

Feature Theory

Origins of feature theory. A very detailed theory of speech articulation and classification of sounds was developed nearly three millenia ago by the ancient Sanskrit grammarians, and this forms the foundation for contemporary feature theory. The main concern of the Sanskrit grammarians was the accurate preservation of the language of the Vedas, whose correct pronunciation was deemed essential to the proper working of the incantations. Modern Western linguistics begins with Sir William Jones' celebrated paper in 1786 on the relationship of Sanskrit to Latin, Greek and Germanic. Jones, a pre-eminent Orientalist, had an extensive knowledge of Sanskrit and the phonetic theories of the Indian phoneticians, and thanks to the introduction of Indian scholarship through the work of people like Jones, that tradition significantly influenced western linguistics. Ancient Indian phonetic theory identified many phonetic categories that are familiar today, such as 'articulator', 'vowel', 'voice': see Allen (1953) for discussion of the Indian Phoneticians.

In addition to developing a descriptive phonetic tradition, the Indian grammarians developed a system of formal grammatical rules — the first generative system, in the sense of Chomsky (1957) — culminating in the *Aṣṭādhyāyī* of Pāṇini. Pāṇini presents an algebraic system of sound classification, one serving the needs of a set of phonological rules in his grammar, in the form of the *Śivasūtras*, a clever 14-line listing of the segments of Sanskrit which reads as:

a-i-uṅ; ṛ-lk; e-oṅ; ai-auc; ha-ya-va-raṭ; laṅ; ña-ma-ṇa-ṇa-nam; jha-bhañ; gha-ḍha-dhaṣ;
ja-ba-ga-ḍa-daç; kha-pha-cha-ṭha-tha-ca-ṭa-tav; ka-pay; ça-ça-sar; hal

This presentation of sounds reflects their grouping in rules (which reflects phonetic properties), and the bold consonant at the end of each line indicates the end of a group. While the grouping is not a perfect mirror of phonetic properties, you can see that the nasals are grouped together, as are the voiced consonants, the stops vs. fricatives vs. sonorants; place of articulation is also reflected in the ordering, though not in a trivial way.

Natural classes of sounds which undergo or trigger rules are encoded by mentioning the first sound in the class and the final consonant at the end of the class. Sanskrit has a rule where syllabic /i u ṛ ḷ/ become nonsyllabic [y w r l] before a vowel or diphthong, stated in the *Aṣṭādhyāyī* as "ik-o yaṅ ac-i", literally "instead of the class 'ik' there is the class 'yaṅ' before the class 'ac'.". The class 'ik' is defined, referring to the first two lines, as everything from [i] through [ḷ] (the sound before the terminator 'k') thus /i u ṛ ḷ/, and 'ac' means everything from [a] to the terminator 'c', meaning [a i u ṛ ḷ e o ai au], viz. all syllabics.

In western linguistics, phonetics, which is the fountainhead of distinctive features, developed through the research of Sievers, Sweet, Jespersen, Jones and others, and phonology is rooted in Sweet's 1887 distinction between sounds that depend on their environment — allophones — and sounds that can establish word differences. N. S. Trubetzkoy, one of the founders of the Prague School of Linguistics, was a major influence in the development of contemporary feature theory. Trubetzkoy (1939) looks beyond the physical aspect of language sound and focuses on how phonetic properties function in a language to define a system of phonological contrasts (oppositions, in his terminology). His taxonomy includes functional classifications such as

bilateral vs. multilateral oppositions, proportional vs. isolated, constant and neutralizable oppositions, as well as privative, gradual and equipollent oppositions. The concept of neutralization is at the heart of phonological analysis: according to Trubetzkoy, we should uncover the underlying, non-neutralized nature of sounds when we analyse a set of alternations.

Two other concepts, **marked** and **privative**, play an important role in phonological theory. For Trubetzkoy, in a privative opposition one member (for example *p* in the pair *p*, *b*) lacks a phonetic mark — voicing — and the other one has the phonetic mark — it is the marked member. The concept of markedness has played a major role in generative grammar (see Chomsky & Halle 1968, Kean 1975, Battistella 1990, 1996), and has taken on a rather different meaning, as being “uncommon”, “at a formal disadvantage”, or “statistically dispreferred”. In Optimality Theory (McCarthy 2002) the concept of a “markedness constraint” is further reinterpreted as meaning “well-formedness constraint”, stating that certain structures (segments and sequences) are dispreferred (this is because OT states what is disallowed, rather than stating what is allowed). While the theory of distinctive features which is presented in Chapter 6, relying on plus and minus values for features, thus views features as being equipollent (two equal values), more recent research, especially underspecification theory and feature geometry (see below) supports a privative view of some features, e.g. *b* being marked with the single-valued feature [voiced] and *p* having no specification for voicing.

