaching also occur in what I call "horse" contexts: coloring occurs around achromatic vowels,

and bleaching around vowels of another color or colors ("a horse of a different color").

We now turn to the discussion of these two fortitions, starting with coloring and ending with bleaching.

4.2.1. Coloring.

The coloring process may be stated as follows:

$$\left[\begin{array}{l} C \\ - \text{Color} \\ \alpha \text{ Pos} \\ ! \text{ Long} \\ ! \text{ lighter} \\ ! / \left[\begin{array}{l} V \\ -\text{Color} \end{array} \right] \end{array} \right] \rightarrow [\alpha \text{ Color}]$$

This statement may be read as: "A consonant which bears no color (i.e., is achromatic) and which is of a given position takes on the color associated with that position, especially when it is long, especially when it is lighter, and especially when it occurs around a vowel which itself bears no color (i.e., an achromatic one)." Again, the exclamation point (!) indicates "especially when."

Here are two general examples of coloring:

- (4.10) a. Polish l became backed to ω . This ω has now become labial w ($w = \omega\omega$, i.e., a velar ω with lip-rounding (labiality), ω). This is my own observation based on the fact that Polish \dagger is now pronounced $[w]$.)
- b. In some dialects of midwestern American English s has become palatal \check{s} in such words as racist, licorice, and grocery: ra $[\check{s}]$ ist, licori $[\check{s}]$, and gro $[\check{s}]$ ry, respectively. In the first two examples, the palatalization is probably due to the adjacent palatal vowels; in the third, it is due to the r, which has the effect of backing (retroflexing) the s , which in turn palatalizes.

In reference to this last example, compare the discussion of American English dialects in sec. 5 below (on lenitions) and the Caipira dialect of Portuguese in 4.2.1.2. (! Fronter).

We now turn to the specific environments which catalyze the process coloring.

4.2.1.1. ! Long.

The use of the term long here is mainly a cover term for two different phenomena that result in coloring (specifically, palatalization). One of these, the first, is an instance of a consonant cluster (sk), which is long on the basis that it is a cluster. The second example is an instance of geminate consonants (nn, ll) which are, properly speaking, specified long. The third example concerns both geminates and clusters.

First, the example concerning the consonant cluster sk:

- (4.11) a. Germanic sk has everywhere become š in modern English and German; in Dutch the sk assimilated to sx (a separate change and not an intermediate stage). Compare Gothic fiskos, modern Dutch vische [v'isxə], modern German Fisch [fɪʃ], and modern English fish (data from Moulton 1972).

Here is an example for long consonants:

- b. In the history of Spanish, the geminate nn and ll inherited from Latin have become palatal ñ (ñ) and l' (ll). In most dialects, l' has simplified to j. Examples include Spanish año [aɲo] from Latin annum 'year' and Spanish ella [eja] 'she' from Latin illa 'this (one) (fem.)' (Passy 1890: 151-52).

In the following example, geminates and clusters take on the color palatal:

- c. In Old French, geminate kk and the clusters sk nk rk rg became palatal k' s' n' r' r'g'; k g also became palatal k' g' initially (cf. Pope 1934:128).

This Old French change started this way:

- (4.12) a. $k \rightarrow k' / k, s, n, r, \# ___.$
 $g \rightarrow g' / r, \# ___.$

The palatal k' g' then developed this way:

- b. $k' \rightarrow t' \rightarrow t\check{s} \ (\rightarrow \ \check{s}).$
 $g' \rightarrow d' \rightarrow d\check{z} \ (\rightarrow \ \check{z}).$

The first two steps both occurred in Old French; the last (in parentheses) is an early modern French change. Here are some examples from Pope (ibid.):

c.

kk	bukka	>	botšə	>	bušə	>	boche	'mouth'
sk	eskina	>	estšine	>	ešine	>	eschine	'spine'
nk	franka	>	frāntšə	>	frānšə	>	franche	'free (fem.)'
rk	arka	>	artšə	>	aršə	>	arche	'arch'
#k	kantare	>	tšānter	>	šānter	>	chanter	'to sing'
rg	larga	>	lardžə	>	laržə	>	large	'wide'
#g	gudja	>	dzoie	>	žuee	>	joie	'joy'.

In all of these Old French examples, dorsal k g become palatal dorsal k' g', and not velar dorsal k^u g^u. If the process of coloring is stated correctly, we should have gotten the latter color (velar) instead of the former (palatal). At this point, I am awaiting evidence

that the sort of relationship holds for palatal j and velar w such that w fronts to j , even when w is attached to a consonant, i.e., kw would become kj ($= k^j$). For the time being, I will assume that only the primary colors (labial and palatal) are produced by coloring.

4.2.1.2. ! Fronter.

Coronals particularly are susceptible to palatalization, as in the following examples:

- (4.13) a. In Russian nursery words, $t d n$ become palatal $t^j d^j n^j$, e.g., [$t^j a^j t^j a^j$] 'father', [$d^j o^j d^j o^j$] 'uncle', [$n^j a^j n^j a^j$] 'nurse' (Jakobson 1968:79).
- b. $t d$ become palatal $t^j d^j$ (or at least are given preferential treatment) in the speech of Gregoire's one-year-old son (Jakobson 1968:78).

Chilics respond to the fronter condition by becoming labial in this example:

- c. In Russian, and in the English speech of Russian speaker YM, $p b f v m$ become labial $p^w b^w f^w v^w m^w$ before the nonpalatal vowels $i \epsilon$ (both occurring in Russian, only the last occurring in English), e.g., Russian bit [$b^w i^w t^j$] 'to be', v \ddot{i} [$v^w i^w$] 'you', and English maybe [$m^w \epsilon b j i$] when it occurs in YM's English.

The vowel ϵ , referred to in the example as "nonpalatal," is in fact nonpalatal to a Russian speaker, its palatal version being realized as [$j\epsilon$]. In Russian the vowel ϵ (nonpalatal) occurs mainly in foreign words.

Here are some other examples, which involve coronal grooved fricatives:

- d. In Old High German apical alveolar \hat{s} became laminal palatal \hat{s} initially before $p t m n l$; e.g., modern German Spiel 'game', Stein 'stone', Schmerz 'pain', Schnee 'snow', Schleim 'slime', all with \hat{s} from earlier \hat{s} (Joos 1952).

- e. In the Caipira dialect of Brazilian Portuguese especially in the town of Piracicaba, state of S\~ao Paulo) $s z$ become palatal $\acute{s} \acute{z}$ before a stop, e.g., ag\~osto [$ago\acute{s}tu$] 'August' (pers. comm., Roy Major).

- f. Also in the Caipira dialect (see above), $s z$ become palatal $\tilde{s} \tilde{z}$ in the same environment in which $t d$ become retroflex $\dagger d$, i.e., after \downarrow (like English \downarrow , or a "growly flap", possibly [$\check{r}\check{r}$]) (pers. comm., Roy Major). Examples include t\~er\~ço [$t\check{e}\check{r}\check{s}\check{u}$] 'third' and quatorze [$k^w a t\check{a} \check{z} i$] 'fourteen' (Rodrigues 1974:159).

4.2.1.3. ! Around an achromatic vowel.

This is the "horse" condition for the process coloring, in this case not a "horse of a different color" (cf. bleaching in sec. 4.2.2 below) but a "horse of no color." This refers to the fact that coloring typically occurs around vowels which bear no color, i.e., achromatic ones. Achromatic vowels are typically central ones, being neither palatal nor labial (nor velar). (Cf. Donegan 1978.)

All of the examples below show dorsal consonants becoming palatal (not velar) when occurring before achromatic vowels. Again, as in sec. 4.2.1.1. above on the tenseness context, a close relationship between palatal and velar coloring is assumed, the velarity of colored dorsal consonants being identified with palatality.

