4 Underspecification and Markedness

DONCA STERIADE

0 Introduction

Not all segments are specified for all features at all times. Theories of underspecification have been proposed which aim to explain the circumstances under which segments might lack feature values, in underlying or derived representations. This chapter surveys the hypotheses about underspecification that have defined the terms of current research.

0.1 The Basic Assumptions

Within the generative phonological tradition initiated by Halle (1959) and Chomsky and Halle (1968), two starting assumptions are generally made:

(1) Lexical Minimality: underlying representations must reduce to some minimum the phonological information used to distinguish lexical items.[1]

(2) Full Specification: the output of the phonological component must contain fully (or at least maximally) specified feature matrices.[2]

One of the functions of the phonological component is then to supply the nondistinctive information missing from the underlying forms. Lexical Minimality requires that the maximal amount of phonological features be left out of the lexical entries, whereas Full Specification dictates that they be present in the input to phonetic interpretation. One way to extract dispensable information from lexical entries is to rely on syntagmatic processes – rules like *Palatalize velar before front vowel* or *Nasalize vocoid after nasal* – which allow us to leave unspecified contextually determined properties like the palatality of velars or vocoid nasality. But syntagmatic processes – the P rules of Stanley

(1967) – cannot be used to rid segments of constant, nonalternating yet predictable features, such as the voicing of sonorants or the continuancy of vowels. This function is then standardly reserved for context-free operations called redundancy rules, which insert the feature values originally left out of lexical entries. The redundancy rules perform functions such as *Mark vowel as [+continuant]* or *Voice sonorant* and thus allow us to eliminate these features too from underlying structures. The study of redundancy rules has been in recent years at the core of research on underspecification. That this class of rules must exist follows, as indicated, from the assumptions of Lexical Minimality and Full Specification. In this sense then, the hypotheses in (1)–(2) are the essence of contemporary theories of underspecification.

It is widely believed that one can observe the effect underspecification has on the phonology of a language by studying the interactions between the P rules and the system of redundancy rules. Why are these interactions revealing? If all redundancy rules precede all P rules, as argued by Stanley (1967), the P rules will apply to fully specified segments and it will be practically impossible to find empirical arguments for or against any individual redundancy rule. The only reason to practice underspecification, in that case, will be to uphold one's belief in Lexical Minimality. Most phonologists would agree, however, that the idea of underspecification can be empirically supported. For instance, in arguing to include *Voice sonorants* in the redundancy rule list, one cites phonological processes that ignore the phonetic voicing of sonorants. Russian, for instance, has a voicing assimilation triggered by voiced obstruents but not by sonorants (cf. Kiparsky 1985). The fact that sonorants are inactive in this process would be explained by letting the Russian voicing assimilation apply before the redundancy rule *Voice sonorants* takes effect. Prior to this redundancy rule, the sonorants are, in accordance with Lexical Minimality, lacking any [voice] values. A schematic derivation illustrating this scenario appears below.

(3) *Voice sonorants* in a language where sonorants do not act as [+voice] segments

THE ORDERED GRAMMATICAL STATEMENTS	THEIR CONSEQUENCES FOR SONORANTS
(i) Underlying representations: according to (1)	sonorants lack [+voice]
(ii) Assimilation: x → [+voice] / ___ [+voice]	fails to apply before sonorants
(iii) Redundancy rule: [+sonorant] → [+voice]	sonorants become voiced

Much of the evidence considered in the underspecification literature is of the type sketched in (3) and has been analyzed in the past in terms of three-step derivations comparable to the one above. Cross-linguistic differences in the effect similar rules have on segments have been attributed to the possibility of switching steps (ii) and (iii) in (3). For instance we may compare Russian with English, where the voicing assimilation targetting inflectional suffixes

like *-s* and *-t* is induced by voiced obstruents and sonorants alike (cf. *bug[z]* and *call[z]* with *chick[s]*). English could be analyzed, without sacrificing Lexical Minimality, by letting step (iii) *Voice sonorants* precede step (ii). By shifting the rule order, we maintain that sonorants lack underlying [voice] values, regardless of phonological patterning.

Because of its reliance on Lexical Minimality and Full Specification, most of the literature on underspecification consists of sequential analyses like (3) which, taken together, read like an advertisement for sequential rule application, rule ordering and, more generally, for a derivational as against a declarative approach to phonology.[3] Even authors who claim that the ordering between phonological rules and redundancy rules is *not* extrinsic – cf. Archangeli (1984), Archangeli and Pulleyblank (1986) – are led to justify the hypothesis of underspecification by formulating derivational analyses in the style of (3). Although the question of derivational vs. declarative phonology cannot be considered directly here, we do have to ask whether the phenomena attributed to underspecification must be analyzed derivationally, by distinguishing an earlier, less-specified stage of the derivation from a more fully specified, later stage. This will turn out to be the key issue.

0.2 *The Writer's Prejudices*

The present chapter has been influenced by the early work on underspecification of Richard Stanley (1967), as well as by Mohanan (1991) and the more recent literature on declarative and harmonic phonology. Stanley's article on redundancy rules was in part a reaction against what I would like to call "opportunistic uses" of underspecification. Some grammatical statements are made simpler if they are assumed to hold of incompletely specified representations: the assimilation rule in (3ii), for instance, might be written as $x \rightarrow$ [*+voice*] / ____ [*+voice, –sonorant*] but it is opportune to simplify this statement by omitting [–sonorant]. We can do so if we rely on the incompletely specified lexical entries assumed in (3i). Stanley thought that the invocation of language-specific convenience is a bad reason to practice underspecification. One hopes, with Stanley, that any discrepancies in feature specification between lexical and surface structure follow from general principles, not descriptive convenience. Lexical Minimality is such a principle but it seems, as we shall see, indefensible insofar as it can be made precise. The search for other principles from which more interesting varieties of underspecification follow has also proved, in my view, fruitless. Sections 2.3.1–2.3.2 develop these points. What remains to be done is to provide plausible alternatives for all analyses relying on derivational scenarios like (3). I cannot undertake here a complete re-evaluation of the evidence, but an initial attempt is made at the end of section 2 and in section 3. In any event, the reader will need to bear in mind, in reading what follows, where my sympathies lie.

0.3 Trivial and Nontrivial Underspecification

I limit the scope of this survey by considering only the evidence for nontrivial – or temporary – underspecification. We will notice repeatedly that there exists abundant evidence for permanent – trivial or inherent – underspecification.[4] Some segments do not carry specifications for certain features, either underlyingly or at any subsequent derivational stage. Thus plain coronals are trivially, inherently, and permanently lacking in specifications for the features [labial] or [tongue root]. Similarly, if [nasal] is a privative feature, oral segments will permanently lack [nasal] values. The extent of their underspecification for [nasal] is settled once and for all when we decide that [nasal] has only one value; having done so, we do not have to wonder when or how oral segments become specified as [–nasal]. Although there are important issues involved in the study of permanent underspecification, they have to do more with the relation between phonology and phonetics than with the question that I consider central to this survey: assessing the validity of derivational scenarios like (3), which invoke *temporary* underspecification. Consequently, I focus here on the evidence for lexical representations from which some *surface-present* feature has been left out.

0.4 Outline

The survey begins by outlining the situations in which features can be said to be predictable, in a syntagmatic or segment-internal context (section 1). I review next theories of underspecification that share the assumptions of Lexical Minimality and Full Specification (section 2) . In the course of the review we will observe that neither Lexical Minimality nor Full Specification can be defended when a closer look is taken at what these ideas entail for the organization of the phonology and for the nature of phonetic representations. The last sections (2.3.3.5 and 3) sketch an alternative, nonderivational view of phonological underspecification.

1 The Facts of Underspecification: Predictability and Inertness

The features of segments are frequently predictable. Sometimes a predictable feature value fails to manifest its presence in a phonological process where it might otherwise be expected to act. In such cases, the practice has been to declare it unspecified and set up derivations similar to (3).[5] I review in this section several classes of feature predictability that can thus be linked to the hypothesis of underspecification. The varieties of redundancy rules corresponding

to these types of predictability were first formulated by Halle (1959) as *Segment Structure* and *Sequence Structure* rules.

1.1 Sources of Predictability: Markedness and Neutralization

1.1.1 Feature Co-occurrence and Context-sensitive Markedness

Feature co-occurrence conditions are formulated (since Halle 1959) to express restrictions on the possible feature combinations within a segment. For instance, the absence of distinctively voiceless sonorants in English may be expressed as a rule (the *Voice sonorants* process mentioned above) or as a filter *[+sonorant, –voice].[6] The consequence of adopting either one is that some feature values become predictable: if, for instance, *[+sonorant, –voice] is part of the grammar, it follows that sonorants are predictably specified as [+voice]. From such facts, Lexical Minimality derives underspecification: in this case, sonorants lacking [voice].