The related concern of the Praguean linguist Roman Jakobson was the question of possible contrasts in human language, and in collaboration with Morris Halle and Gunnar Fant (Jakobson, Fant & Halle 1952; Jakobson & Halle 1956) developed the theory of distinctive features. This research identified 12 to 15 features which were claimed to handle all phonological contrasts in human languages. The Jakobsonian features were typically defined in terms of acoustic properties rather than articulatory ones, since acoustic research was quite innovative at the time. Another significant contribution of this theory was to give all features two values, plus and minus. Rather than a gradual (scalar) opposition of vowel height with four values, the Jakobsonian system defined vowel height by two features [diffuse] (≈[high]) and [compact] (≈[low]) which says whether there is a concentration of acoustic energy in the central region.

The SPE theory of features. The version of feature theory presented in Chapter 6 is an adaptation of Chomsky & Halle 1968 — the so-called SPE features — which was the next step in the development of feature theory, and still serves as a theoretical baseline for current work on features. The differences between the SPE features and the Jakobson & Halle features are more along the lines of a refinement. Greater attention was given to adding articulatory definitions of features in the SPE system which were largely lacking from the previous definitions, and the features themselves were revised and expanded in a number of ways. In the course of a decade of research on phonology, a number of inadequacies in the Jakobsonian system were uncovered, for example the existence of non-strident affricates which posed a problem for the earlier theory of affricates, or the discovery that [flat] was inadequate as a model of retroflex, round, uvular and pharyngeal consonants, and these are addressed in the SPE feature system.

The SPE theory of features differed from its predecessors in a significant way, that rather than being designed just to handle phonemic *contrasts*, it is intended to describe all linguistic phonetic properties:

The total set of features is identical with the set of phonetic properties that can in principle be controlled in speech; they represent the phonetic capabilities of man and, we would assume, are therefore the same for all languages. (p. 294-5)

Binarity of features is assumed for lexical entries but not throughout the grammar: they state (p. 297) that “the phonetic features are physical scales and may thus assume numerous coefficients, as determined by the rules of the phonological component”, and certain rules in SPE operate in terms of specified numeric stress coefficients. See Clifton (1976) and Johnson (1981) and references therein for examples of phonetic implementation rules stated with numeric coefficients.

Notation. The major notational devices of the SPE theory are presented in Chapter 6, but there are a few other notations and points regarding notation that need brief mention. First, the SPE theory also defined an “angled bracket” notation, which is used to express discontinuous dependencies, i.e. “if X is true here, Y is true there”. Unfortunately, this notation is subject to slightly different interpretations within SPE, so what is given here represents “common usage”. One typical use of angled brackets is to express the fact that in Old High German, back vowels become front before a front vowels but in addition, if the vowel is a short low vowel then it becomes non-low. Thus, /wurm-i/ → [würmi] “worms”, /not-i/ → [nöti] “need”, /ta:t-i/ → [tæ:ti] “deed” but /gast-i/ → [gæsti] “guest”. This dependency can be expressed in the rule with angled brackets:

$$\left[\begin{array}{l} + \text{ syl} \\ \langle + \text{ low} \rangle \\ \langle - \text{ long} \rangle \end{array} \right] \rightarrow \left[\begin{array}{l} - \text{ back} \\ \langle - \text{ low} \rangle \end{array} \right] / \text{---} \left[\begin{array}{l} + \text{ syl} \\ - \text{ back} \end{array} \right]$$

Generally, an expression of the form

$$\left[\begin{array}{l} \text{X} \\ \langle \text{Y} \rangle \end{array} \right] \rightarrow \left[\begin{array}{l} \text{Z} \\ \langle \text{Q} \rangle \end{array} \right] / \text{---} [\text{R}]$$

expands to the set

$$\left\{ \begin{array}{l} \left[\begin{array}{l} \text{X} \\ \text{Y} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{Z} \\ \text{Q} \end{array} \right] / \text{---} [\text{R}] \\ \left[\text{X} \right] \rightarrow \left[\text{Z} \right] / \text{---} [\text{R}] \end{array} \right\}$$