Here are the examples:

- (4.14) a. I have observed that dorsal k g become palatal dorsal $\acute{k} \acute{g}$ initially before stressed [o] in the informal speech of speakers from New York City, especially in those dialects which are "r-less", e.g., Carter [$\acute{k}^h'o:\check{r}\check{e}$], garden [$\acute{g}'o:dn$].
- b. The Old French change given above in sec. 4.2.1.1 involves Latin k g becoming $\check{t}\check{s} \check{d}\check{z}$ initially before stressed [o], probably through an intermediate stage of $\acute{k} \acute{g}$, e.g., Latin kantare > Old French chanter [$\check{t}\check{s}...$] 'to sing'; Latin gamba > Old French jambe [$\check{d}\check{z}...$] 'leg' (cf. Pope 1934:128).

This last change is usually explained by saying that initial sequences like $k\check{o}$ became $k\check{a}$ (a = lax low palatal vowel), which then became $\acute{k}\check{a}$. This palatal \acute{k} can then be explained as an assimilation to the palatal quality of the following a. The weakness of this argument is that it is then necessary to say that palatal a then changed back to a, leaving the palatality behind: $\acute{k}\check{a}$ (cf. Pope *ibid.*).

Another change in Old French converts γ (from k g intervocalically) to j before what Pope (1934:128) identifies as a and after i e, a. So, Latin baka 'berry' gives Late Latin [ba γ a] and Old French [baj ϵ], later [baie]. In these forms from Pope (*ibid.*), she has assumed a fronting change of [a] to [a].

A similar change is reported for Slavic (Channon 1972:34, quoting some rules attributed to Halle). The rule in effect is as follows:

$$(4.15) \quad k \ g \ x \rightarrow \acute{k} \ \acute{g} \ \acute{x} / C \check{i} \left(\left\{ \check{r} \right\} \right) _ + \check{o}$$

In this rule C = any consonant and + = a morpheme boundary. So, a form like otik+os 'Atticus' becomes otik \acute{o} s, eventually developing into Old Church Slavonic ot \check{c} cb (< ot \check{i} c \check{i} where c = [ts] 'father'); cf. Russian otec 'father'.

The symbol \check{o} appearing at the end of the rule above actually reflects Indo-European long and short \check{o} as well as long and short \check{o} (Kondrašov 1962:30). So, perhaps the process involved here responds not only to achromatic vowels like [a] but also to labial vowels like [o]. The labiality of o opposes in terms of color the palatal \acute{k} which develops. We might

assume, therefore, that what Slavists presume to be labial \check{o} was actually achromatic \check{o} .

This completes the discussion of coloring.

4.2.2. Bleaching.

The process bleaching, which is a sonorization, may be stated this way:

$$\left[\begin{array}{c} C \\ \alpha \text{ Color} \\ ! \text{ wider} \\ ! \text{ darker} \\ ! / \quad V \\ [-\alpha \text{ Color}] \end{array} \right] \rightarrow [-\text{Color}]$$

This statement is to be read as: "A consonant which bears any color loses that color (i.e., becomes bleached), especially when it is wider (i.e., more sonorant), especially when the color combined with it is darker, and especially when it occurs around a vowel of a different color (the "horse" condition)." Again, the exclamation point (!) is to be read as "especially when."

Here are a few general examples of bleaching:

- (4.16) a. Joan Velten at age 1;10 regularly bleached the palatal coronals \check{s} \check{z} to ts z (the latter voicing to z before a vowel). At age 2;0 (and presumably also at 1;10) she bleached palatal j to z .
 Examples: touch [$\text{d}\check{o}\text{ts}$], brush [$\text{b}\check{o}\text{s}$], shoes [$\text{z}\check{u}\text{s}$];
yellow [$\text{z}\check{o}\text{w}\check{o}$], yard [$\text{z}\check{o}\cdot\text{d}$] (Velten 1943).

Apparently the bleaching of j to z proceeds this way: (1) j narrows to z , (2) z assibilates to \check{z} , (3) \check{z} bleaches to z . Only in the last step, (3), does bleaching occur (cf. Stampe 1973).

- b. Arthur as late as age 5 bleached \check{s} to t , apparently with ts as an intermediate step, e.g., church [$\text{t}\check{t}\text{o}\cdot\text{t}$].

This same sort of thing appears not only in child language but also in the natural history of languages of the world:

- c. In western Romance (Gallo- and Ibero-Roman), the \check{s} developed from Latin \acute{k} (from k before palatal vowels) bleached to ts , giving rise to s in Old French (Pope 1934:125f.; Joos 1952) and also Old Spanish, which later gave Castillian θ (Harris 1969:196-98). The \check{z} from Latin \acute{g} , however, remained palatal.

We now turn to the specific environments in which bleaching is catalyzed.

4.2.2.1. ! Wider.

Both of the examples given here affect coronal approximants, the most sonorant (widest) consonants for the position coronal.

- (4.17) a. Belo-Russian and Slovak ř bleaches to r in all positions, e.g., Belo-Russian mora [morə] 'sea' as compared with Russian more [mɔřə] 'sea' (Kondrasov 1962:7, 113f., 151).
- b. In some dialects of British (and Colonial) English, r~l as early as the end of the seventeenth century became bleached to r l, e.g., blue [blju:] became [blu:], brute [brju:t] became (ultimately) [bu:t] (Wright 1924:56, 103).

This last change is related to the northern and midlands American English bleaching of all original palatal coronals before a labial vowel (see sec. 4.2.2.3 below). Since the change started with the widest consonants, applying them first in chronological terms, we can assume that wideness is nevertheless a catalyzing environment for bleaching.

4.2.2.2. ! Darker.

This context for bleaching refers to the darkness of the color associated with the consonant. The colors referred to here mainly include the retracted colors velar, pharyngeal, and rhotacized. The bleaching statement is then generally intended to exclude advanced colors like palatal and labial.

The loss of rhotacism mainly occurs in vowels, e.g., British and American English "r-less" dialects in which ɞ (= ər) becomes simply ə. But here is an example of loss of pharyngeality:

- (4.18) a. Maltese Arabic has lost all traces of pharyngeal coloring on consonants, substituting plain (non-colored) sounds for the standard Egyptian pharyngealized sounds (Borg 1973).

Examples of velarity being lost are rarer. Here is an example for one sound in which it typically occurs, namely labiovelar w (= ɰω).

- b. For Yiddish (Judeo-German), Sapir reports that in the Swabian dialects "w seems...to have become b," e.g., [leb] 'lion', [iŋbr] 'ginger', both from earlier Middle High German forms with w (Sapir 1915 [1949:264-65]).

I suspect that this w in the last example that "seems to have become" b, is really first a loss of velarity (w to ω) then a subsequent loss of labial color (ω to υ). The υ so produced might give a b-like effect. At any rate, velarity is lost in either interpretation.

4.2.2.3. ! Around a vowel of a different color.

In the examples which follow, palatal consonants become plain before labial vowels, and labial consonants become plain before palatal vowels. This catalyzing environment represents the "horse" condition, in which color is lost around a vowel of a different color ("horse of a different color").

- (4.19) a. In the northern and midland dialects of American English, palatal t' d' θ' ξ η have become bleached to t d θ s n before the labial vowel u , e.g., tune, dune, enthusiasm, sewer, new. (The palatality is preserved in certain dialects of southern American English, e.g., tune [$\dagger juyn$], etc.)
- b. Based on observations of my own, the labiovelars k^ω g^ω of Latin have become k g (and then k' g') before the palatal vowels i e in both Italian and Spanish. Yet in both languages k^ω g^ω remain before (stressed) a . Examples: Latin quid [$k^\omega id$] 'what' becomes Italian chè [$k'e$] and Spanish qué [$k'e$]; yet Latin quantu(m) [$k^\omega antu$] 'how much' becomes Italian quanto and Spanish cuanto, both [$k^\omega anto$].

Compare both these changes with the pre-Greek change of Indo-European labial dorsals k^ω g^ω g^ω^h to Greek t d $\dagger h$ before the palatal vowels i e . Apparently, the labial dorsals took on palatality before these vowels (i.e., colored), lost labiality (partially bleached) and became palatal dorsals before these vowels, fronted to palatal coronals, and these palatal coronals lost their palatality (i.e. bleached). Using k^ω as representative, this series of processes may have occurred: $k^\omega > k^y > k' > t' > t$. Thus two bleaching processes were involved in this Greek change.

This completes the discussion of bleaching.