The vast majority of feature cooccurrence statements have some cross-linguistic validity, in the sense that they characterize the unmarked combination of articulatory gestures for the relevant segmental class. Thus, a statement like *Sonorants are voiced* is justified not only internally to English but also cross-linguistically, as a markedness statement: it characterizes the normal state of the glottis in the sonorants of all languages. The assumption is frequently made (e.g., Archangeli 1984, 1988; Kiparsky 1985; Calabrese 1987; Archangeli and Pulleyblank 1992) that cross-linguistically valid feature co-occurrence conditions are part of Universal Grammar and do not contribute to the complexity of individual linguistic systems. If so, any language learner can be relied upon to know that *Sonorants are voiced,* in advance of exposure to data, and to draw from this principle the conclusion that the voicing of sonorants need not be lexically encoded. One of the tasks of markedness theory is to document the validity of such universal statements of feature co-occurrence and seek an explanation for their universal status. We return to the issue of markedness in section 2.3.3.[7]

It must be noted, however, that any co-occurrence condition engaging two features will derive two distinct and sometimes incompatible patterns of underspecification. For instance, *[+sonorant, –voice] renders predictable not only the voicing of sonorants but also the sonority of voiceless segments. Must we then leave the [–sonorant] value of *p, t, k* out of the underlying representations? There seems to be little evidence for such a move and we need to ask why, especially as there exists substantial evidence for leaving the sonorants unspecified for [voice] (Kiparsky 1985; Ito and Mester 1986). In a different case, that of the [round] /[back] relation in triangular vowel inventories like {a, e, i, o, u}, we observe that either [round] can be left out of underlying representations or [back], since a statement like *[αback, –αround] allows us

to derive the values of one feature from those of the other: but we cannot eliminate both. We must therefore, in accordance with Lexical Minimality, choose one feature as basic and derive the other. But the filter does not establish an asymmetry between the features involved in it: the values of back and round are mutually predictable, not arranged in some obvious pattern of ordered dependency. The choice of either [round] or [back] as the underlying value is not dictated by the co-occurrence condition itself and thus remains arbitrary. It appears that any grammar that fully complies with Lexical Minimality has to contain a number of arbitrary decisions of this sort.[8]

1.1.2 Context-free Markedness

Cross-linguistic asymmetries can be observed not only in the distribution of feature combinations but also in that of individual feature values. For instance, nasal segments have, in the vast majority of cases, a more limited distribution than corresponding oral ones; any aspirated or glottalized segment is more limited in occurrence than its nonaspirated, nonglottalized counterpart. Observations of this sort are also the province of markedness theory. A possible connection they have to underspecification has been formulated by Chomsky and Halle (1968, chap. 9), Kiparsky (1981, 1985, 1988), and assumed by others: Universal Grammar provides every feature with a marked and an unmarked value. The unmarked value of any feature corresponds to the normal, neutral state of the relevant articulator. For [nasal] the unmarked value is [–nasal], for [spread glottis] it is [–spread glottis] and so forth. Only one value for any given feature need be present underlyingly. Since features are assumed to be strictly binary, the other value can always be predicted by a context-free rule that mirrors the relevant markedness statement. Thus, corresponding to the statement that *Segments are normally oral,* we can have the universal rule [ ] → [–nasal]. The intention is to let this rule, like other redundancy rules, insert [–nasal] only in segments lacking a [nasal] value. A lexical contrast between *b* and *m* can then be represented as follows:

(4) A context-free redundancy rule: *[ ] → [-nasal]*

underlying:	b	vs.	m	
			[+nasal]	
surface:	b	vs.	m	
	[–nasal]		[+nasal]	

The use of markedness-based context-free redundancy rules of this sort appears to express the derivational transition between an underlying system in which all features are privative and a surface system in which all features are binary.[9] We may question the justification for this class of redundancy rules in two ways. First, we may ask whether the asymmetric distribution between the marked and the unmarked value justifies eliminating the unmarked value from underlying structure. Is underlying privativity the faithful

representation of markedness facts? Second and more important is the need to question the assumption of surface binarity: should the unmarked value be represented at all on the surface? The second issue will be addressed below, in section 2.3.3. We can anticipate here the conclusion that the facts of context-free markedness do not in fact motivate any *universal* redundancy rules of the type in (4).

1.1.3 Syntagmatic Predictability of Features: Positional Neutralization[10]

A third variety of feature predictability involves not the markedness of segment-internal feature combinations or individual feature values, but the neutralization of featural contrasts in certain positions. This case has received little attention since Trubetzkoy (1939); and its particulars will be documented more carefully in section 3.

Segments identifiable as marked may, if allowed at all in a given language, be restricted to certain salient positions within the word: the syllable peak (rather than the margin), the onset (rather than the coda), the stem (rather than the affix), the stressed syllable, or the edges of the word. Maidu (Shipley 1956) allows laryngeally-specified consonants (ejectives or implosives) only in the syllable onset. Guaraní (Kiparsky 1985 and references there) allows underlying nasal vowels, but only in stressed syllables. Ancient Greek allowed aspirated vowels word-initially but not elsewhere. Copala Trique, an Otomanguean language, allows a wide range of segmental distinctions to surface only in the last syllable, which may be the one carrying stress: lenis/fortis contrasts, tonal distinctions, the laryngeals ʔ and h, and nasalized vowels occur only in the final syllable (Hollenbach 1977). Chumash (Applegate 1971) stem vowels are drawn from a crowded inventory which contains {a, e, o, i, i̤, u}; affixal vowels, however, are underlyingly limited to the peripheral set {a, i, u}. In languages like Bashkir (Poppe 1962) and Vogul (Kálmán 1965), round vowels are restricted to initial position in both underlying and derived representations.

In all these cases, the absence of a contrast renders predictable whatever phonetic value the relevant feature might take on in the position of neutralization. Thus Bashkir and Vogul non-initial vowels are predictably unrounded; Guaraní stressless vowels are predictably oral; Chumash affixal vowels are predictably peripheral. The predictable value frequently coincides with the one normally identified as "universally unmarked," although in cases like Chumash such identification is difficult to implement. We retain, however, the need for grammatical statements which express the impossibility of a featural contrast in a given position. Such statements may take the form of filters, as in (5) (an example of which is Ito's (1986) Coda Condition) or of positive licensing conditions, as in (6) (cf. Goldsmith 1990, p. 123ff.).

(5) Positional neutralization: filter version
*αF in x where x is defined prosodically or morphologically.

(6) Positional neutralization: licensing version
αF must be licensed in $\bar{x}$, where $\bar{x}$ is defined prosodically or morphologically.

The presence of statements like (5) or (6) in a grammar induces predictability: no value for F is necessary in x, the unlicensed position, since only one value is allowed. If we adopt Lexical Minimality, underspecification follows as well: no value for F is possible in x. We return to the formalization of conditions like (5)–(6) in section 3, where their consequences for underspecification are explored in detail.

1.2 *Varieties of Phonological Inertness*

The purpose of the preceding sections was to inventory the circumstances under which a feature value might be predictable and hence, *potentially* unspecified. (I stress *potentially* because underspecification does not directly follow from predictability. It follows only if we subscribe to some further principle such as Lexical Minimality.) How do we determine, then, independently of the merits of Lexical Minimality, that a predictable feature is *actually* unspecified? How is the existence of underspecification diagnosed by phonologists? In what follows I will try to make explicit the reasoning that might lead to the adoption of underspecification.

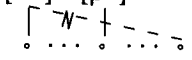
Phonologists determine the presence of underspecified representations by observing facts which appear to contradict three basic expectations: that rules will apply locally, with maximal generality, and that lexically specified features will change only in response to clearly defined considerations of well-formedness. Consider first the hypothesis of local rule application:

(7) Locality: Phonological rules apply between elements adjacent on some tier.

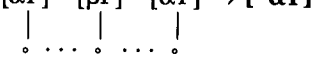
Assimilations and dissimilations are expected, according to this statement, to involve prosodic positions, anchoring nodes, or terminal features that are adjacent. The formulation of Locality given above is deliberately vague, since phonologists disagree on the analysis of several apparently nonlocal phenomena.[11] Opinions converge, however, on the only aspect of Locality that is relevant to our concerns: no phonological rule is expected to "skip" specified features on the tier involved in assimilation or dissimilation. The operations diagrammed below are considered impossible, because they violate all proposed versions of Locality:

(8) Where skipping is impossible

Assimilation: *Spread [αF]*

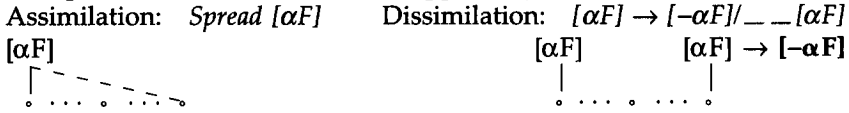


Dissimilation: *[αF] → [−αF]/_ _ [αF]*



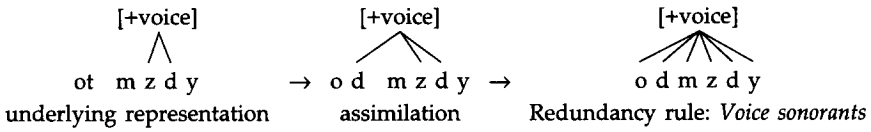
There exist, however, processes which do have the appearance of these prohibited operations. Such cases have been analyzed in the past by distinguishing an earlier unspecified stage (shown below) where the operation of the rules shown can be viewed as local, and a later stage, during which some intervening segment acquires, through redundancy rules, the [β F] value.

(9) Underspecified structures behind an apparently non-local rule application



An illustration of (9) is the Russian rule of voicing assimilation mentioned earlier. It is not only the case that Russian sonorants fail to trigger voicing (*ot melodii* "from the melody" not **od melodii*) but also that Russian consonantal sonorants may intervene between the obstruent triggering voicing assimilation and the obstruent undergoing it: *iz Mtsenska* "from Mtsensk" becomes *is Mtsenska*, while *ot mzdy* "from the bribe" becomes *od mzdy*. The derivation in (10) reconciles these facts with our belief in Locality:

(10)



Consider next the hypothesis of generality in rule application: all other things being equal, one expects that rules which spread, dissimilate, or are otherwise conditioned by [αF] will apply to all segments possessing [αF].

(11) Generality: If some process manipulates [αF], then all segments possessing [αF] will participate in it.