Thus concretely, the Old High German rule expands to:

$$\begin{bmatrix} + \text{syl} \\ + \text{low} \\ - \text{long} \end{bmatrix} \rightarrow \begin{bmatrix} - \text{back} \\ - \text{low} \end{bmatrix} / \text{---} \begin{bmatrix} + \text{syl} \\ - \text{back} \end{bmatrix} \quad \text{and} \quad [+ \text{syl}] \rightarrow [- \text{back}] / \text{---} \begin{bmatrix} + \text{syl} \\ - \text{back} \end{bmatrix}$$

One problem with the notation is that it is not clear whether it expresses a significant generalization needing to be captured in rule formalism. While it would clearly miss a generalization to claim that Old High German had two unrelated umlaut rules, it is possible that the second, more general rule is the correct statement of OHG umlaut per se, and that the appearance of [ɛ] rather than [æ] reflects an additional fact of the language, that such a segment does not exist. Hence the effect explicitly encoded in the rule with angled brackets may be an automatic structure-preserving side-effect created by the lack of short [æ]. Such effects may arise because of specific rules applied later in the derivation, i.e. /gasti/ → gæsti and then later [gæsti] because of a rule [æ] → [ɛ]. Although there are many possible examples of angled brackets in rules, there do not seem to be any truly compelling examples which resist reduction to one general rule and a “clean-up” rule of this type; hence angled brackets was not a widely embraced notation.

Another notation used in SPE is the “bracket-dash”¹ notation, exemplified by SPE’s Rounding Adjustment rule:

$$\begin{bmatrix} \alpha \text{round} \\ + \text{back} \\ \text{V} \end{bmatrix} \rightarrow [-\alpha \text{round}] / \left\{ \begin{array}{l} \left[\text{---} \right] \\ [- \text{tense}] \\ \left[\text{---} \right] \\ \beta \text{low} \\ \beta \text{round} \\ \left[+ \text{tense} \right] \\ \text{---} \\ \text{V} \end{array} \right\}$$

This states that a back vowel takes on the opposite roundness, when it satisfies any one of three conditions. First, it may be lax; second, it may itself be tense and either low, round or non-low, non-round; or, it may be before another vowel. The properties contained in [] on the left side of the arrow are combined with those contained in [] after the slash. The point of the notation is to allow you to factor out certain properties of the class of target segments, especially when the structural change is held in common, and collapse rules which could not otherwise be collapsed. Whether or not this descriptive power is a good thing is another matter: feature theory says that lax back vowels, tense low round vowels and tense non-low non-round vowels are not a natural class, so it seems ill-advised to include a notation that negates the predictions of feature theory, so that three rules can be collapsed into one.

¹ This notation did not receive an official or popular name, to the best of my knowledge.

Another notation commonly used in formulating phonological rules was the mirror-image notation. A number of languages have rules that lower high vowels to mid vowels either before or after uvular consonants (examples: Lushootseed, Inuit, Cuzco Quechua). This can be expressed with braces, as in:

$$[+syl] \rightarrow [-hi] / \left\{ \begin{array}{l} \left[\begin{array}{l} -syl \\ +back \\ -hi \end{array} \right] \\ \left[\begin{array}{l} -syl \\ +back \\ -hi \end{array} \right] \end{array} \right.$$

but this makes no connection to the fact that the triggering segments are the same whether it precedes or follows the vowel. The convention developed of replacing “/” with “%” to mean “this sequence or its mirror image”, thus:

$$[+syl] \rightarrow [-hi] \% \left[\begin{array}{l} -syl \\ +back \\ -hi \end{array} \right]$$

Finally, one last type of rule is worthy of mention, namely so-called transformational rules. The conventional form of a rule in phonology is $X \rightarrow Y / Z _ Q$. This is equivalent to a rule of the form $ZXQ \rightarrow ZYQ$, the difference being that in the conventional format, the changing segment is factored out from the non-changing segments. The conventional format is appropriate for the typical context-sensitive phonological rule (“context-sensitive” refers to a particular formal type of rule in what is known as the Chomsky hierarchy, where context-free rules are the least powerful and unrestricted rewrite rules are the most powerful).

Certain processes cannot be formalised using the context-sensitive notation — such rules affect more than one segment at once. Any rule of phonological movement, such as metathesis, tone-movement (as we will see in Chapter 10) or a rule which combines two segments into a third segment must be stated with this transformational notation, as an unrestricted rewrite rule. Terms in such a rule are often numbered, so that it is possible to identify what segment is being referred to. Thus a rule which transposes consonant across vowels when the first is [+coronal] and the second is [+back] — i.e. /taki/ → [kati] — would be stated as:

$$\begin{array}{ccccccc} [+cor] & V_0 & C & & & & \\ & & [+back] & & & & \\ 1 & 2 & 3 & \rightarrow & 3 & 2 & 1 \end{array}$$

The objection to this notation is that you can do “anything” with it. This is not literally so, because there remain vast numbers of imaginable rules which cannot be stated using the standard notation (examples: string reversal on strings of arbitrary length; swapping segments numbered 2^{2^n} and $2^{2^{n+1}}$ counting from the left), but still this does give phonological rules considerable generative power, and the question arises whether this power is justified.