5. Examples of Lenitions.

5.1. Some sonority lenitions.

The processes which will be discussed in this section all have the four properties of lenitions (syntagmatic processes). These properties are: (1) Lenitions are sequence-optimizing and assimilative, i.e., they are strategies for pronouncing sequences of sounds based on the concept of least effort (cf. Grammont 1933). (2) They are always context-sensitive, unlike fortitions, which can apply in a context-free way. (3) They are weakening, i.e., they generally produce outputs that are wider or laxer than the input. But even when they produce narrower or tenser outputs, such a change constitutes a compromise with the context and is done for the sake of ease of articulation. (4) They are most apparent in hypoarticulate (casual, careless, or lazy) speech.

So, while fortitions shape the conscious half of phonology, lenitions make those fortitive underlying representations into understandable segmented speech. Given a simple phonemic form like $/l\uparrow t\partial/$ little or $/m\partial n/$ man, certain lenitions add properties to these forms to make them easier to pronounce. For little, lenitions add flapping of the t , and simultaneously they add loss of schwa, syllabification of the l , and its

"subsequent" velarization. The resulting form is [lɪʀ̥ɥ] little. In the case of man, a nasalization lenition (assimilation) gives [m̃æn] man. The main purpose of the lenition, then, is to facilitate the pronunciation of segments occurring in sequence.

Naturally, the adjustments that a speaker makes during the progress of the speech event are very numerous and, more than that, vary from one speech event to the next. Aside from such seemingly unnoticeable speech phenomena (whose statement and laboratory analysis would fill volumes), there are certain grosser aspects, or processes, to which we can give some consideration here in this short space. We will accordingly consider six groups of phenomena, to be presented in succeeding subsections: (1) lenitions affecting stricture, i.e., narrowing/widening assimilations, (2) lenitions affecting orientation (to that stricture), i.e., tensing/laxing assimilations, (3) lenitions affecting glottal events, such as aspiration, glottalization, and voice, (4) other lenitions which affect the sonority of the sound to which they apply, e.g., lengthening and shortening, simplification, and loss.

5.1.1. Lenitions affecting stricture.

The two fortitions affecting stricture were narrowing and widening. Accordingly, the two lenitions affecting stricture in opposing ways are called narrowing assimilation and widening assimilation.

5.1.1.1. Narrowing assimilation.

Processes which narrow stricture in a dissimilative way are fortitions, e.g., the change of English of sθ and sf to st and sp, respectively, as in [sɪksθ] sixth and [spɪj] sphere. But there is another change in English that, while it brings about a narrowing of the stricture, is nevertheless a lenition. I am referring to the change of tʃ to tʃ̥ as in nature, or across a word boundary as in got you (to give gotchoo). The change in nature [n'eɪtʃ̥] is a lexicalized phonemic one: /ne'čəj/. But that of got you/gotcha is a morphophonemic sandhi rule, generally under the control of the speaker using it. The change of tʃ to tʃ̥ narrows the stricture of the second sound from an approximant j to a fricative ʃ̥. Nevertheless, it is a lenition since ʃ̥ represents a contextually based compromise of the j in the direction of the stop t.

5.1.1.2. Widening assimilation.

The process changing the combination dt to st, which apparently occurs in an early stage of the development of Indo-European, is a fortition. For instance, the root for 'eat', *ed, combines with the suffix for third person singular, *-ti, to produce *esti 'he eats'; cf. Old Church Slavonic estv 'he eats'. It is a fortition because it dissimilatively widens (and devoices) the d to s.

But a similar change in early Latin which widens t to s before s is a lenition because it acts in an assimilative way. Thus, the combination of ment, the root for 'mind', and the nominative singular suffix (athematic), namely -s, produces not *ment-s but mens-s (which simplifies to mens). (Cf. Stampe 1973.)

While the western Romance change of pt and kt to ft and xt, respectively, as reported on in sec. 3.1.2.1 above, represents what may or may not be a fortition (depending on one's interpretation), the similar

Slovak change of v to u in a closed syllable is clearly a lenition, brought about by the preceding vowel. For example, the Slovak word ovca is pronounced ['outɕə] and oblokov 'of windows' is ['oblokoʊ] (Kondrašov 1962: 151).

Also compare the development of f θ x from what must have been three different sources in the early stages of Germanic: (1) The occurrences of p t k plus one of the laryngeals H produced the strong aspirates ph th kh . (2) The voiceless stops p t k were generally weakly aspirated to p^h t^h k^h , except after s and in combinations the p^h t^h k^h , and probably early fell together with the strong aspirates. (3) The combinations p t k , not otherwise affected, dissimilated directly to ft xt , this change becoming phonemic only after the widening of the aspirates, from both sources, to f (θ) x . This last change, the widening of the first element of p t k to ft xt , was probably helped along by the influence of the preceding sound, whether vowel or resonant.

Other lenitions in this category include the Chipewyan "smoothing" of the grooved sounds ts ds dz s z to $t\theta$ $d\theta$ $d\theta$ θ θ (Goodman 1968), an assimilation to the vowel. Also compare my stepson Arthur's age 5 change of $t\check{s}$ to t , narrowing assimilation which goes in the other direction and constitutes an assimilation of \check{s} to the tangency ("touchingness") of the t .

5.1.2. Lenitions affecting orientation.

The two fortitions which affected the orientation, in terms of the tense/lax distinction, of a sound to its stricture were tensing and laxing. Similarly, the two lenitions affecting orientation are called tensing assimilation and laxing assimilation.

5.1.2.1. Tensing assimilation.

The tensing and subsequent devoicing of g to k in combination with a following t occurs in pre-Latin and is an example of tensing assimilation. By way of an example, Latin ag- 'to do, or move' combined with -tos, the marker of the past passive participle (masc. nom. sg.). This combination gives not *ag-tos but actus 'done, moved' in Classical Latin.

This sort of tensing, which is dependent upon the context, is different from the fortitive type, like the tensing and devoicing of Old High German voiced stops, a kind of tensing that affected the whole system.

5.1.2.2. Laxing assimilation.

The laxing of voiceless stops in Danish when they occur around fricatives (Spore 1965) is clearly a fortition, e.g., gisp [gisɕ] 'groan' (but Spore states that voiceless ɕ is voiced b in this position!). But the similar laxing of voiceless stops in a postvocalic and pre-stop context, which occurs in southern dialects of American English is probably a lenition. A typical example is the pronunciation of the word Baptist as [bæɔ̃tɪst] or even [bæbtɪst]. Cf. the Puerto Rican Spanish laxing, voicing, and widening of p to β in septiembre [sɛβtjɛmbɾɛ] 'September'.

5.1.3. Lenitions affecting glottal events.

The discussion here will be divided into two different sorts of

glottal events: (1) aspiration and glottalization (glottal enhancements), and (2) voicing and devoicing (voice).

5.1.3.1. Aspiration and glottalization.

The early Germanic aspiration of voiceless stops is clearly a fortition, a change that affected the whole system in a general way. And the present-day process that produces aspirated voiceless stops before stressed vowels or initially in English, German, and other Germanic languages is also obviously a fortition. Similarly, the process that produces glottalized voiced stops initially before stressed vowels in French and Spanish is also a fortition.

But the process in English that adds a separate (nonglottalizing) glottal gesture, as for a voiceless stop in sentence-final position, is a lenition. An example is the last \dagger in the phrase "And that's that!" [...ðæt^ˈ] and the \dagger is said to be unreleased. According to Catford (1977) sounds like \dagger^h actually involve an inserted glottal stop, but timed so as not to produce actual glottalization. Sometimes full glottalization results from this insertion, e.g., "I can't!" [...kæ.t̚ʔ], or "Just think!" [...θɪŋk̚ʔ]. This is the result of a lenition, which makes the utterance-final stop easier to say by relaxing the requirement to release that stop. Yet something "has to give," so to speak. The result is an inserted glottal closure ?.