There are many phenomena which appear to contradict this expectation. Most, if not all, could be understood as stemming from underspecification. An example has already been provided in the analysis of Russian voicing: voiced sonorants do not trigger the rule (recall *ot melodii* and *is Mtsenska*) because they lack any [voice] value when the rule applies (Kiparsky 1985). A distinct example is that of Lamba height assimilation. This process, widespread among Bantu languages, turns a high suffixal vowel into a mid one when preceded by a mid vowel. The spreading feature appears to be [–high]. But low vowels – which are necessarily [–high], since they cannot be otherwise – do not trigger this rule. The facts, as recorded in Kenstowicz and Kisseberth (1977, p. 72), appear below:

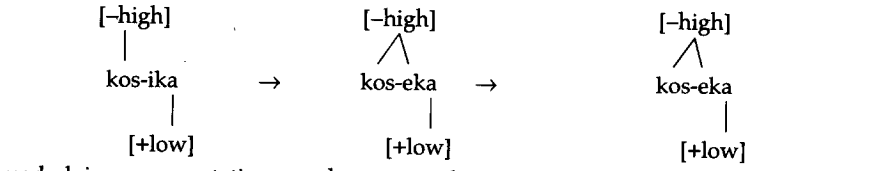
(12)

Past	Neuter	Applied	Gloss
tul-a	tul-ika	tul-ila	"dig"

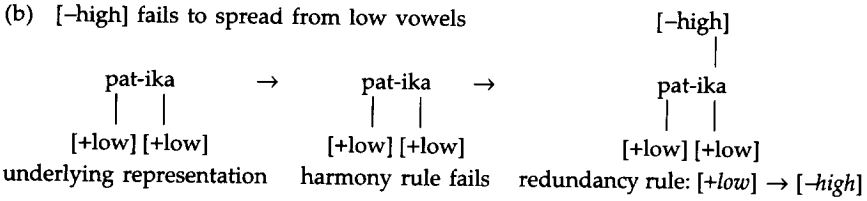
fis-a	fiš-ika	fiš-ila	"hide"
kos-a	kos-eka	kos-ela	"be strong"
sek-a	sek-eka	sek-ela	"laugh at"
pat-a	pat-ika	pat-ila	"scold"

An analysis of this data which appeals to underspecification in order to preserve Generality has been proposed in Steriade (1987b): low vowels cannot be [+high] for obvious articulatory reasons. Therefore their [high] values are predictable. Since predictable, they are unspecified: they have no height value to spread. In contrast, non-low vowels – *i, u, e, o* – have distinctive height values: no feature-co-occurence statement can predict whether a non-low vowel will turn out as [+high] or [–high]. In particular, mid vowels must be specified as [–high], which is the spreading feature in this case. Two three-step derivations illustrating this analysis appear below. Although this is immaterial to the argument, I ignore the [+high] values which high vowels might posess.

(13) (a) [–high] spreads from mid vowels



(b) [–high] fails to spread from low vowels



A final assumption from which one may draw conclusions about underspecification is the idea of invariance. As mentioned earlier, there is considerable consensus that assimilation affects mostly segments lacking values for the spreading feature. The failure of specified segments to assimilate could be attributed to Invariance:[12]

(14) Invariance: Lexically specified features prefer to remain unchanged.

I state (14) as a preference, since Invariance is not upheld at all costs: dissimilation processes like that in (9) switch or delete a feature value under the compulsion of more urgent constraints, such as the Obligatory Contour Principle.[13] Where segments are observed to undergo assimilation, this fact is made compatible with Invariance by assuming underspecification. Most such cases will be discussed in section 3, where we will conclude that the facts do not necessarily lead to the conclusion that assimilation targets are underspecified.

2 Derivational Theories of Underspecification: "Radical" or "Contrastive"?

We have reviewed so far the forms of reasoning that lead phonologists to postulate underspecified representations. Considered next are a number of ideas about the nature and ordering of redundancy rules that have been explored during the last decade. All theories articulated so far can be generally referred to as *derivational*, since they all rely on the possibility of ordering redundancy rules before or after P rules. The effect of underspecification can be determined, according to these theories, only when the rule order is settled. The first two underspecification systems to be reviewed have been known as radical underspecification.[14] The term "radical" refers perhaps to the fact that these theories uphold Lexical Minimality, by eliminating from underlying representations both feature values predictable from co-occurrence conditions and those predictable from context-free markedness statements. A distinct line of thinking, contrastive or restricted underspecification, limits the degree of underspecification in lexical forms by eliminating only feature values predictable on the basis of feature co-occurrence.

2.1 *Radical Underspecification within Lexical Phonology: Kiparsky (1981, 1985)*

At the source of all current work on underspecification lie Paul Kiparsky's ideas on lexical phonology and their applications to the principles of Locality and Invariance identified earlier. It is Kiparsky who first pointed out (in Kiparsky 1981) that apparent violations of Locality can be interpreted as stemming from the possibility of skipping segments which are predictably specified for a relevant feature. The connection between underspecification and Invariance was also made there. In a later study (1985), Kiparsky presents the hypothesis that the operation of different kinds of redundancy rules is tied to specific derivational levels. In a nutshell, the claim is that nondistinctive values (those predictable by the co-occurrence filters discussed in sec. 1.1.1) will occur only postlexically. We examine now the specifics of the model incorporating this idea.

Kiparsky assumes, following Chomsky and Halle (1968), a theory of markedness which provides a universal list of marked feature combinations and marked feature values.[15] Corresponding to the marked combinations there are filters similar to the *[+sonorant, –voice] condition discussed earlier. It is implied that, although universal, the filters are violable; a segment inventory will then be defined by the set of filters violated.[16] Corresponding to the marked feature values there are redundancy rule applications, such as [ ] → [–nasal], which insert unmarked specifications. Underlying representations obey Lexical Minimality: they lack features insertable by redundancy rules.

The distinguishing aspect of Kiparsky's proposal is the principle of Structure Preservation. It dictates that lexical rule applications, those taking place within the lexical rather than postlexical component of the phonology, will not have outputs violating the filters obeyed in the underlying representations of the language. A possible formulation is (15):

(15) Structure Preservation: No lexical rule application will generate structures prohibited underlyingly.

P rules as well as redundancy rules are subject to Structure Preservation. The result aimed at is that of obtaining a limited degree of underspecification in the lexical component, by barring the lexical application of the redundancy rules linked to feature co-occurrence conditions. To understand how the system works, consider again the case of Russian voicing. There are two relevant markedness facts in this case. One is the fact that [–voice] is the unmarked value in obstruents; the other is the filter prohibiting [–voice] in sonorants. This filter is formulated as (16) by Kiparsky; it states that *no value* for [voice] is well formed in sonorants. This filter is accompanied by a unique redundancy rule, (17):

(16) $*\begin{bmatrix} \alpha \text{ voice} \\ +\text{ son} \end{bmatrix}$

(17) [α sonorant] → [α voice]

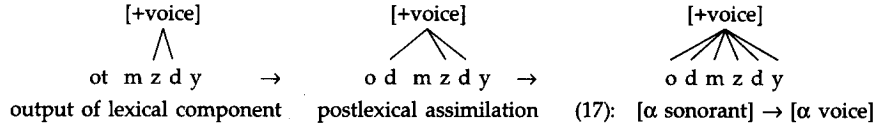
The consequence of combining (16) with Structure Preservation is that no lexical rule applications will insert either [+voice] or [–voice] in sonorants. Thus (17) will be unable to mark sonorants as [+voice] within the lexical component. It may however apply to obstruents, marking them [–voice], since no value for voicing is prohibited in obstruents. Postlexically, (17) will apply to sonorants as well. The resulting system predicts the following generalizations:

(18) (a) No lexical rule application will encounter – or generate – a [voice] value in sonorants.
(b) Obstruents will emerge fully specified, [+voice] or [–voice], from the lexical component.
(c) Sonorants will be specified for [voice] at some point in the postlexical component.

Are the facts of Russian consistent with these predictions? This is not entirely clear. As noted earlier, the sonorants do not undergo, trigger or block voicing assimilation. Now, Kiparsky claims that voicing assimilation applies both lexically and postlexically in Russian: the sonorants, however, fail to trigger, block or undergo the rule in *either* component. This is not exactly what (18) predicts: (18c) leads us to expect that the voicing of sonorants will manifest itself postlexically. The facts can be made to fit the model only if an additional ordering condition is imposed: in every component, the voicing

redundancy rule (17) follows the rule of voicing assimilation. The derivation in (19), modified below, illustrates the postlexical order.[17]

(19)



output of lexical component | postlexical assimilation | (17): [α sonorant] → [α voice]

Given that the order between Assimilation and (17) must be stipulated, we should ask what empirical considerations support the assumptions in (15)–(16). Could we handle all the Russian facts by simply ordering Assimilation before (17), without appeal to (16) and Structure Preservation? The answer is not clear and the issue cannot be pursued here.[18] The major point to emerge, however, is that Structure Preservation contributes minimally, if at all, to the account of the Russian facts. The main element in the analysis is the extrinsic ordering between assimilation and redundancy rules.[19]

Similar analyses are applied in the 1985 article to a number of other phenomena in which underspecification appears to play a role. We will briefly consider here the analysis of Catalan consonantal place features, based on the original account of Mascaró (1976). The significant fact of Catalan is the contrast between alveolar coronals and all other consonants in triggering and undergoing place assimilation. The contrast is illustrated below: only the alveolars undergo place assimilation; and only the nonalveolars trigger it. Some details of the paradigm are omitted here.