One class of rules that seem to require this notation is the rather large set of fusion rules, e.g. glide formation with compensatory lengthening of a neighboring vowel; consonant deletion with compensatory nasalization; deletion of one vowel with transfer of its tone to a neighbor. These kinds of processes receive an elegant explanation in non-linear phonology — see Chapter 10. The other class of attested rules apparently demanding the power of unrestricted rewrite rules is the class of phonological movements — metathesis, primarily. There were a number of responses to metathesis rules in the literature, primarily exile (decreeing that the process in question was not part of phonology and was part of a separate component, morphology, which might have such power) or denial (reanalysing the phenomenon so that unrestricted rewrite rules are not needed to accomplish apparent movement). The denial approach is valid in some cases, and seems potentially appropriate for a number of CCV \rightarrow CVC cases. Many of these can be reanalysed as a combination of harmonic epenthesis and apocope, i.e. CCV \rightarrow CV_iCV_i \rightarrow CV_iC (Maltese Arabic or Sierra Miwok provide examples of this type). Not all metathesis cases can be reduced to plausible multi-step derivation, thus in Faroese (as well as Lithuanian), /skC/ becomes [ksC], cf. *fesk* ‘fresh (fem)’, *fesk-or* ‘fresh (masc)’, *feks-t* ‘fresh (neut)’. Examples of metathesis are documented extensively at <http://www.ling.osu.edu/~ehume/metathesis>, Elizabeth Hume’s metathesis web page.

Formal Simplicity. Many aspects of phonology in the 60’s are best understood in terms of the importance of evaluating grammars in terms of the number of symbols used in their formalization. One of the fundamental ideas of the theory of generative grammar is that a theory should contain an evaluation metric which distinguishes between empirically equivalent accounts of a language. The evaluation metric articulated in SPE states (p. 334) that ‘The “value” of a sequence of rules is the reciprocal of the number of symbols in its minimal representation’. Hence a rule which requires 12 symbols to state is less highly valued than a rule which requires only 8 symbols to state. It is not too hard to extend this idea to lexical representations, so that one can analogously distinguish between highly valued underlying representations which use few symbols, and less-valued representations which use more symbols.

The notational conventions proposed in SPE and subsequent publications were strongly driven by the consideration of formal simplicity. To take an example, a common rule assimilating nasals in place of articulation to the following consonant may need to assimilate 4 separate features. In order to explicitly list each of the feature configurations, 16 separate collections of 8 feature-mentions would be needed, resulting in an additional cost of 120 symbols — Greek letter variable notation allows a significant simplification in the statement of such a rule.

The idea of economy of formulation was not new to generative grammar — it was directly inherited from the preceding taxonomic theory, and was epitomized by the radical formal simplicity of Pāṇini’s *śabdakṛta lāghava* style — it is said that a Sanskrit grammarian would rejoice more over saving a half a syllable than over the birth of a son. The problem posed by the

simplicity metric is that it tended to be seen as a primary end in itself. Numerous rules in the SPE account of English had a form similar to that seen in the Rounding Adjustment rule above — many symbols glued together into one rule, unified by a notational device which nets a savings of some half-dozen symbols. This does not mean that simplicity is an invalid criterion for choosing between analyses, and as Chomsky & Halle emphasize (p. 331), a correct theory of grammar must:

develop a system of formal devices for expressing rules and a set of general conditions on how these rules are organised and how they apply. We postulate that only grammars meeting these conditions are “entertained as hypotheses” by the child who must acquire knowledge of a language.

What remains, then, is a set of grammars that are in the domain of grammar-evaluation. It follows that if phonological theory does not contain certain devices, or restricts their usage in well-defined ways, then a child cannot entertain hypothetical grammars with such symbol-saving devices in them, no matter how many symbols are saved. The role of formal simplicity is to eliminate from consideration those grammars which are “theoretically possible” accounts of a set of data, but which use more symbols in their formalization than are required in order to generate the language. The more highly articulated the metatheory of grammar is — the more restrictive the notion “possible grammar” can be made — the smaller the role becomes for the notion of simplicity in language acquisition.