A clearer example of this is the lenition of $p + (k)$, all going to glottal stop $ʔ$ before a homorganic nasal, in some forms of American English. Examples include cap'm [kæp̚m] 'captain', buttin' [bʌt̚n̩] 'butting', and I c'n [aɪk̚n̩] 'I can'. It regularly happens for \dagger before syllabic η in such words as button, cotton, satin, and for some dialects in words like important, sentence, pittance. In some dialects it also applies to \dagger before syllabic l : bottle [bɑt̚l̩] subtle [sʌt̚l̩], little [lɪt̚l̩], etc. All these are the result of a lenition, assimilating the stop in the direction of the previous vowel. As further evidence that this is a lenition, the glottal stop so produced is often lost outright, e.g., little [lɪt̚l̩] becomes li'l [lɪl̩].

This same lenition, which might be called glottal lenition, also probably accounts for the weakening or loss in Chinese of final voiceless stops (which existed in Ancient Chinese). The Cantonese dialect preserves the Ancient Chinese final consonants, the Suchow dialect preserves an intermediate glottal stop stage, but in the Peking dialect (Mandarin), these glottal stops are lost (data from Forrest 1948:Appendix II, 307ff.):

Ancient Chinese	Canton	Suchow	Peking	gloss
-jɛp	-ji:p	-jɛʔ	\jɛ	'leaf'
<u>-bjɛ†</u>	-pi:t̚	-biʔ	/pjɛ	'different'
-tɛk	-tak	-tɛ	/tɛ	'get'

The lines indicate tones. Compare example (4.9f) above in sec. 4.1.2.3.

From my own observations of the dialect of Atlanta, Georgia, we can establish yet another stage for glottal lenition. In that dialect, there is free variation between a glottal stop produced by lenition and a glottal stop accompanied by a partial, approximant-like gesture of the original voiceless stop. So the phrase Atlanta Hawks comes out [ˈæt̚l̩.ə.t̚ə hoʊz̩]

(careful speech) or [æʔ|ǎǎ hʊʔs] (casual speech). Notice the partial t gesture, signified by superscript ʔ, and the partial k gesture, signified by superscript ɥ, in the first (careful speech) example. So, for instance, for k we can establish these stages of weakening:

$$(5.1) \quad k \rightarrow kʔ(k^1) \rightarrow ɥʔ \rightarrow ? \rightarrow \emptyset$$

This series of processes is based on our cross-linguistic observations of English and Chinese.

Aspiration also occurs lenitively. In English, as we have noticed, it typically occurs before a stressed vowel and here it is a fortition. But aspiration lenition occurs when the amount of subglottal air pressure produced by the lungs is greater than that needed to pronounce the word. The result is word- (or sentence-) final aspiration. For instance, Tojolabal, a Mayan language, has an alternation between nonfinal t (and other voiceless stops?) and aspirated word-final t^h, e.g., /čataʔ/ [čataʔ^h] 'a kind of plant' (Gleason 1965:56).

A similar change, that can be interpreted as sentence-final aspiration lenition, occurs in Spanish of Mexico City, e.g., señor [senjoɾ^h]. This was reported above in sec. 3.2.2) as an example of devoicing. It is just possible that final ɾ becomes aspirated to ɾ^h and then only does the h-aspiration assimilate to the apical grooved ɾ to produce a period of apical grooved voicelessness in the form of ɶ.

5.1.3.2. Voice.

The following relationships are probably true: (1) Devoicing of obstruents is often a lenition. (2) Devoicing of resonants (and vowels) is always a lenition. (3) Voicing of obstruents is always a lenition. (4) Voicing of resonants (and vowels) is almost never a lenition. Schematically:

(5.2)		Obstruents	Resonants
	Devoicing	often	always
	Voicing	always	almost never

The devoicing of obstruents, as in Old High German, has already been discussed (see sec. 3.2.1). It is a fortition. But word-final devoicing of obstruents in languages like German and Russian is a lenition, being an assimilation to the following voiceless pause.

Devoicing of resonants, specifically the change of m n ŋ l r to m̥ n̥ ŋ̥ l̥ r̥, occurs in Welsh in initial position. These voiceless sounds are written mh nh ngh ll rh in Welsh (Lewis and Pedersen 1937:48). This is the result of a lenition, due to the assimilation of these sounds to the voicelessness of the preceding pause. Compare the lack in English of initial zb zd zg even though sp st sk do exist. Also compare the devoicing of vowels in Southern Paiute (Sapir 1933), which is due to a lenition.

The voicing of obstruents in Late Latin, specifically p + k to b d g between vowels, was a lenition (even though p + k were first laxed to p̥ t̥ k̥ by what was probably a laxing fortition, affecting all voiceless stops). For instance, Latin ripa became early Gallo-Roman *riba (later giving Modern French rive 'shore, bank') (Pope 1934:137).

Shortening, or flapping, of resonants is also a lenition. Flapping in English involves three steps, all lenitions:

- (5.4) a. $t d \rightarrow$ Short ($\check{t} \check{d}$)
b. $\check{t} \check{d} \rightarrow$ Gru (App) ($\check{r} \check{r}$)
c. $\check{r} \check{r} \rightarrow$ Voi (\check{r}).

In (a), the shortening occurs. In (b) these shortened stops become grooved, and thus, in effect, approximants. In (c), the voiceless flap gets voiced.

5.1.4.2. Simplification, loss, and other types of assimilation.

Simplification of clusters is a lenition, since it represents a loss of information. (Fortitions generally simply act to reorganize the information present in a sound system.) An example is the change of Latin initial $p|$ and $k|$ to l' in Spanish (spelled ll). Eventually, the l' delateralized to j , e.g., Latin plovere and Spanish llover [$j\text{ɔ}\beta'\text{e}\check{r}$] 'to rain' (pers. comm., David Griffin); Latin klamare and Spanish llamar [$\text{jam}'a\check{r}$] 'to call'.

Of loss, Hyman (1975:165) has said that a lenition is any change on its way to zero. Examples of loss include the deletion of word-final glottal stops in Mandarin (see above) and loss of flapped $t d$ in English child language.

5.2. Some position and color lenitions.

The lenitions to be dealt with will be presented in two sections: first, position assimilations, and second, lenitions affecting color. This last section will deal mainly with the effect of vowel color on consonants. We turn now to the discussion of position assimilation.

5.2.1. Lenitions affecting position.

Hutcheson's (1973) Ph.D. dissertation dealt specifically with the problem of consonant assimilation. He divided such assimilations into three groups: (1) partial assimilations, (2) fortuitous complete assimilations, and (3) complete assimilations. Partial assimilations dealt with clusters, as I have been using the word here. For instance, given a cluster like ks , a partial assimilation might produce xs . Here the stop k has become partially assimilated to the fricative s by becoming the fricative x .

Fortuitous complete assimilations result in geminate consonants (as do, as we shall see, complete assimilations). But they are fortuitous in that the two consonants involved are already similar, requiring the change of but one feature in a statement of the process. In the terms I have been using, fortuitous complete assimilations apply to affricates or co-articulations, or their mirror image clusters. For instance, pf becoming ff (what Hutcheson calls a "one-step change") is a fortuitous complete assimilation, because p and f were already very similar, i.e., differed (mainly) by one feature.

Complete assimilations, on the other hand, convert clusters differing both by sonority and position into geminate consonants, e.g., pn becoming nn or pp. A fourth type, which Hutcheson didn't consider (probably because he didn't consider them assimilations), would be the change of such clusters to a geminate composed of neither of the two members of the cluster but representing a common ground between them, e.g., pn becoming mm, or sk becoming šš. I know of no evidence, however, which would indicate the plausibility of these last two changes as a one-step change (but see below).

It is clear that Hutcheson considers only the last type, complete assimilation, of any theoretical interest. He seems to consider complete assimilation in terms of no intervening steps, all done in one change. However, the concept of two consonants being surreptitiously jammed together, as a case in point, does not seem too elegant.

I am inclined to think that the other two types, partial assimilation and fortuitous (complete) assimilation, are theoretically more significant. The case of partial assimilation tells us something significant about consonant categories. For instance, the change (as above) of ks to xs tells us that x has properties in common with both k and s. It has the position of k and the sonority of s. This is not an altogether obvious fact. The fact that ks becomes xs tells us that the x and s share properties even though x is made with a flat ("smooth") tongue and s with a "grooved" one. This problem is a moot one but ultimately an important one.