(20)

Underlying alveolar in coda		Underlying labial in coda		Other codas	
son amics	"they are friends"	som amics	"we are friends"	tiŋ pa	"I have bread"
som pocs	"they are few"	som pocs	"we are few"	aŋ feliç	"happy year"
son dos	"they are two"	som dos	"we are two"		
soŋ grans	"they are big"	som grans	"we are big"		

The first task is to explain the difference between the alveolar point of articulation and the others. The explanation proposed by Kiparsky is that a marking condition, (21) below, prohibits the lexical appearance of [+coronal]. From this filter, we deduce a corresponding redundancy rule, (22):

(21) *[+coronal]

(22) [ ] → [+coronal]

The consequence of (21) is that the coronals are placeless in underlying representations, as well as throughout the lexical component. Structure Preservation

ensures that (22) will apply only postlexically. The intention is to let the place assimilation take effect before (22). If so, we can explain both facts illustrated by (20). Only coronals undergo the assimilation rule because only they are place-unspecified: the principle of Invariance (14) applies here. Coronals fail to trigger the rule for the same reason. We note, however, that the Catalan analysis too relies on an ordering relation unrelated to Structure Preservation, or any other principle: within the postlexical component, (22) must *follow* place assimilation. The opposite order would also be consistent with Kiparsky's theory, but does not generate the data. If the alveolars became specified *before* the postlexical place assimilation, there would be no representational difference between them and the other Catalan consonants. Here too, the actual contribution of Structure Preservation is hard to pinpoint;[20] the bulk of the paradigm is handled by old-fashioned rule ordering.

There are, then, two elements in the analyses of the 1985 study: a feature co-occurrence condition (e.g., (21) or (16)) and an ordering statement which places a redundancy rule *after* some assimilation rule. We have suggested that the ordering itself cannot be derived from higher principles and, although central to each analysis, remains unconnected to Structure Preservation. We examine next the form taken by Kiparsky's feature co-occurence conditions, which are used to predict lexical underspecification. The discussion of this point owes much to Mohanan (1991).

The initial attractiveness of filters like (16) is that some statement is needed, independently of the issue of underspecification, to characterize the Russian segment inventory: any description of Russian must note the absence of voiceless sonorants. Filter (16) is presented as recording this fact. In fact, however, it does not: (16) prohibits not only [+sonorant, –voice] – the combination that is provably absent from underlying structure – but also [+sonorant, +voice], a combination that is, naïvely speaking, well formed even if nondistinctive. In the discussion of feature co-occurrence (section 1.1.1), we noted that a filter like *[+sonorant, –voice] renders predictable, and hence underspecifiable, the [+voice] value in sonorants. Kiparsky, however, is trying to derive a stronger result: the aim is not only to eliminate [+voice] from the underlying representation of sonorants but also to keep it out of the sonorants for the entire lexical component. Whether or not this is the right idea, it clearly cannot be implemented by using the well motivated filter *[+sonorant, –voice]: this filter will not prevent the redundancy rule in (17) from generating voiced sonorants in the lexicon. The hypothesis of lexical underspecification can be implemented only by adopting the extended filter (16), *[+sonorant, α voice]. In this extended form, however, the filter is motivated only by the drive to generate an underspecified lexical component, not by independent considerations. To put it plainly, the argument for lexical underspecification in Russian based on (16) is circular. This becomes fully obvious when we consider Kiparsky's observation that the sonorants are subject postlexically to [–voice] spreading (from a preceding obstruent, as in i[s m̥ts]enska "from Mcensk") but that they undergo this rule *gradiently*. Kiparsky suggests that gradient application is a possible symptom of postlexical rules. Note, however, that the application of

the voicing redundancy rule (17) to sonorants is equally a postlexical process, on Kiparsky's analysis, in virtue of Structure Preservation: yet the [+voice] value normally appearing in the Russian sonorants is anything but gradient. The right conclusion to draw from this is that gradient applications mark only violations of a lexical filter, and the right lexical filter is the well-motivated *[+sonorant, –voice] not *[+sonorant, α voice].

A different criticism must be leveled at (21), the filter barring [+coronal] from the lexical component of Catalan. A statement of markedness might have to note the fact that alveolars are more common than other consonantal points of articulation: (21) appears to derive its force from this consideration. But (21), i.e., *[+coronal], does not express any observable property of Catalan consonants, nor any clearly defined markedness fact: coronals are neither impossible in Catalan nor undesirable universally. If the intended general principle is to rule out unmarked values from the lexical component, not just underlying representations, then this principle conflicts with Kiparsky's analysis of Russian, where [–voice] obstruents appear lexically, and with his analysis of Guaraní, where stressed oral vowels are lexically represented as [–nasal]. (On Guaraní, see below section 3.) Once again, we observe that the filter from which lexical underspecification is meant to follow is only supported by the need to uphold lexical underspecification in individual analyses.[21]

Aside from this criticism of the specifics, we must note that Kiparsky is seeking to address *the* essential question in the theory of derivational underspecification: that of predicting the degree of segmental specification at a given derivational level and hence of limiting the interactions between phonological rules and redundancy rules. The intuition expressed in the 1985 study is that of an orderly progression from maximally underspecified lexical entries to fully specified surface structures.[22] This progression comes in two major blocks of redundancy rule applications. First are inserted the feature values corresponding to context-free markedness preferences (such as *[–voice] is unmarked in obstruents*) and then, only postlexically, come the feature values corresponding to filters (such as *only [+voice] is possible in sonorants*). We will see in the next sections that later hypotheses about underspecification represent divergent developments of Kiparsky's views.

2.2 *Radical Underspecification Outside of Lexical Phonology: Archangeli (1984, 1988)*

2.2.1 *Ordering Redundancy Rules and Phonological Rules: The Redundancy Rule Ordering Constraint (RROC)*

The project of predicting the interactions between redundancy rules and phonological rules has been continued by Archangeli (1984, 1988), Pulleyblank (1986, 1988a, 1988b) and Archangeli and Pulleyblank (1986, 1989), within the general framework defined by Lexical Minimality. These writers seek to

strenghten Kiparsky's proposals by disallowing any extrinsic rule ordering between phonological rules and redundancy rules. According to Archangeli and Pulleyblank, all ordering matters between these two types of rules are settled by one principle, the Redundancy Rule Ordering Constraint (RROC). The intended effect of the Redundancy Rule Ordering Constraint is to apply all redundancy rules inserting [αF] before any phonological rule mentioning [αF].

(23) The Redundancy Rule Ordering Constraint (RROC) (Abaglo and Archangeli 1989, p. 474)
A redundancy rule inserting [α F] is assigned to the same component as the first rule referring to [αF].

The formulation in (23) assumes that, within any component of the phonology, the applicable redundancy rules apply as anywhere rules: that is, they precede all phonological rules and continue to apply wherever new eligible inputs are created. To understand the functioning of the Redundancy Rule Ordering Constraint in a grammar, we consider an example from Yoruba provided by Archangeli and Pulleyblank (1989). In this language, the feature [–ATR] spreads leftward from a low vowel *a*, which is noncontrastively specified as [–ATR], as well as from the mid [–ATR] vowels ɛ and ɔ, whose tongue root position is distinctive. The phonemic vowel system of Yoruba is {a, ɛ, ɔ, e, o, i, u}. There are excellent reasons, outlined by Archangeli and Pulleyblank (1989, pp. 184–187), to believe that the harmony system of Yoruba is asymmetric: the vowels {a, ɛ, ɔ} are triggers, while {e, o} are undergoers and {i, u} neutral segments. One reflex of this is the fact that [–ATR] spreads leftward across compound boundaries, while [+ATR] does not (data from Archangeli and Pulleyblank 1989, pp. 189–190; tones omitted):

(24) ogbo "old" + ɛni "person" → ɔgbɛni "sir"
ogũ "twenty" + ɛta "three" → ɔgɔta "sixty"
ɔkɔ "husband" + olobĩrĩ "married man" → ɔkɔlobĩrĩ "married man"

This pattern is interpreted as indicating that only [–ATR] values, the dominant ones here, exist in underlying representations. Lexical Minimality requires us to eliminate the [–ATR] feature of *a*, because it can be predicted from the feature co-occurrence *[+low, +ATR]. A further principle, discussed below, dictates that the inert [+ATR] values of {e, o, i, u} be inserted by the context-free redundancy rule [ ] → [+ATR]. This generates the underlying array of height/ATR values given below. I represent them accompanied by the corresponding redundancy rules (from Pulleyblank 1988a, p. 238).

(25)

	a	ɛ	ɔ	e	o	i	u
ATR		–	–				
low	+						
high		–	–	–	–		

(i) [+low] → [–ATR]
(ii) [ ] → [+ATR]
(iii) [ ] → [+high]
(iv) [+low] → [–high]

Within this system, we must assume that [–ATR] spreads *after* the redundancy rule in (i) – since the non-contrastively retracted *a* does trigger [–ATR] harmony – but *before* the redundancy rule in (ii) – since the [+ATR] vowels undergo and fail to trigger harmony. (A tacit appeal to the principles of Generality and Invariance is being made here.) The Redundancy Rule Ordering Constraint predicts exactly this desired order, if we assume further that harmony belongs to the lexical component of the phonology. In virtue of the RROC, [–ATR] harmony and redundancy rule (i) must belong to the same component. Moreover, we know that, within a given component, redundancy rules are the first to apply. These assumptions derive the order *(i) ≪ [–ATR] harmony*. The low vowel *a* will now be able to propagate [–ATR]. In addition, we assume that no Yoruba lexical rule mentions [+ATR], and the RROC will therefore fail to place rule (ii) within the lexical component. Its application must be assumed to come postlexically. It thus follows that the order must be *[–ATR] harmony ≪ (ii)*: the high and mid tense vowels {i, u, e, o} will correctly lack any [ATR] values when [–ATR] harmony applies. The sequence of rule applications appears below, along with the principles motivating each ordering statement:

(26)

	Order	Motivating principles
lexical stratum	[+low] → [–ATR]	RROC: [–ATR] mentioned in lexicon.
	[–ATR] harmony	P rules follow first application of redundancy rules in any given stratum.
postlexical stratum	[ ] → [+ATR]	RROC: [+ATR] not mentioned lexically.