Post-SPE developments in features. Certain features are proposed tentatively in SPE but were not generally adopted, including [suction], [pressure] and a distinction between [delayed primary release] i.e. [delayed release] for affricates, and [delayed secondary release] which is primarily relevant for clicks — see Sagey (1986) for a reanalysis of click “effluxes”. The SPE feature [covered] is the same as Advanced Tongue Root, which is now considered to be the standard name for the feature. Halle & Stevens (1971) introduce the features [c.g.] and [s.g.], replacing [heightened subglottal pressure], [pressure] and, in part, [suction], and also introduce two features [stiff vocal cords] and [slack vocal cords] to describe voicing: the features [c.g.] and [s.g.] are now standard, whereas [stiff] and [slack] have not met with wide acceptance. Other features have been argued for after the publication of SPE, most notably [labial] (Anderson 1974, Campbell 1974) and [grave] (Hyman 1973, Vago 1976, Odden 1978). See Keating 1988 for a detailed overview of distinctive features in the post-SPE era.

A major change in phonology began in the mid 1970’s with research by Leben, Goldsmith, Clements and others on the representation of tone. The autosegmental model of representation was proposed to resolve an accretion of technical problems with the standard SPE model especially as it applied to tone. The proposal was made that features organised into smaller sub-segmental units, which function like whole segments (segments in their own right, i.e. autosegments). In the first decade of autosegmental phonology, the features remained largely those of SPE, though each feature had a formal independence from other features, in that feature specifications and segments are not in a 1-to-1 relationship as they are in SPE-style representations.

One outgrowth of autosegmental theory was that prosodic properties — length, syllabicity and stress — were reconceptualised as hierarchical structural properties in a separate representational dimension, thus outside the domain of feature theory. Accordingly, ‘syllabic’ is the suprasegmental relation of being the ‘peak’ (or non-peak) in a syllable and not a feature at all; length is represented by a separation of segmental content and ‘timing’ where long segments have two timing units and short segments have one; stress is a rhythmic prominence relation between syllables in a word so that one syllable is stronger than another. See Broselow (1995), Blevins (1995), Hayes (1995) for further details. The features for tone, which were never developed in SPE, are a matter of current debate: see Clements (1983), Pulleyblank (1986), Hyman (1993), Bradshaw (1999).

Two recurring ideas which were part of the autosegmental theory resulted in a significant change in feature theory. One was the realisation that while all features are partially independent of others, there are still valid multi-feature units, thus the features for place of articulation can function as a unit in rules. In other words, there is a constituent structure to feature relations. This resulted in the theory of Feature Geometry, discussed in Chapter 10, which has also resulted in the adoption of certain new features, especially for place of articulation.

Underspecification. The second such idea was that not all segments have values for each feature. It was particularly obvious in the case of tone that vowels but not consonants have tone, and it was sometimes difficult to formulate tone rules correctly if consonants get in the way. A similar problem is often observed in vowel harmony, where consonants typically are completely irrelevant to assimilations between vowels, and for that matter vowels are usually irrelevant to (less frequent) long-distance consonantal assimilations. The problem of material intervening between target and trigger was a long-standing one in phonology (see Howard 1972, Jensen 1973, Odden 1977, 1994), and autosegmental representations coupled with Underspecification Theory (Archangeli 1988, Steriade 1995) provided a formal context for saying that consonants do not matter for vowel harmony because consonants and vowels are usually specified with different features. The problem of neutral segments extends well beyond skipping consonants under vowel harmony, because sometimes, languages have so-called “neutral” vowels, where vowel harmony can operate through such vowels. One example is front-back vowel harmony in Finnish, which operates over the vowels *i*, *e*. Explanations of neutral vowels begin from the position that certain segments may be underspecified for particular features, thus in Finnish *i*, *e* can be left unspecified for backness, given that the language does not contrast *i*, *e* and **ɨ*, *ə*.