The case of fortuitous assimilation is also of significant theoretical interest in that it tells us something about phonetic identity. For instance, the change (as above) of pf to ff reveals that the two instances of ff are both occurrences of the same consonant sound, differing in phonetically predictable ways, i.e., the f's would be treated as instances of the same phoneme. Again, a moot point. But a very basic concept, that of phonemic identity or sameness, is behind it.

The case of complete assimilation then takes on an importance relative to the two preceding types. In this thesis, I have dealt with processes in an atomistic way, i.e., as producing small, discrete changes. The case of complete assimilation is then, to me, an occurrence of two partial assimilations, either occurring sequentially over time (in which case we notice their partialness) or occurring simultaneously in the mind or mouth of the speaker (cf. Donegan and Stampe 1979). In the latter case, of course, no external perception of the simultaneity in question would be possible. Two steps are involved but they both occur together.

The conjecture which supports the simultaneity of a two-in-one change is ineluctably its converse, i.e., the sequentiality of the two changes not occurring simultaneously but separately over a period of time. Two sequential changes show us things that one simultaneous two-in-one change does not. For one thing, they show us the possible "moves" that sounds can make. For another, they show us what "maps" we should draw, as linguistic cartographers, to delineate these changes, and perhaps to explain them.

A case in point is pn, as in Indo-European *swopnos 'sleep, dream'. In Latin it had become somnus, pn having partially assimilated to mn. In Italian the assimilation is complete: Latin somnus has become Italian

sonno, mn have "fortuitously" completely assimilated to nn (cf. French somme with mm). We notice two things: both changes were regressive and both were changes into a nasal, first being partially assimilated and later, completely.

It is altogether possible that pn could have become nn in one "fell swoop," to use Hutcheson's phrase. But rather than think of this as one change, I would consider it two simultaneous ones. The simultaneity of such a change is contingent on such factors as maintaining phonemic distinctions and certain cultural and societal considerations. Phonemic distinctions are, of course, important for keeping messages distinct: if a "fell swoop" change would destroy too many distinctions, it would be suppressed. The simultaneity of "fell swoop" changes, then, is largely happenstance, and sequentiality is more revealing.

I turn now to a discussion of general problems concerned with recognizing lenitions as distinct from fortitions and also to a position assimilation in Mandaic due to Malone (1971, 1972) (and cited by Hutcheson 1973). The Mandaic assimilation is one of glottals to buccals, e.g., ?t to tt, a change on which I have not previously touched.

5.2.1.1. Recognizing lenitions which affect position.

Based on an atomistic process approach to sound change, we would normally expect only fortitions to be of the "fell swoop" variety. This statement refers only to changes which "jump across" other categories of sound without being affected by them. For the sake of an example, consider the change of k^{ω} to p. Not so obviously, there are no intervening steps: the labial dorsal stop is replaced by a plain bilabial stop, corresponding in terms of position to the labial coloring that was originally part of the combination k^{ω} . This change is due to the fortition fronting, which especially affects consonants bearing a color. When this occurs, and there are no for-sure examples, the action is that of a fortition.

But the change of k^{ω} to p could also be viewed as a lenition in that the labial color and stop sonority of k^{ω} come together, as it were, to produce a "labial" stop, p.

Stampe has suggested (pers. comm.) that the change of k^{ω} to p is "typically" (I would say "sometimes") the result of an assimilation whereby the labial color ω narrows to p to produce kp, possibly a coarticulation at this point. The combination then simplifies to p.

Another possibility that I would consider (and which is not due to Stampe) is this series of changes: $k^{\omega} > kp^{\omega} > p^{\omega} > p$. Nandris (1963:112) cites a case somewhat like this one for the combination $b^{\prime} > bd^{\prime} > bd\check{z} > d\check{z}$. (The sequencing of the changes is my own conclusion, based on his observations on Rumanian dialects.) Granted that the change above for k^{ω} to p is only remotely possible, what makes this last change, of b^{\prime} to $d\check{z}$, quite possible? The answer is simple: the change of k^{ω} to p is an instance of fronting; the change of b^{\prime} to $d\check{z}$ is an instance of backing.

Fronting normally proceeds directly with color; the change of k^{ω} to p would be obviated by this correlation between fronting and color. Backing normally proceeds without color. The change then proceeds in assimilative increments, eventually "winding up" at a coronal palatal position. I would say on this account that reports of changes involving frontings that go through a variety of stages are probably misfounded. As an example, k^{ω} proceeds directly to p as a fortition, or k^{ω} assimilates

to kp and then simplifies to p as a series of two lenitions. Both, I think, are possible, but the lenition solution depends on the existence of intermediate kp in some dialect or speech style.

5.2.1.2. Position assimilation (glottals).

Malone (1971, cited in Hutcheson 1973) presents the following data on Mandaic, a language in which glottals assimilate to buccal sounds. In Mandaic, a glottal stop ? or glottal fricative h metathesizes with a preceding consonant C preceded in turn by a vowel:

(5.5) a. C Glot → Glot C / V _____

Then, for glottal stop, it either assimilates completely to the consonant or the vowel:

b. ? C → C C / V _____
V ? → V̄ / _____ C.

For the glottal fricative h, it may assimilate to either optionally. So these three possibilities exist:

c. h C → C C / V _____
V h → V̄ / _____ C
V h C remains.

Hutcheson (1973:66) says of this example that instances of both forms of substitution for glottal stop (and presumably all three forms for h) occur within the same dialect. Malone (1972, cited by Hutcheson) notes that the change of V? Vh to V̄ generally happens in prejunctural position: otherwise, ?C hC becomes CC (or C in these environments: CCV_____ or V̄CV_____, i.e., where the preceding C is geminate or the next preceding V is geminate). There are other limitations on the application of this process that need not concern us here.

This gives the idea of glottality as a full-blown position, a point which has not been previously touched upon. Hutcheson (1973) also gives examples of glottals becoming buccal sounds in Southern Paiute, Ancient Greek, and Finnish.

We might want to compare this glottal assimilation to the reverse process of introduction of glottals into original geminates. This occurs in Icelandic and there intervocalic pp tt kk become hp ht hk, e.g., kakka [kʰakka] 'to heap up'. Garnes (1976:13-22) describes such changes in connection with other related phenomena, reflected in such words as hekla [hɛkʰla] 'to crochet', henta [hɛntʰa] 'to be suitable, convenient', brúnna [brʰutna] 'brown (gen. pl.)', and alla [oʰʎa] 'all (gen. pl.)'. Here the spellings reflect an earlier form of pronunciation, which has been modified today to the form given in brackets.

5.2.2. Lenitions affecting color.

This topic will be discussed in two parts: (1) color matching and (2) vowel color phenomena.

5.2.2.1. Color matching.

Color matching, a lenition, may be stated this way:

$$(5.6) \quad C \rightarrow \alpha \text{ Color} / \left[\begin{array}{c} V \\ \alpha \text{ Color} \end{array} \right]$$

An obvious example is consonants becoming palatal before front (palatal) vowels. This situation holds for dorsal k g in English, which become k̟ ġ before i ɪ e ε æ a (all palatal vowels). This process also produced the situation in the Slavic languages, although the reflexes of the simple palatalizing process have become phonemic, thus making the situation more complex. For instance, in Russian, the high front vowel i has as its nonpalatal counterpart the high central vowel ɨ. All the other vowels ε α ɔ u have palatal counterparts ʲε ʲα ʲɔ ʲu, which palatalize preceding consonants (i varies freely with ʲi) (cf. Kondrašov 1962:86).

Another example is consonants becoming labial before labial vowels (generally back ones). So northern (Mandarin) Chinese -ʂɔ 'say, tell' is realized phonetically as [ʂʷɔ]; hence, the Yale spelling of this word as shwō. This assimilation proceeds regularly for other consonants appearing with -ɔ, regardless of tone. But note that -ʂu 'book' is not * [ʂʷu] but [ʂu]. The process is not triggered by -u itself (cf. Tewksbury 1948).

Another example would be consonants becoming pharyngeal before pharyngeal (low back) vowels. This is not the only source for pharyngeal consonants, however. The reader should try saying the sequence [sxα>] quickly; the result generally comes out [s^αα>]. This reveals a possible fricative origin for pharyngeal consonants.