But the Redundancy Rule Ordering Constraint, while successful in Yoruba, is inconsistent with the Russian voicing facts. When we compare assimilated forms such as *od mzdy* (from /ot mzdy/) to unassimilated *ot melodii* (**od melodii*) we must conclude (a) that [+voice] spreads postlexically and (b) that it spreads only from obstruents, with the sonorants behaving as voice-neutral. The RROC will predict that any P rule mentioning either [+voice] or [α voice] will be sufficient to trigger the prior application of the redundancy rule specifying the sonorants as voiced; but this means that the sonorants will become voiced *before* voicing assimilation, and thus able to trigger and block the rule. In a different context, Archangeli (1988, p. 199) suggests that problems of this general class may be resolved by appeal to Structure Preservation. Perhaps Structure Preservation takes precedence over the RROC in prohibiting the insertion of the nondistinctive [+voice] value in sonorants. Russian voicing is not the example considered by Archangeli, but we can use it to illustrate the pitfalls of this proposal. First, the Russian sonorants must be voice-unspecified *postlexically*, as indicated by the comparison between the two phrases *ot melodii* and *od*

mzdy. Structure Preservation is surely not responsible for this. Second, if Structure Preservation applies to block the insertion of noncontrastive values in Russian, then it should have the same effect in the case of Yoruba ATR: the predictably retracted *a* should be unable to trigger [–ATR] harmony, at least within the Yoruba lexical component. This is not the case. No account based on the RROC can be combined with the hypothesis of Structure Preservation to derive both the Russian and the Yoruba patterns of underspecification.

2.2.2 *Markedness Reversals*

A second theme of the underspecification model developed by Archangeli is the idea of parametrizing markedness. Recall that the analysis of Yoruba ATR harmony requires reference to [–ATR] as the active value. In other systems [+ATR] must be assumed to be active;[23] these include the ATR harmonies of Akan (Clements 1978), Igbo (Ringen 1979), and others. The version of markedness theory espoused by Kiparsky (1985) provides for one and the same value in every feature being designated as universally marked. This marked value is the only one allowed underlyingly, in all languages, but the problem is that we need this value to be [–ATR] in Yoruba and [+ATR] in Akan. To resolve such conflicts, Archangeli (1984, 1988, pp. 193–196) proposes to consider the marked status of certain feature values as reflecting preferences, not invariant facts of Universal Grammar (UG). We may think of these preferences as the optimal, but not unique, values of universal parameters. Thus UG is said to prefer [+high], [–low], [–back], [–ATR] values in vowels. This means that, *in the normal case*, the opposite values ([–high], [+low], [+back], [+ATR]) will be marked and, hence, candidates for underlying specifications. The resulting system is merely the null hypothesis, not the only possibility entertained by a language learner, and so language specific facts may override the null hypothesis and lead to an alternative analysis of the data. In particular, evidence that a feature value such as the [–ATR] of Yoruba is phonologically active may be taken as sufficient grounds to set up an underlying inventory of values which reverses some universal preference. The UG markedness system will be able to manifest itself only in the absence of clear-cut data overriding it. The descriptive system resulting from this idea of parametrized markedness is flexible enough to handle a great deal of cross-linguistic variation in the behavior of features like [ATR]. The analyses of individual languages are, in principle, learnable, since the UG preferences act as guidelines to the language learner whenever the system of underlying values remains indeterminate.

My view of this is that parametrization is clearly involved in the Yoruba and Akan [ATR] systems; however, it is far from obvious that what is being parametrized is markedness. In the case of [ATR], we simply have no cross-linguistic basis on which to claim that one or another value is marked: [ATR] displays a cross-linguistic distribution that does not involve the sort of

context-free implicational relations upon which claims of markedness are normally based. We say that laryngealized sounds are marked because their presence in a system always implies that of the corresponding non-laryngealized ones. The presence of either [+ATR] or [–ATR] vowels does not imply cross-linguistically the presence of the other value: only certain combinations between ATR and height values can be said to be marked. And, if [+ATR] by itself is no more or less marked than [–ATR], then this is hardly the feature on which to base claims about markedness variation. Indeed, we can note that features like [nasal], for which implicational statements of the form *[αF] implies [–αF]* can be formulated, are *never* involved in markedness reversals. [ATR] simply does not belong in this set. Understanding what is cross-linguistically invariant in matters of segmental markedness, what is not, and why is the critical issue here; we return to it in section 2.3.3.3, where the theory of markedness is more fully explored.

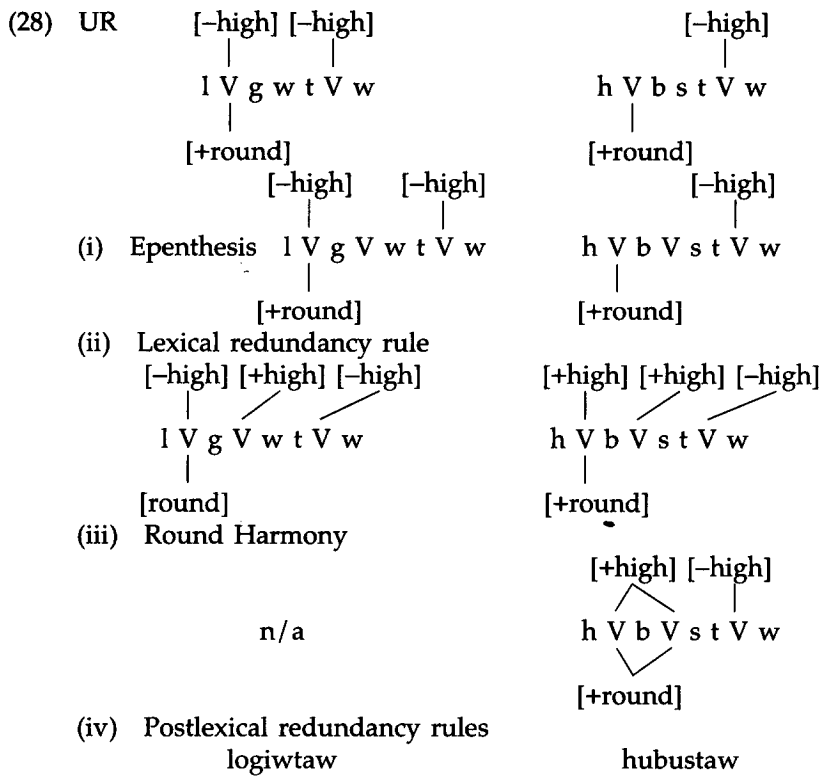
2.2.3 *Null Segments*

The link between underspecification and markedness reversals is also explored by Archangeli (1984, 1988) and Pulleyblank (1988a) in their discussion of *null segments*. These are segments carrying full surface specifications but behaving phonologically as if they lack many or most feature values. The typical null segment is the epenthetic vowel. By definition, this is not an underlying segment: under any reasonable account of vocalic epenthesis, the element first introduced by epenthesis into the string is a segmentally empty place-holder (Itô 1989). Only later rules specify the quality of the inserted nucleus. What are these rules? Archangeli (1984) points out that the unmarked assumption should be that these are independently needed mechanisms: the redundancy rules of the language. Yokuts (Newman 1949; Archangeli 1984) illustrates this connection between epenthesis and redundancy rules. Its underlying vowel system is {a, o, i, u}, analyzed in (27) in accordance with Archangeli's (1984) remarks:

(27)

	a	o	i	u	[ ] → [–round]	[+round] → [+back]
high	–	–			[ ] → [+high]	[+low] → [+back]
round		+		+	[–high, –round] → [+low]	[ ] → [–back]

In Yokuts, a vowel is inserted before an otherwise unsyllabifiable consonant (28i); this vowel is then subjected to the lexically applicable redundancy rule (28ii), after which Rounding Harmony applies between vowels of equal height (28iii). The remaining redundancy rules apply postlexically (28iv). The distinction between lexical and postlexical redundancy rules is dictated by the RROC: the lexical rule of harmony mentions [α high] (it spreads [round] only between [α high] vowels) and thus triggers the lexical application of [ ] → [+high]. All other vocalic redundancy rules are left to apply postlexically.



What evidence, other then the descriptive coherence of the system, do we have that this is the right analysis of Yokuts epenthesis? The evidence, as I see it, is that the epenthetic vowel is an undergoer, not a trigger of harmony; it is not a trigger because it lacks the harmonically active [+round] value. It lacks it because [+round] cannot be supplied by redundancy rules; it is a marked value. Beyond Yokuts, the hypothesis that redundancy rules are entirely responsible for the segmental makeup of epenthetic vowels makes the frequently correct prediction that these segments will be phonologically inert; they will never carry more specifications than other vowels and, frequently, they will carry fewer. They are typically non-triggers and undergoers, as in the Yokuts case.

Once we appreciate the attractions of this analysis of null segments, we can observe that it commits one to the hypothesis of markedness reversals, for it is clearly the case that languages with essentially identical inventories select different qualities for their epenthetic segments. A relevant comparison is that between Yokuts and Mohawk. The vocalic system of Mohawk contains {a, e, o, i, u} along with a unique nasalized vowel $\tilde{\text{a}}$, which could perhaps be viewed as altogether lacking an oral articulatory target. Like Yokuts, Mohawk epenthesizes a vowel in clusters that cannot be properly syllabified, but the

vowel epenthesized is *e*, not *i*. It is possible to analyze Mohawk *e* as underlyingly featureless, but that can only be done if we assume that Mohawk lists [+high], not [–high], as a marked value. The resulting system appears below:

(29)

	a	e	o	i	u
high				+	+
round			+		+
low	+				

[ ] → [–round] [+round] → [+back]
[ ] → [–high] [+low] → [+back]
[ ] → [–low] [ ] → [–back]

Pulleyblank's (1988a) discussion of Yoruba *i*, another possible null segment, clarifies an essential aspect of the analysis: the null segment need not be epenthetic. It can be underlyingly present as a placeless vowel, a (mostly) bare root node. What the typical null segment is claimed to lack invariably are the place features; they are missing because they can be inserted by redundancy rules. Given Lexical Minimality and the possibility of postulating a redundancy rule of the form [ ] → [αF] for every underlying [–αF], the theory of radical underspecification predicts that every language will possess a placeless vowel and perhaps an underlyingly placeless consonant, whether or not the language has epenthesis rules. Before looking into the evidence bearing on this point, we should however analyze more closely the logic of these predictions and the mechanisms they require.