In addition to being crucial to solving the segment transparency problem, underspecification may explain recurring problems of asymmetrical phonological behavior in segments. It is extremely common, almost universal, that sonorants do not devoice in the way that obstruents do, as seen in numerous final-devoicing and voicing assimilation processes presented in the book and elsewhere; furthermore, sonorants typically do not trigger voicing despite being voiced — see examples in Hungarian in Chapter 4 and Lithuanian in Chapter 5. Similarly pointing to the asymmetry of sonorants in terms of their patterning with other voiced consonants, a voicing rule of Japanese, *Rendaku* (see Ito & Mester 1995), voices a root-initial obstruent in the second member of a compound as long as the compound contains no voiced obstruents — sonorants do not affect the rule one way or the other. By that rule, /*yu-toofu*/ → [*yudoofu*] ‘boiled tofu’ but

there is no voicing in [onna-kotoba] ‘feminine speech’ because of following *b*, and in /te-kami/ → [tegami] ‘letter’, phonetically voiced [m] does not block voicing. These facts can be (partially) explained if sonorants are generally not specified for a value of [voice], since it is not contrastive, thus they would not be phonologically voiced. However, it is important that sonorants eventually be assigned a value for voicing, since sonorants do sometimes trigger voicing assimilation — see Kipsigis (Chapter 3) and Armenian (Chapter 5).

The idea of underspecification has a reasonably ancient pedigree in generative phonology, since the Praguean concept of an “archiphoneme” was incorporated in substance into early generative representations, where segments could have blank values for features. In the theory of Jakobson, Fant & Halle (1951) and Jakobson & Halle (1956), phonemes were considered to be redundancy-free phonetic matrices, having only the invariant phonetic properties specified. In the generative account of Halle (1959), this was extended to the level of morphophonemic (underlying) representations.² In light of the simplicity criterion, it was considered metatheoretically desirable for representations to be economical, specifying only those elements which cannot be predicted, thus redundant (predictable) values could and should be removed, resulting in a simplification of the grammar. A powerful objection to blank values was articulated by Stanley (1967), who shows that their existence is tantamount to giving features ternary specifications, essentially a repudiation of binarity distinctive features. Thus in terms of a simplicity criterion for fully-specified representations, it was proposed that the evaluation should be framed in terms of blanks that “could have resulted” by removing predictable values.

Another consideration arguing against grammar-simplification as a basis for omitting features is presented in Odden (1992). The classical 5-vowel system /i u e o a/ provides an exemplification of the multiple-analysis problem in underspecification. In such a system, [back] can be predicted from [round] or [low]; or, [low] and [back] can be predicted from [round]. But clearly one cannot totally eliminate both [back] and [round] from underlying forms. One feature must serve as the predictor for the values of the other, and that feature must remain specified. Thus some order of simplification is necessary — either [back] is preferentially retained at the expense of [round], or vice versa. The computational modeling study of Odden (1992) shows that there are huge numbers of distinct alternative underspecification schemes for phonemic inventories, for example there are *at minimum* over 1,900,000 different arrangements of fully-underspecified English segment matrices, resulting in an average of 188 irreducible specifications. A miniscule percentage of these underlying systems — .003% — can reach a maximum degree of feature economy in expressing the English segment inventory by making use of just 169 specifications. This indicates that vast numbers of hypotheses need to be considered to reach the most economical underlying representation, and the savings resulting from such a search is about 10%. This raises the question whether a child could ever discover the most parsimonious system of underlying representations.

Autosegmental (non-linear) phonology provided a compelling basis for re-accepting blank values. In autosegmental phonology, it was recognised that not every segment has a tonal specification — only syllabic segments do. It was obvious that consonants should not be speci-

² That this should be so follows from the fact that generative phonology does not recognize a distinctive phonemic level of representation between the underlying and phonetic levels.

fied for tones, since specifying them for tone results in unconscionable complications of tonal processes which would need to cluttered with numerous conditions of the kind “ignoring any tone specifications of consonants”. More important than complication, because of fundamental formal conditions of the theory, it turned out to be impossible to state the most mundane tonal assimilation between vowels, if intervening consonants also have tonal specifications (see the discussion of the No-Crossing Constraint in Chapter 10). Second, the autosegmental theory formally separated segments and tones into autonomous representations, which made it difficult to enforce a requirement that all vowels must bear an underlying tone. The fact of vowels having tones at all is the result of a later process of “association”, which is why it was non-trivial to require all vowels to have an underlying tone. It follows quite naturally in the theory that a vowel might “have” underlying [+H] tone, or might “have” underlying [-H] tone, but it also might have no tone at all. The upshot of this is that three-way specifications were easy to derive, at least for autosegmental tone. It did not hurt one bit that, as we will see in Chapter 10, Margyi actually has a three-way contrast between H roots, L roots, and variable roots. Other examples of actual ternarity, while rare, do put paid to the objection that underspecification leans to ternarity. See Nosske (1996) for evidence that Turkana makes a three-way distinction between [+ATR], [-ATR] and [ØATR], Clements & Sezer for (1982) arguments for a distinction between palatalized velars, regular “blank value” velars, and marked [+back] velars in Turkish, as well as a three-way contrast between front, back and unspecified suffix vowels, and Inkelas (1995) for a three-way contrast in stem-final position between voiced, voiceless and unspecified obstruents in Turkish.