Similarly, pharyngeal(ized) consonants may result from sequences of C + α, palatal(ized) consonants from C + j, and labial(ized) ones from C + w.

5.2.2.2. Vowel color phenomena.

Just as consonants can become colored next to vowels of the same color, so vowels can take on the color of an adjacent consonant. This usually applies to vowels that are lengthened, e.g. i becoming i:, and then broken (diphthongized), e.g. i: becoming ij̥ or iɛ̆. For the sake of the following statement, let us assume that the vowel broken off is always nonsyllabic ɛ̆. The following is a typical situation:

$$(5.7) \quad \varepsilon \rightarrow \alpha \text{ Color} / \text{---} \left[\begin{array}{c} C \\ \alpha \text{Color} \end{array} \right]$$

As an example, consider the following processes (which merely suggest the solution and are not entirely accurate), all of which apply to most dialects of American English:

- (5.8) a. $\begin{array}{l} C \\ \text{Dor} \end{array} \rightarrow \text{Pal} / \begin{array}{l} V \\ \text{Pal} \text{---} \end{array}$
- b. $V \rightarrow \text{Short} / \text{---} \begin{array}{l} \text{Stop} \\ \text{-Voi} \end{array}$

$$(5.8) \text{ c. } \left[\begin{array}{c} \text{V} \\ \text{-Short} \end{array} \right] \rightarrow \left[\begin{array}{c} \text{V} \\ \text{Short} \end{array} \right] \quad [\text{ə}] / \left[\begin{array}{c} \text{Pal} \\ \text{-Tns} \\ \text{---} \end{array} \right]$$

$$\text{d. } [\text{ə}] \rightarrow \text{Pal} / \begin{array}{c} \text{C} \\ \text{Pal} \end{array}$$

Rule (a) palatalizes the dorsals k g ŋ after a palatal vowel, e.g., sick [sɪk̟]. Rule (b) shortens vowels before voiceless stops; in effect, it lengthens them everywhere else, e.g., bid [bɪ:d], red [ɹɛ:d]. Rule (c) diphthongizes (breaks) nonshort vowels V̄ to short V̄ plus ə, provided they are palatal and lax. Rule (d), like the process in (5.7), makes ə palatal (typically ī following r̄ ε: [r̄ī] [ɛī], and ə following æ: [æə]). This is similar to the change of vowel V to diphthong Vi before palatals like dž ḥ (ṭṣ) in Sora and its sister Munda languages (pers. comm., David Stampe). Forms can be derived as follows:

(5.9)	bæg	wɪš̃	sɪk
(a)	bæḡ	---	sɪk̟
(b)	---	---	sɪk̟
(c)	bæəḡ	wɪəš̃	---
(d)	bæəḡ	wɪīš̃	---

The letters (a-d) refer, of course, to the four processes given above.

In the speech of DLH, who grew up in Brooklyn, NY, rule (c) above is modified to read:

$$(5.8) \text{ c' . } \left[\begin{array}{c} \text{V} \\ \text{-Short} \end{array} \right] \rightarrow \left[\begin{array}{c} \text{V} \\ \text{Short} \end{array} \right] \quad [\text{ə}] / \left[\begin{array}{c} \text{-Tns} \\ \text{---} \end{array} \right]$$

The only change is in the rightmost set of brackets: the feature Pal has been removed. Thus, any lax vowel is affected, but [ə] after labial (back) vowels would only become palatal before š̃ ž̃, as per rule (d) above, since back vowels would not make k g ŋ palatal as front vowels did by rule (a).

So she has the following diphthongs before š̃: push [pʰuɪ̄š̃], Josh [dʒɔɪ̄š̃], wash [wɔɪ̄š̃]. The combination [aš̃] doesn't occur in her dialect, but I once taught her the Russian word karandaš̃ 'pencil', native pronunciation [kəɾənd'aš̃]. She said it correctly several times and then submitted it to substitution: [kəðənd'əəš̃] (note the diphthong before s).

Most dialects have different strategies for back lax vowels. For instance, push is [pʰuuš̃], reflecting the lip-rounding that often occurs with s, e.g., ship [ʃɪp]. Labiality seems to be connected with ɹ also, e.g., rip [ɹɪp], and in fact replaces the expected wash [wɔuš̃] of such dialects with [wɔɹš̃] 'warsh'. Dialects around Atlanta, Georgia, also have water [wɔɹtəɹ], and ought [ɔɹt].

Retroflexion seems to be connected with r-coloring, retractedness, and palatality, as in the ruki rule. The ruki rule affected eastern Indo-European, mainly Indic (Sanskrit) and Balto-Slavic. In Sanskrit, s became retroflex ś when it occurred after r, u, k, or i. Zwicky (1970) reports that after k, there are no exceptions, all instances of ks being realized as kś. The other cases are more exceptional.

In Slavic, s becomes retracted to x (perhaps through intermediate s̄) after r u k i, except before the voiceless stops p t k. Compare Latin muscus '(house) fly' and Old Church Slavonic михъ (or мѣхъ), same meaning (Kondrašov 1962:31).

In Baltic, Lithuanian (but not Latvian) has s becoming palatal š, but only after k and r, e.g., Lith. veřsis 'ox, calf' but Latvian vērsis; Lith. āugštas 'high' but Latvian aūgsts (Endzelins 1971:50).

The relationship between r-coloring (or retroflexion) and palatality is also reflected in those dialects which have lost r- coloring and replaced it with palatality, e.g., worse [wəɹs]. There is, then, a reciprocal influence of consonants on vowels and vice versa.

This completes the discussion of lenitions.

6. Conclusion.

The basic premise of natural phonology is that sound substitution proceeds most often due to the nature of the sounds themselves. With this theme in mind, we might expect that the vast phonological differences which exist in the languages of the world are due to a discrete set of relatively predictable and relatively available natural phonological processes which are at the basis of all sound change. To validate this notion, I believe it is sufficient to state what these processes and their explanations are, and I have attempted to do this in this article.

As for the relative predictability of these natural processes, I have dealt with this in the form of the catalyzing environments in the discussion and explanation of the eight fortitions described in sections 3 and 4. As for the relative availability of these processes, I have drawn data from both children's and adults' speech, from both synchronic alternations and diachronic change, and from diverse languages and dialects and styles of speech.

The explanation for the relative predictability of these natural processes is made on the basis of the opposing teleologies of each pair of fortitions, so that half of them predict the augmentation or addition of one phonological parameter and the other half predict the augmentation or addition of some other (opposing) parameter. The explanation for the relative availability of these processes is in terms of their innateness within the individual, being part of the nature of the individual from his birth and being modified in a systematic way by the exigencies of his linguistic experience.

The phonological parameters upon which the predictability of these processes is based consist of these three: sonority, tonality, and chromaticity. Sonority distinguishes consonant from vowel and the various types of consonants and vowels from one another, e.g., obstruent versus resonant consonants or high versus low vowels. Tonality distinguishes position among consonants (and possibly among vowels) so that consonants which have been given a sonority type have also been given a position type, e.g., chelic, coronal, etc. Sonority works as the basic concept of the syllable, more sonorant sounds (e.g., vowels) being the syllable carriers and less sonorant sounds (e.g., consonants) being the syllable satellites. Tonality works within the domain of sonority to arrange the sequences of sounds within carrier and satellite. Chromaticity is most easily identified with color type, e.g., labial, palatal, etc. It works within the domain of tonality to increase the intensity of distinctons based on tonality.

These three parameters just mentioned work on the basis of progressing streams of speech. But such speech is based on, or made up of, various sounds which are differentiated on the basis of their inclusion or non-inclusion within certain classes of sounds. The various instances of sounds which are used by a given language to differentiate forms is called a sound system. The various properties which are inherent in these instances of sounds and which are used to differentiate members of classes of sounds are called sound features. Sound systems are, then, defined by sound features.