The first point in need of examination is the principle in virtue of which place features can be left out of the underlying structure of null segments. Consider Yokuts again and its null vowel *i*. In this language there exist underlying *i*'s as well as epenthetic ones. What distinguishes an underlying placeless vowel from the utter absence of a segment? The answer, according to Pulleyblank (1988), is the root node. What justifies the existence of an underlying root node to which no place features are associated? The underlying presence of some stricture feature, say [+sonorant] or [–consonantal]. We come now to the critical question: why is this stricture specification present underlyingly? Clearly, it too could be inserted by a redundancy rule: [ ] → [+sonorant]. Moreover, both [+sonorant] and [–consonantal] values are predictable. Thus [+sonorant] follows, in the case of *i*, from [–consonantal]: there are no obstruent vocoids. The value [–consonantal] can be predicted from either the absence of a supralaryngeal component, as in the case of *h* and *ʔ*, or from typically vocalic place specifications such as [+round], [+high], [+low]: there are no consonants bearing such features in Yokuts. Markedness does not tell us whether to predict place from stricture or stricture from place; if anything, the fact that we cannot determine the stricture degree of a constriction unless we know where it is made suggests that we should try to predict stricture from place, not vice versa. In any case, the language learner needs to find some good reason *not* to eliminate [+sonorant] and signal the existence of

the underlying segment *i* by using some other feature, say [+high]. I submit that no such independent argument can be found. Pulleyblank, who touches on this issue in the course of analyzing Yoruba *i* as a null segment, suggests that Yoruba refers to [+sonorant] in one of its morpheme structure conditions. Therefore the need for underlying [+sonorant] is guaranteed, and every other insertable feature must be eliminated. But this answer, not persuasive in its details, is limited to Yoruba; it does not extend to languages lacking the relevant morpheme structure condition. The consequence is that the choice between marking an underlying null segment by using a stricture feature like [+sonorant] or by using a place feature like [+high] remains arbitrary.[24] No credible principle will lead us to the desired conclusion.

2.2.4 *Placeless Segments and Phonetic Underspecification*

The second point to consider is the functioning of the redundancy rules needed to provide the surface values of a null segment. Assume that Yokuts *i* starts out as a [+sonorant] root node. To be realized as [+high, –back, –round, –low], this segment has to be subjected to the application of some context-free redundancy rules. The rules were given in (27). We sketch below two steps of the filling-in process:

(30) [ ] → [+high]: $\text{root}_{[+son]}$ . . . Place . . . [+high]
[ ] → [–back]: $\text{root}_{[+son]}$ —— Place —— [+high]
. . . [–back]

Within the analysis proposed by Pulleyblank and Archangeli, it is critical that at least some of the redundancy rules involved in specifying null segments be context-free, for there are, by hypothesis, no input place features in a null segment that could condition the redundancy rules. It is this very assumption of context-free redundancy rules that turns out to be unworkable. The source of the difficulty are the segments, found in practically every language, which *start out and remain fully or partially placeless* on the surface: the laryngeal consonants *h* and *ʔ*, the schwa-like vowels, the central vowels lacking both [round] and [back] values, the segments displaying surface fluctuation in their exact constriction site (cf. the discussion of Gooniyandi below). We may call these segments *permanently placeless* and compare them to the null segments, which are believed to be only temporarily placeless. The supposedly cost-free, universal redundancy rules of the form [ ] → [+high], [ ] → [–back], [ ] → [+anterior] have inexplicably failed to affect the permanently placeless segments.

Let us consider one specific instance of this problem. We know that the laryngeal consonants *h* and *ʔ* are phonologically placeless (Steriade 1987a; McCarthy 1988; Pulleyblank 1988b); they must be represented as root nodes, possessing a laryngeal branch but no supralaryngeal node or values. We also know that they are phonetically placeless; they lack oral articulatory targets

(Keating 1988). We have established, then, that they never undergo any redundancy rules specifying place. The question then is what aspect of the structure of *h* tells context-free redundancy rules like those in (30) *not* to apply to it? How is *h* relevantly different from *i*? Is it the fact that *h* possesses a laryngeal feature? That suggestion makes the unlikely prediction that the inventory of aspirated or laryngealized vowels – in languages like Zapotec (Jones and Knudson 1977; Lyman and Lyman 1977) or Acoma (Miller 1965) – will somehow differ from that of modal vowels, in that the presence of aspiration will block the redundancy rules from applying to an aspirated V. In fact, however, the inventories of modal and laryngeally specified vowels are identical. Perhaps then *h* fails to undergo the vocalic redundancy rules because *h* is a consonant. But its status as a consonant refers to its possible syllabic position, not to its feature values: syllabicity should be irrelevant to the application of redundancy rules. We know, in any case, that vowels too can be permanently placeless: this is the case of schwa, which phonologists have analyzed as featureless (cf. Jakobson 1938, p. 129; S. Anderson 1982). As in the case of *h*, we can show that schwa remains featureless into the phonetic component; the discussion provided by Browman and Goldstein (1992) and their commentators indicates that English schwa either lacks a specified articulatory target or possesses a target corresponding to a vocalic neutral position "the mean tongue tract-variable position for all the full vowels" (Browman and Goldstein 1992, p. 56). Why have the universal context-free redundancy rules [ ] → [+high], [ ] → [–back], etc., failed to apply to schwa?[25]

There are several ways to look at the problem just identified. One is to characterize it as a technical issue: we must then find ways of reformulating the relevant redundancy rules in such a way as to allow certain segments to emerge as placeless from the phonological component. I doubt that a meaningful solution will be found along these lines. Another way is to abandon the hypothesis of Full Specification ((2) above) and, with it, the idea that context-free redundancy rules are needed in order to generate fully specified representations. If context-free redundancy rules are eliminated, the only feature-filling mechanisms will be of the context-sensitive category, e.g., [+back] → [+round]. They will therefore be applicable only if the segment contains some place feature to begin with; and consequently, genuinely placeless segments like *h* will be left placeless. This solution leads directly to the model of Contrastive Underspecification discussed below. The third possibility is to view the difficulty raised by context-free redundancy rules as an indication that no workable theory of redundancy rules can be formulated. This is, for different reasons, Mohanan's (1991) conclusion. We will see that the bulk of the evidence supports it. For the moment, though, we return to our discussion of Radical Underspecification to address a distinct issue: leaving aside matters of principle and technique, we should ask whether the available evidence supports the hypothesis of *temporarily* placeless segments. (I emphasize here *temporarily*, since that is the only category of placeless segment that supports the existence of context-free redundancy rules.) The next section takes up this question.

2.2.5 *The Evidence for Temporarily Placeless Segments: Epenthesis and Asymmetry Effects*

There are two sorts of arguments for postulating derivations like (30), in which a placeless segment acquires feature values through the application of context-free redundancy rules. One has been mentioned above: the scenario in (30) gives us a satisfying account of epenthesis, in that it uses only independently motivated mechanisms to specify the quality of inserted vowels. If we give up on context-free redundancy rules, epenthetic vowels like Yokuts *i* will have to emerge from the application of language-particular specification rules. The obvious concern here is that such specification rules are ad hoc mechanisms, revealing nothing about markedness or the functioning of individual sound systems. The other argument for context-free redundancy rules was developed at some length by Pulleyblank (1988b) and summarized by Archangeli (1988, p. 200) as *the asymmetry effect*: a variety of rules reveal that one and only one segment in a language is a selective undergoer of assimilation, a non-trigger of other processes, and a generally inert element in the system. This segment must then be underlyingly present but null: context-free redundancy rules will be needed to give it surface features.