Various proposals were advanced concerning this re-emerging underspecification, in light of the concerns raised by Stanley. Kiparsky (1982) and Archangeli (1984, 1988) formulated a version of underspecification which precluded underlyingly specifying both plus and minus values in the same context, and the Redundancy Rule-Ordering Constraint was proposed which required language-specific or universal fill-in rules to apply before any rules that refer to the “unmarked” value. As the concept of underspecification and feature organisation developed, it became clear that a distinction had to be made between temporary underspecification, where a segment lacks a specification of some feature underlyingly but acquires one in the course of the derivation, versus permanent non-specification, where a segment simply never has any value for some feature. An example of the former would be the situation where /s/ lacks an underlying specification for voicing, perhaps in a language that has no voiced fricatives, but presumably acquires a [-voice] specification at some point. An example of the latter kind of underspecification would be tonal specifications for voiceless obstruents, which never are specified for tone features, or vowels, which never have a specification for [strident]. A variety of permanent non-specification is the possibility of intrinsically unary features, that is features which do not have both plus and minus values to which rules can refer, but rather are present (implicitly, [+X], to which rules can refer) or lacking — a condition to which rules cannot, by hypothesis, directly refer. Such features are said to be monovalent or privative.

The fundamental schism regarding underspecification is whether contrastive feature values can be removed from underlying representations (“no”, according to the theory Contrastive Underspecification — Steriade 1987, Clements 1988), or must some surface value for a given feature always be removed, thus allowing only underlying [+voice] or only [-voice] but never both (“yes”, according to the theory Radical Underspecification — Archangeli 1984, 1988).

Resolution of this controversy depends on answers to a number of other questions. For example, if indeed *labial* is a single-valued feature, then it follows that something resembling Radical Underspecification is definitionally closer to the mark, since only one value — [labial] — can ever be specified either underlyingly or on the surface for that feature. Such examples would constitute examples of (permanent) *non*-specification rather than (temporary) *under*-specification. Quite a number of features were claimed to be privative, under some version of feature geometry, so the question then arises whether the entirety of underspecification can be reduced to privativity and permanent non-specification. The answer to this question turns out to be “no”, first because some (contextually redundant) marked underlying values have to be underspecified and surface specified (preconsonantal nasality in Kikongo: see Ao 1991), and second because even though it is rare to find active evidence for both values of a feature, crosslinguistically, it does happen. Thus both [+lateral] and [-lateral] are referenced in Latin and Georgian, respectively (Odden 1994); yet these feature values are required for [r] and [l] but not ever provably for [p,t,k], so that lateral may be permanently unspecified outside the set [r,l]. While [+voice] is the typically best-attested underlying specification, references to [-voice] are also sometimes required, for example for Dahl’s Law in Kikuria (Odden 1994) and a rule devoicing nasals before voiceless consonants in Toba Batak (Hayes 1986).

Readings

- Allen, W. S. 1953. *Phonetics in ancient India*. London Oriental Series vol. 1. Oxford University Press: London.
- Anderson, S. 1974. *The organization of phonology*. Academic Press: New York.
- Ao, B. 1991. Kikongo nasal harmony and context-sensitive underspecification. *Linguistic Inquiry* 22: 193-196
- Archangeli, D. 1984. “Aspects of Underspecification Theory”. *Phonology* 5: 183-207.
- Archangeli, D. 1988. Aspects of underspecification theory. *Phonology* 5: 183-208.
- Battistella, E. 1990. *Markedness: the evaluative superstructure of language*. SUNY Press: Albany.
- Battistella, E. 1996. *The logic of markedness*. Oxford University Press: New York.
- Blevins, J. 1995. The syllable in phonological theory. J. Goldsmith (ed.) *The handbook of phonological theory*, 206-244. Blackwells: Oxford.
- Bradshaw, M. 1999. *A crosslinguistic study of consonant-tone interaction*. Doctoral dissertation, OSU, Columbus.
- Broselow, E. 1995. Skeletal positions and moras. J. Goldsmith (ed.) *The handbook of phonological theory*, 175-205. Blackwells: Oxford.
- Campbell, L. 1974. Phonological features: problems and proposals. *Language* 50: 52-65.
- Chomsky, N. & M. Halle. 1968. *The sound pattern of English*. Harper and Row: New York.
- Chomsky, N. 1957. *Syntactic structures*. Mouton: The Hague.
- Clements, G. N. 1983. The hierarchical representation of tone features. I. R. Dihoff (ed.) *Current approaches to African linguistics* vol 1, 145-176. Foris: Dordrecht.
- Clements, G. N. 1988. “Towards a substantive theory of feature specification”. in J. Blevins & J. Carter (eds), *Proceedings of NELS 18*: 79-93. Amherst: GLSA.