Within the sound system and by means of the sound features differentiating the individual members of the sound system, certain sounds or contingent groups of sounds are distinctive. They account for the phonological representation of sounds within a given language, i.e., they are phonemic. Phonological representation is determined on the basis of the typical pronunciation of a given phoneme, e.g., the English phoneme /p/ is a bilabial voiceless stop, typically aspirated. In cases where phonemes appear to "overlap", i.e., fall together in a given context, they are defined in terms of the sound intended (mental intention), represented in fully defined phonemic terms.

Sound systems may change due to shifts (transphonologizations, to use Jakobson's term), mergers (dephonologizations), and splits (phonologizations). As a subset of shifts, we may also consider the changes, typically of position, due to the principle of maximal differentiation, e.g., the backing of *š* in Old Spanish and of *š ž* in Russian to prevent their merger with frontier sounds.

Fortitions and lenitions apply in a contingent order, either over a period of time (diachronically) or simultaneously (synchronically). The essential unity of diachrony and synchrony can be seen as a free relationship in which pronunciation may vary over time or alternate at one time within the speech of a given individual or group of individuals.

I have attempted in this article to do four main things: (1) to establish the existence and relevance of a set of eight paired fortitions which describe the major class changes of consonant sounds, (2) to establish, at least tentatively, the opposing teleologies of each pair of fortitions, (3) to establish the three parameters on which consonant distinctions are based, and (4) to set up a system of features capable of reproducing these distinctions. In doing so, I have also talked about the basic premises of the study of natural phonology.

The four sonority fortitions, presented in sec. 3, concern changes in stricture (narrowing/widening) or changes in orientation to a stricture (tensing/laxing). The two position fortitions and the two color fortitions, presented in sec. 4, concern changes of tonality (fronting/backing) and changes of chromaticity (coloring/bleaching). They occur under typical relatable ways, i.e., narrowing and fronting typically occur for tense sounds, and widening and backing for lax sounds; and tensing and fronting occur typically for narrower sounds, and laxing and backing for wider sounds.

Lenitions have been dealt with sparsely here; the lenitions that were discussed in sec. 5 are probably only the tip of the iceberg. The fortitions dealt with here certainly do not represent the entire set of such processes. For one thing, they do not account for the metathesis and insertion of consonants.

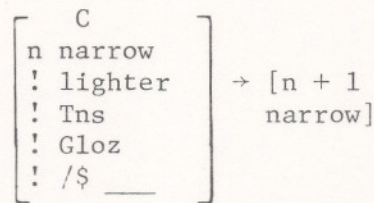
The study of phonology has in the past decades discredited such ideas as sound changes being due to climate or to racial membership. In the space of this paper, I have attempted to discredit the idea that sound change is formal or mechanical (as opposed to instinctive or intuitive) or that it results from the history of the language (as opposed to the intuition of the individual). Thus, language change is more to be identified with psychology than with computer and information science. And drift is then to be identified with similar linguistic intuitions, bringing about similar changes. The idea that Germanic and Armenian must have had a common Urheimat simply because they exhibit similar sound shifts should now be discredited. With the denial of such ideas, further endeavors concerned with the study of universal phonology should proceed on the basis of the innate availability and teleological predictability of natural phonological processes.

7. Appendix: Summary of processes.

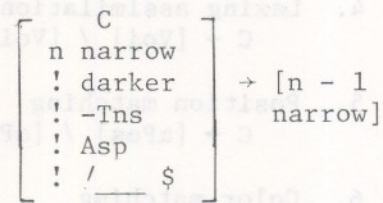
7.1. Fortitions.

C following the name of a fortition means chromatization: S means sonorization.

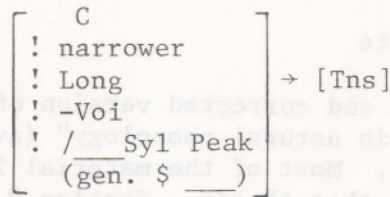
(7.1) Narrowing (C)



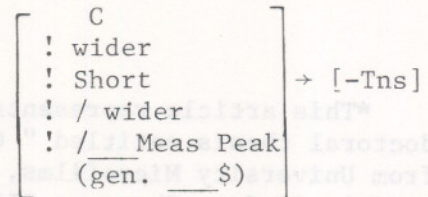
(7.2) Widening (S)



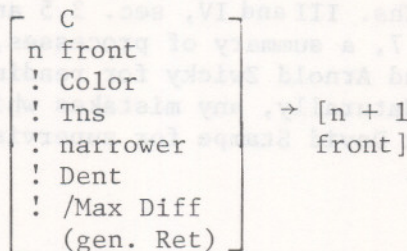
(7.3) Tensing (C)



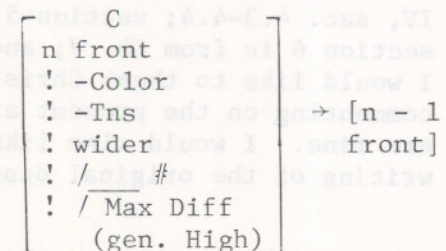
(7.4) Laxing (S)



(7.5) Fronting (C)



(7.6) Backing (S)



(7.7) Coloring (C)

$$\left[\begin{array}{c} C \\ - \text{Color} \\ \alpha \text{ Pos} \\ ! \text{ Long} \\ ! \text{ lighter} \\ ! / [V] \\ ! [-\text{Color}] \end{array} \right] \rightarrow [\alpha\text{Color}]$$

(7.8) Bleaching (S)

$$\left[\begin{array}{c} C \\ \alpha \text{ Color} \\ ! \text{ wider} \\ ! \text{ darker} \\ ! / [V] \\ ! [-\alpha\text{Color}] \\ ! (\text{"diff.color"}) \end{array} \right] \rightarrow [-\text{Color}]$$

7.2. Lenitions.

1. Narrowing & widening assimilation
C → [n narrow] / [n narrow]
2. Partial narrowing assimilation
C → [n + 1 narrow] / [n + 2 narrow]
3. Tensing assimilation (devoicing)
C → [-Voi] / [-Voi]
4. Laxing assimilation (voicing)
C → [Voi] / [Voi]
5. Position matching
C → [αPos] / [αPos]
6. Color matching
C → [αColor] / [αColor]

Footnote

*This article represents a revised and corrected version of my 1979 doctoral thesis entitled "Consonants in natural phonology" (available from University Microfilms, Ann Arbor). Most of the material in this article is from Chapters III and IV of that thesis. Section 0 is new; section 1 is from Ch. III, sec. 3.1-3.2; section 2 is from Ch. IV, sec. 4.1-4.2; section 3 is from Ch. III, sec. 3.3-3.4; section 4 is from Ch. IV, sec. 4.3-4.4; section 5 is from Chs. III and IV, sec. 3.5 and 4.5; section 6 is from Ch. V; and section 7, a summary of processes, is new. I would like to thank Chris Farrar and Arnold Zwicky for reading and commenting on the present article. Naturally, any mistakes which remain are mine. I would also like to thank David Stampe for supervising the writing of the original dissertation.

References

The abbreviation CLS refers to Papers from the Regional Meetings of the Chicago Linguistic Society.