We consider the asymmetry effect first. A distinction must be made at the outset between arguments establishing that an *underlying* segment of the language behaves as an all-around inert element and arguments showing that an *epenthetic* segment is inert. Abaglo and Archangeli (1989) demonstrate, for instance, that Gengbe *e* is inserted in a variety of circumstances, to satisfy minimal word constraints or provide a proper syllabification to underlying consonants. This *e* is also asymmetrically targeted by assimilation rules. The asymmetry effect in this case could well be due to the fact that the relevant instances of *e* contain no segment whatsoever, not to the presence of an underlying null segment. Details of Abaglo and Archangeli's derivations of Gengbe *e* (1989, p. 474) indicate quite clearly that the inert *e* starts out as a bare mora, not as a bare root node: it fails to undergo redundancy rules at the same time as other vowels of the language, presumably because there is no segment there to receive the redundant values. This case of asymmetry then boils down to the observation that epenthetic segments are not present underlyingly. We cannot dispute this fact, but the extent to which it provides an argument for context-free redundancy rules can only be determined when we look at other properties of epenthesis. A very different situation is that of Yoruba *i*, analyzed in detail by Pulleyblank (1988b). This segment *is* present underlyingly. Pulleyblank shows that its distribution is unpredictable and that it possesses a root node capable of spreading onto adjacent positions. It is then highly significant that there exist rules in Yoruba which single out this *i* as inert. The most revealing one is a phrasal-level regressive assimilation, triggered by all vowels but not by *i* (Pulleyblank 1988b, p. 238). Some examples of this optional process appear below:

(31)	V_1 # V_2 ($V_2 \neq$ i)			V_1 # i	
	owó adé	~ owá adé	"Ade's money"	ará ìlú (*arí ìlú)	"townsman"
	owó ɛmu	~ owɛ̃ ɛmu	"wine money"	ɛrù igi (*ɛrì igi)	"bundle of wood"
	ilé ayɔ̀	~ ilá ayɔ̀	"Ayo's house"		

Pulleyblank formulates the regressive assimilation as spreading the place node of the second vowel: if *i* has no place node, it cannot trigger the rule and the pattern in (31) is explained. The facts, however, are somewhat more complex. The only word-initial high vowel allowed in Standard Yoruba is *i*. The other high vowel, *u*, cannot occur initially. Thus we cannot tell whether the failure to trigger Regressive Assimilation characterizes all high vowels, or just *i*. According to A. Akinlabi (personal communication, 1992) the evidence of *u*-initial loanwords and dialectal forms suggests that *all* high vowels are non-triggers. This cannot be explained by claiming that both *i* and *u* are placeless. In addition, Akinlabi (1993) documents a lexical process in Yoruba whose application is triggered by high vowels: *r* deletes when next to *i* or *u*. This phenomenon calls for a reference to [+high] values in the lexical component of Yoruba, a fact which remains at odds with Pulleyblank's contention that only postlexical redundancy rules insert [+high].

Consider now the argument for context-free redundancy rules based on epenthesis. Clearly, the optimal outcome would be for epenthetic segments to be filled-in by processes needed elsewhere in the language. The question is whether this hypothesis can be worked into coherent accounts of individual systems. Several observations make this an unlikely prospect. First, there exist languages with *two* epenthetic vowels, both lexically derived: Hindi inserts *i* before *s*-obstruent clusters, and schwa between other C sequences (Mahajan 1993), while Hungarian inserts *a* before certain nominal suffixes generating disallowed CCC clusters and *o* into most other instances of CCC (Vago 1980, p. 53). Both these vowels cannot be due to the application of the same redundancy rules. Second, there exist languages where a null vowel – identifiable as such by the fact that it is a selective target of assimilation – contrasts with a non-null vowel of *identical* surface quality. Thus Hualde (1991) shows that Basque possesses a bare mora, which selectively undergoes a close-range assimilation from an adjacent vowel. Where no adjacent vowel exists, this slot is realized as *e*. Basque also possesses an underlyingly specified *e*, which does not pattern as a null vowel in assimilation. As Hualde points out, we need here a specification rule of the form *empty* $\mu \rightarrow$ *[e]*, not a set of context-free redundancy rules from which all *e*'s would emerge as null segments. A very similar argument was formulated by Hume (1992, pp. 273–274) for Maltese Arabic: one morpheme class contains vowels that are selective undergoers of assimilation. Where assimilation is inapplicable, these vowels surface as *i*. But not all Maltese *i*'s are selective undergoers; most are not. We cannot therefore analyze the null vowel of Maltese as undergoing general rules of the form [ ] $\rightarrow$ [+high], [ ] $\rightarrow$ [–back], and we cannot attribute the insertion of *i* to the fact that [+high] and [–back] are absent across the board from the underlying

representations of the language. Here, too, we need a specification rule that cannot be equated with an redundancy rule.

A different issue is raised by languages like Tiberian Hebrew, where epenthetic and reduced vowels receive a range of distinct surface values, depending on where in the string they occur. The data has been analyzed by Rappaport (1981), who notes that epenthetic vowels are realized as *e* in closed final syllables (/melk/ → [melek] "king"), as *i* in initial closed syllables (/k sbii/ → [kisbii]), and as schwa elsewhere (e.g., intermediate *mVlaakiim* → [məlaakiim], or underlying /b šalom/ → [bəšalom]). It is impossible to claim that the independently needed redundancy rules of Hebrew – whatever they may be – are responsible for the [+high] quality of inserted *i* in [kisbii] *and* for the [–high] of the *e* in *melek*. An added twist is that most Hebrew epenthetic and reduced vowels surface as schwa, the only plausible analysis of which is that it is placeless. There are no context-free redundancy rules operating here; there are only specification processes associating the features of *i* or *e* to certain empty nuclei.

If we conclude, then, that the quality of epenthetic segments should not be attributed to redundancy rules, we must explain several of the observations that made Archangeli's underspecification theory initially attractive. One point is that epenthetic vowels are typically targets of assimilation. The other is that when inserted vowels have a definite quality, they are generally drawn from the set {i, e, ɨ}: certain feature values (like [+round] or [+low]) and certain feature combinations (like [+round, –back] or [+low, –back]) are normally absent from the composition of an epenthetic vowel. This fact is clearly related to the markedness of these values. How then can we explain the fact that specification rules do not normally insert segments with the quality of u, ü, œ, ɑ or æ?

A possible line of investigation starts by observing that epenthetic vowels originate as schwa-sounding releases or transitions from one consonantal gesture onto the next. Many languages stop at this stage and leave the inserted vowel placeless, or subject to coarticulatory specification from neighboring sounds. In languages where a definite quality is eventually associated to the inserted segment, this quality could result from an attempt to identify the schwa-sound with a vowel quality that is phonemically present in the language. This is an instance of what Kiparsky (1968a, 1988) has dubbed "imperfect learning": the language learner comes to expects all vowels to be drawn from a certain fixed inventory and is therefore led to mistakenly identify schwa as being a deviant instance of some other vowel. If so, the crucial question will be, what underlying vowel does this schwa sound most like? It is here that considerations akin to markedness play a role: schwa is the vocalic neutral position and will therefore be identified with sounds that are in one respect or another closest to the neutral position. The absence of rounding, of front-rounding, or of a low jaw position (i.e., [+low]) can be straightforwardly explained in this way: these articulations involve a significant departure from the neutral position. We may finally speculate that the frequent choice of a

high vowel – typically *i* or *ɨ* – indicates a preference for the vowels that are phonetically shortest, perhaps because schwa itself starts out as a brief transition between consonantal gestures.

Whether or not this alternative interpretation of epenthetic vowels can be successfully developed, we should note that any connection between markedness theory and the quality of an inserted segment remains precarious in the context of Archangeli's version of Radical Underspecification. The reader will recall that one of the essential aspects of that model is that markedness facts are reversible. Thus, even though [–round] is the normally unmarked value, language-specific circumstances may require underlying reference to [–round] and hence a redundancy rule inserting [+round]. To the extent then that markedness plays a constant role in selecting the quality of epenthetic vowels, the theory cannot characterize this fact, for it cannot exclude the possibility that a language learner will find some reason to postulate a redundancy rule system from which *ɶ* or *ɑ* will emerge as the null vowel.

2.3 *Contrastive Underspecification*

One intuition that underlies Kiparsky's views on lexical underspecification is that redundancy rules based on context-free markedness principles, such as *[–voice] is the unmarked value of [voice]*, will be able to apply lexically. Only the feature values subject to *co-occurrence filters*, such as [voice] in **[–voice, +sonorant]*, will be uninsertable lexically. The prediction, then, is that postlexical rules will encounter segments that are fully specified for some feature F, provided that F is not subject to a co-occcurrence filter within that segment class, i.e., provided that F is distinctive. Thus Russian obstruents will emerge from the lexicon fully specified for voicing, but sonorants, not being distinctively voiced, will have no lexical values for [voice]. The theory of Contrastive Underspecification originates as an attempt to verify this prediction of Kiparsky's model, in somewhat weakened form. The survey of underspecification cases presented in Steriade (1987b) was undertaken initially in order to observe whether (32) holds generally:

(32) Priority for distinctive values: At every level of phonological analysis at which we find a feature F specified in a segment for which F is *not* contrastive, we will also find *both* values of F present in segments for which it *is* contrastive. However, the converse does not necessarily hold: we may find feature F behaving as a bivalent feature at a level at which feature F is not specified in segments where it is not contrastive.

As noted, (32) is simply a weaker version of the prediction made by Kiparsky's model, weaker because it does not mention anything about the boundary between the lexical and postlexical components. The survey results turned out to be largely compatible with (32): no credible cases were

encountered in which a distinctive F value was absent at a derivational stage where nondistinctive values for F were already present. Concretely, no cases were found (at least in 1987) in which, for instance, [+voice] values would have to be present in sonorants but [–voice] would be absent in obstruents.

A striking fact emerged from that survey that was, however, inconsistent with Kiparsky's views. It was the surprising scarcity of cases illustrating distinctive underspecification, i.e., the possibility that unmarked but distinctive F values might be missing underlyingly. This observation led to the hypothesis that distinctive F values, whether marked or unmarked, are always specified underlyingly. The clearest formulation of this position is Calabrese's (1988), paraphrased below:

(33) The Hypothesis of Contrastive Underspecification:
  (a) Feature values predictable on the basis of universal co-occurrence conditions can be omitted from underlying representations.
  (b) No other features may be underspecified.

The statement in (33) requires that both contrastive[26] values of a binary feature be present underlyingly. Thus, if [voice] is binary, obstruents must be specified as [+voice] and [–voice] respectively in underlying structure; only sonorants can be left unspecified for voice, since they alone are governed by a co-occurrence filter, *[+sonorant, –voice]. More generally, (33) bars underspecification of any feature F, binary or not, unless in segments where some *[(α)F, βG] condition can be validated, and F's values are predictable from G's. As we shall see in 2.3.2, (33) cannot be viewed as dictating the elimination of *all* feature values predictable from co-occurrence filters; it only indicates where we may find underspecification, not where we must.