- Clements, G. N. and Engin Sezer. 1982. "Vowel and consonant disharmony in Turkish." *The Structure of Phonological Representations, part 2*, ed. H. van der Hulst and N. Smith, 213-55. Foris: Dordrecht.
- Clifton, J. 1976. Downdrift and rule ordering. *Studies in African linguistics* 7: 175-194.
- Halle, M. & K. Stevens. 1971. A note on laryngeal features. QPR 101: 198-213. Research Laboratory of Electronics, MIT. Cambridge.
- Halle, M. 1959. *The sound pattern of Russian*.
- Hayes, B. 1986. "Assimilation as spreading in Toba Batak". *Linguistic Inquiry* 17: 467-499.
- Hayes, B. 1995. *Metrical stress theory: principles and case studies*. University of Chicago Press: Chicago.
- Howard, I. 1972. *A directional theory of rule application in phonology*. Doctoral dissertation, MIT. Cambridge.
- Hyman, L. 1973. "The feature [grave] in phonological theory". *Journal of Phonetics* 1: 329-337.
- Hyman, L. 1993. Register tones and tonal geometry. in K. Snider and H. van der Hulst (eds.), *The phonology of tone: the representation of tonal register*, 75-108. Mouton de Gruyter: Berlin.
- Inkelas, Sharon. 1995. "The consequences of optimization for underspecification." ROA 40-1294.
- Ito, J. & R. Armin M. 1995. "Japanese Phonology". J. Goldsmith (ed.) *The handbook of phonological theory*, 817-838. Blackwells: Oxford.
- Jakobson, R. & M. Halle. 1956. *Fundamentals of language*. Mouton: The Hague.
- Jakobson, R., G. Fant & M. Halle 1952. *Preliminaries to speech analysis*. MIT Press: Cambridge.
- Jensen, J. 1973. A constraint on variables in phonology. *Language* 50: 289-330.
- Johnson, R. 1981. *The role of phonetic detail in Coeur d'Alene phonology*. Doctoral dissertation, WSU. Pullman.
- Kean, M. L. 1975. *The theory of markedness in generative grammar*. Doctoral dissertation, MIT. Cambridge.
- Keating, P. 1988. A survey of phonological features. Distributed by IULC, Bloomington.
- Kiparsky, P. 1982. Lexical morphology and phonology. I.-S. Yang (ed) *Linguistics in the morning calm*, 3-91. Seoul: Hanshin.
- McCarthy, J. A Thematic Guide to Optimality Theory. Cambridge: Cambridge University Press.
- Noske, Manuela. 1996. "[ATR] harmony in Turkana". *Studies in African Linguistics* 25: 61-99.
- Odden, D. 1977. Overspecification and variables in phonology. *Linguistic analysis* 3.177-196.
- Odden, D. 1978. Further evidence for the feature [GRAVE]. *Linguistic inquiry* 9: 141-144.
- Odden, D. 1992. "Simplicity of underlying representation as motivation for underspecification." *Papers in Phonology*, 85-100. OSUWPL 41, ed. by E. Hume.
- Odden, D. 1994. Adjacency parameters in phonology. *Language* 70: 289-330.
- Pulleyblank, D. 1986. *Tone in lexical phonology*. Kluwer: Dordrecht.
- Sagey, E. 1986. *The representation of features and relations in nonlinear phonology*. Doctoral dissertation, MIT.
- Stanley, R. 1967. Redundancy rules in phonology. *Language* 43: 393-436.

- Steriade, D. 1987. "Redundant values". A. Bosch et al. (eds) *Papers from the parasession on autosegmental and metrical phonology*, 339-362. Chicago: CLS.
- Steriade, D. 1995. Underspecification and markedness. J. Goldsmith (ed.) *The handbook of phonological theory*, 114-174. Blackwells: Oxford.
- Sweet, H. 1887. *A handbook of phonetics*. London.
- Trubetzkoy, N. S. 1939. *Principles of phonology*. Translated by C. Baltaxe, 1969, University of California Press: Berkeley.
- Vago, R. 1976. "More evidence for the feature [grave]". *Linguistic Inquiry* 7: 671-674.