- Allen, W. S. (1953). *Phonetics in Ancient India*. London: Oxford University Press.
- Anttila, Raimo. (1972). *An Introduction to Historical and Comparative Linguistics*. New York: Macmillan.
- Biggs, Bruce. (1971). The languages of Polynesia. In J. Donald Bowen, et al., Eds., *Current Trends in Linguistics*, Vol. 8, Part 1, 466-505. The Hague: Mouton.
- Borg, Alexander. (1973). The segmental phonemes of Maltese. *Linguistics* 109.5-11.
- Catford, J. C. (1977). *Fundamental Problems in Phonetics*. Bloomington: Indiana University Press.
- Channon, Robert. (1972). On the Place of the Progressive Palatalization of Velars in the Relative Chronology of Slavic. The Hague: Mouton.
- Chomsky, Noam, and Morris Halle. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Davison, Alice. (1971). A problem concerning the relative naturalness of phonological rules and phonological representation. CLS 7.332-39.
- Donegan, Patricia Jane. (1978). On the natural phonology of vowels. *Ohio State University Working Papers in Linguistics* # 23.
- Donegan, Patricia Jane, & David Stampe. (1978). The syllable in phonological and prosodic structure. In Alan Bell & Joan B. Hooper, Eds., *Syllables and segments*, 25-34. Amsterdam: North-Holland.
- Donegan, Patricia Jane, & David Stampe. (1979). The study of natural phonology. In Daniel A. Dinnsen, *Current Approaches to Phonological Theory*, 126-173. Bloomington: Indiana University Press.
- Edwards, Mary Louise. (1973). The acquisition of liquids. *Ohio State University Working Papers in Linguistics* 15.1-54.
- Endzelins, Janis. (1971). *Janis Endzelins' comparative phonology and morphology of the Baltic language*. Tr. by William R. Schmalstieg & Benjamins Jegers. The Hague: Mouton.
- Forrest, R. A. D. (1948). *The Chinese Language*. London: Faber & Faber.
- Garnes, Sara. (1976). *Quantity in Icelandic: Production and Perception*. Hamburg: Helmut Buske Verlag.
- Gleason, Henry A. (1965). *Workbook in Descriptive Linguistics*. New York: Holt, Rinehart & Winston.
- Goodman, Felicitas. (1968). *Proto-Athapaskan Vowels*. Unpublished M.A. thesis, Ohio State University.
- Grammont, Maurice. (1933). *Traité de phonétique*. Paris: Librairie Delagrave.
- Gray, Louis H. (1934). *Introduction to Semitic Comparative Linguistics*. New York: Columbia University Press.
- Harris, James W. (1969). *Spanish Phonology*. Cambridge, Mass.: MIT Press.
- Hockett, Charles F. (1958). *A Course in Modern Linguistics*. New York: Macmillan.
- Hutcheson, James W. (1973). *A Natural History of Complete Consonantal Assimilations*. Unpublished Ph.D. Dissertation, Ohio State University.

- Hyman, Larry M. (1975). *Phonology: Theory and Analysis*. New York: Holt, Rinehart & Winston.
- Jakobson, Roman. (1968). *Child Language, Aphasia, and Phonological Universals*. Tr. by Alan Keiler. The Hague: Mouton.
- Jakobson, Roman, and Morris Halle. (1956). *Fundamentals of Language*. The Hague: Mouton.
- Jones, Daniel. (1914). *The Pronunciation of English*. 2nd ed. London: Cambridge University Press.
- Joos, Martin. (1952). The medieval sibilants. *Language* 28.221-31. Also in Martin Joos, ed., 1957. (see); 372-78.
- Joos, Martin, ed. (1957). *Readings in Linguistics*. Chicago: University of Chicago Press.
- Kondrašov, N. A. (1962). *Slavjanskije jazyki*. 2nd ed. Moscow: Gosudarstvennoe učebno-pedagogičeskoe izdatel'stvo.
- Kruisinga, E. (1924). *A Grammar of Modern Dutch*. London: George Allen & Unwin.
- Ladefoged, Peter. (1971). *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Lehiste, Ilse. (1970). *Suprasegmentals*. Cambridge, Mass.: MIT Press.
- Leopold, Werner F. (1947). *Speech Development of a Bilingual Child: A Linguist's Record*. Vol. 2. Evanston, IL: Northwestern University.
- Lewis, Henry & Holger Pedersen. (1937). *A Concise Comparative Celtic Grammar*. Göttingen: Vandenhoeck & Ruprecht.
- Major, Roy C. (1977). Phonological differentiation of a bilingual child. *Ohio State University Working Papers in Linguistics* 22.88-122.
- Malone, Joseph L. (1971). Systematic metathesis in Mandaic. *Language* 47.394-415.
- Malone, Joseph L. (1972). The Mandaic syllable-adjustment circuit and its historical origins. *CLS* 8.473-81.
- Mandelbaum, David G., ed. (1949). *Selected Writings of Edward Sapir*. Berkeley: University of California Press.
- Meillet, André. (1936). *Esquisse d'une grammaire comparée de l'arménien classique*. 2nd ed. Vienna: Pp. Mekhitharistes.
- Morris Jones, John. (1913). *A Welsh Grammar: Historical and Comparative*. Oxford: Oxford University Press.
- Moulton, William G. (1972). The Proto-Germanic non-syllabics (consonants). In Frans van Coetsam & Herbert L. Kufner, Eds., *Toward a Grammar of Proto-Germanic*, 141-173. Tübingen: Max Niemeyer Verlag.
- Nandris, Octave. (1963). *Phonétique historique du roumain*. Paris: Librairie C. Klincksieck.
- Nguyen Dang Liem. (1970). *Vietnamese Pronunciation*. Honolulu: University of Hawaii Press.
- Passy, Paul. (1890). *Étude sur les changements phonétiques et leurs caractères généraux*. Paris: Librairie Firmin-Didot.
- Pokorny, Julius. (1959). *Indogermanisches etymologisches Wörterbuch*. Bern: Francke Verlag.
- Pope, Mildred K. (1934). *From Latin to Modern French with Especial Consideration of Anglo-Norman*. Manchester: Manchester University Press.
- Pribsch, R., & W. E. Collinson. (1934). *The German Language*. 6th Ed. London: Faber and Faber.
- Prokosch, Eduard. (1939). *A Comparative Germanic Grammar*. Philadelphia: LSA, University of Pennsylvania.

- Rodrigues, Ada Natal. (1974). O dialeto caipira na região de Piracicaba. São Paulo: Editora Ática.
- Ruhlen, Merritt. (1976). A Guide to the Languages of the World. Stanford: Stanford University Language Universals Project.
- Sapir, Edward. (1915). Notes on Judeo-German phonology. The Jewish Quarterly Review, n.s. 6.231-66. Reprinted in Mandelbaum (1949):252-72.
- Sapir, Edward. (1933). The psychological reality of phonemes. Mandelbaum (1949), (see), 46-60. Originally published as La réalité psychologique des phonèmes. Journal de psychologie normale et pathologique 30.247-265.
- Saussure, Ferdinand de. (1959). Course in General Linguistics. New York: McGraw-Hill.
- Skorik, P. Ja. (1968). Čukotskokamčatskie jazyki: voedenie. Jazyki narodov SSSR, vol. 5, 235-247. Leningrad: Izdatel'stvo "Nauka".
- Spore, Palle. (1965). La langue danoise. Copenhagen: Akademisk Forlag.
- Stampe, David. (1973). A Dissertation on Natural Phonology. Ph.D. Dissertation, University of Chicago. (1980), New York: Garland.
- Tablan, Andrea A. & Carmen B. Mallari. (1961). Filipino-English/English-Filipino Dictionary. New York: Washington Square Press.
- Tewksbury, M. Gardner. (1948). Speak Chinese. New Haven: Yale University Press.
- Thompson, Laurence C. (1965). A Vietnamese Grammar. Seattle: University of Washington Press.
- Trubetzkoy, N. S. (1969). Principles of Phonology. Tr. C. A. M. Baltaxe. Berkeley: University of California Press.
- Velten, H. V. (1943). The growth of phonemic and lexical patterns in infant language. Language 19.281-292.
- Vogt, Hans K. (1974). Armenian language. Encyclopaedia Britannica, 15 ed., vol. 2 (Macropaedia), 23-24. Chicago: Encyclopaedia Britannica.
- Walker, Willard. (1975). Cherokee. In James M. Crawford, Ed., Studies in Southeastern Indian Languages, 189-236.
- Williams, Edwin B. (1962). From Latin to Portuguese. 2nd ed. Philadelphia: University of Pennsylvania Press.
- Wright, Joseph, & Elizabeth Mary Wright. (1924). An Elementary Historical New English Grammar. London: Oxford University Press.
- Wright, Joseph, & Elizabeth Mary Wright. (1928). An Elementary Middle English Grammar. 2nd ed. London: Oxford University Press.
- Zwicky, Arnold M. (1970). Greek-letter variables and the Sanskrit ruki class. Linguistic Inquiry 1.549-55.
- Zwicky, Arnold M. (1972). Note on a phonological hierarchy in English. In Robert P. Stockwell & Ronald K. Macaulay, Eds., Linguistic Change and Generative Theory. Bloomington: Indiana University Press.