In reviewing the evidence for (33), we must consider two types of arguments: those that establish (33b) by showing that underspecification is not found in certain cases, and those that establish (33a), by showing that feature co-occurrence conditions lead to underspecification. I emphasize again that our interest here is in the instances of temporary or nontrivial underspecification predicted by (33a). It can be anticipated now that these will be hard to come by. Trivial cases will turn out to be abundant.

The most significant aspect of Contrastive Underspecification is its rejection of Lexical Minimality. That principle is clearly being violated if we specify underlying voiceless obstruents as [–voice], instead of leaving them blank and using a redundancy rule such as [ ] → [–voice]. We must consider therefore first (section 2.3.1) the arguments for weakening or eliminating Lexical Minimality. A further point to be discussed is the necessary connection between Contrastive Underspecification and some theory of privative features: many potential counterexamples to the hypothesis in (33) were analyzed (by Mester and Itô 1989 and Steriade 1987b) as involving permanently privative rather than binary features. Privative features generate systems of lexical contrast that will appear to contradict (33): whether such evidence actually falsifies

(33) depends on the prospects for a coherent theory of privativity. A sketch of such a theory is presented in section 2.3.3, but its existence turns out to eliminate most of the evidence originally presented as supporting (33). Our general conclusion on Contrastive Underspecification will be that the work done in that framework was useful insofar as it shows why Lexical Minimality must be abandoned, but that the evidence for (33) as a principle inducing underspecification remains minimal.

2.3.1 *The Evidence against Lexical Minimality*

We consider now the evidence backing up the second half of (33): underspecification is not found in a large number of cases, where it could have served the purposes of Lexical Minimality.

At the outset a correction must be made in (33), without which the claim of restricted underspecification will appear to be patently false. In Steriade's and Calabrese's formulations, no account was taken of the instances of positional neutralization mentioned earlier (section 1.1.3), from which a great deal of underspecification appears to follow. Thus, for instance, [+ATR] is not licensed in Akan affixes (Clements 1978; Kiparsky 1985): the result is that the affix vowels are predictably [–ATR] in the absence of harmony. Harmony may be said to proceed from root to affix precisely because the affixes have predictable, hence underspecified, values for this feature: underspecification follows from predictability on the assumptions outlined earlier. Because the Akan affixal underspecification for [ATR] does not fall under the provision of (33a), Contrastive Underspecification appears to wrongly exclude it. This, I would suggest, was more in the nature of an oversight than an intended result; (33) can be corrected to allow all varieties of underspecification rooted in context-sensitive constraints. A revision apears below:

(34) The Hypothesis of Contrastive Underspecification: revised
   (a) Feature values predictable on the basis of universal co-occurrence conditions or on the basis of positional neutralization statements can be omitted from underlying representations.
   (b) No other features may be underspecified.

Having thus extended the scope of Contrastive Underspecification, we note that the study of positional neutralization reveals several interesting difficulties for the principle of Lexical Minimality. What we observe is that redundancy rules of the form [ ] → [coronal], [+anterior] are *not* being used to eliminate the point of articulation features of alveolar coronals from underlying representations, despite the fact that such redundancy rules would vastly simplify the underlying consonantal inventory and individual lexical entries. Plain alveolars appear to be place-specified, at least as [coronal], frequently as [coronal, +anterior]. Place underspecification in the coronal class is encountered, but only in the cases of features subject to positional neutralization.

In an important study of aboriginal Australian coronal systems, Hamilton (1993) points out that many Australian languages do not allow their full inventory of coronal features to surface in initial position. Thus Gaagudju, a language where apical consonants may be distinctively alveolar or retroflex, neutralizes this anteriority contrast in initial position: all initial apicals are, in the general case, alveolar. However, when an initial apical is followed within the word by a retroflex consonant, it surfaces as retroflex too:

(35) (a) naːwu 3rd sg. masc. pronoun
(b) deːn̪mi → d̪eːn̪mi "again, as well'
(c) niːɲja "just"

The full range of facts considered by Hamilton suggests that certain place distinctions are perceptible, in these aboriginal languages, only postvocalically, perhaps because the transitions from a preceding vowel carry clues indispensible to the nature of the articulation. We could record such facts in the form of a licensing condition. Thus Gaagudju apicals license [–anterior] only postvocalically. In initial position, therefore, the apicals are predictably [+anterior]: they must be alveolars, since the other value for [anterior] is unlicensed in the absence of a preceding vowel. It is these apicals with predictable anteriority that are being targetted for anterior assimilation; the word-medial anterior apicals, whose anteriority is distinctive, do not undergo assimilation. One may interpret this fact, in accordance with the principle of Invariance, as indicating that the distinctive word-medial apical anterior segments (d, n) are fully specified as [–laminal, +anterior]: they do not undergo anteriority harmony *because* they are fully specified. The initial apicals are predictably [+anterior] and hence unspecified for anteriority; because unspecified, they can undergo assimilation. This analysis explains why assimilation targets precisely the site of positional neutralization. It entails, however, that no general redundancy rule of the form [ ] → [coronal, –laminal, +anterior] has been employed in Gaagudju. Apico-alveolars may well be unmarked, but they are not underlyingly placeless. This contradicts Lexical Minimality, in the form in which this principle has been invoked as the fundation of Radical Underspecification.

Even more revealing is the case of Gooniyandi, also analyzed by Hamilton (1993). Here initial apicals are realized with free variation betwen an anterior and a nonanterior (retroflex) constriction site, while in other positions Gooniyandi has a lexical contrast between alveolar and retroflex apicals. Like Gaagudju, Gooniyandi displays assimilation between an initial apical and a following one. In such cases, the articulation of the initial apical is invariably alveolar or retroflex, depending on the point of articulation of the following apical:

(36) (a) duwu or d̪uwu "cave"
(b) d̪iṛipindi (no diṛipindi variant) "he entered"
(c) dili (no d̪ili variant) "flame, light"
(d) laŋgiya or l̪aŋgiya "midday"

It appears that in Gooniyandi neither [anterior] value is licensed initially. In this case, the phonological representation of the initial apicals lacks any value for anteriority and the phonetic component inherits structures that it is free to interpret variably. Harmony targets the initial apicals precisely because they lack [anterior] values. We have then two distinct ways to identify [anterior] unspecified apicals in Gooniyandi: they are subject to [anterior] harmony and, when harmony is inapplicable, they lack constant values for [anterior]. By these criteria, the medial apico-alveolars of Gooniyandi, e.g., *d* in *ḍiṛipindi* "he entered", emerge as fully specified [coronal, –laminal, +anterior]; they are non-targets of assimilation and they do not fluctuate in their anteriority values. Once again, the data contradict Lexical Minimality. They demonstrate that a perfectly coherent, markedness-based redundancy rule ([coronal, –laminal] → [+anterior]) has not been used to simplify lexical representations, for, if such a redundancy rule had been used, the medial apicals would have been underspecified too. Only *contextually* predictable features are omitted.[27]

Comparable arguments against Lexical Minimality are provided by many systems of vocalic harmony. One example, drawn from work by McCarthy (1979) and Flemming (1993), illustrates the entire class. Tigre has a series of long vowels {aː, eː, iː, oː, uː} and a series of short ones {ə, ɐ}. The short vowels contrast in relative height, but not in localization; both ə and ɐ are central. It is these short central vowels, and only they, that undergo a harmony spreading [–back] and [+back]. To express the connection between harmony and the underlying lack of backness distinctions in the short series, we may want to write a feature-filling harmony which spreads palatality from {i, e} and velarity from {u, o} onto the vowels unspecified for [back]. (An additional filter will be needed to prohibit the co-occurrence of [low] with [α back] in underlying or derived representations, since the central vowel *aː* does not undergo harmony.) Notice then that long vowels possess both backness values, since ə, ɐ are fronted before *iː, eː* and backed before *oː, uː*, whereas *aː* and ə, ɐ possess neither. Lexical Minimality is violated here by the failure to eliminate one or the other of the two backness/rounding values from the lexicon. (This would have been feasible in this system of vocalic contrasts; among the long vowels, the system could have been based on the feature values [+low], [+high], and [+round] or [+low], [+high], and [–back]. Among the short vowels, only a height contrast would be needed.) We may also view this case, like that of Gooniyandi above, as an argument against the assumption of Full Specification (2): where harmony is inapplicable, the central vowels fail to acquire *any* [back] specification.[28] This observation points to the same conclusion as the previous ones: there are no context-free redundancy rules of the form [ ] → [αback]. For the long non-low vowels of Tigre, such redundancy rules are unnecessary, since both backness values of {i, e, o, u} are phonologically active. For the Tigre central vowels *a*, ə, ɐ, such redundancy rules would be positively harmful, since they will prevent the central vowels from ever surfacing as central.

A very similar argument involving the palatality and velarity of Barra Gaelic

consonants has been formulated by Clements (1986b). In Barra Gaelic, too, one must assume that a three-way distinction between palatalized, velarized, and plain consonants exists underlyingly and persists into the surface representations. Here too we discover lexical entries that are less than minimally specified; here too we must assume that context-free redundancy rules like [ ] → [αback] are inoperative.[29]

2.3.2 *An Excursus on Liquid Underspecification and Liquid Transparency*

Mester and Itô (1989) had formulated an argument against Lexical Minimality that is very similar to those reviewed in the preceding section. The argument is based on the observation that the *r* of Japanese, whose coronality is predictable from its liquid status, functions as placeless, in contrast to the obstruent coronals. Mester and Itô suggest that obstruents like *t* are not placeless because their place features are distinctive, unlike those of *r*. Their overall conclusion is that, even when a coherent analysis could be built on the context-free redundancy rules mandated by Lexical Minimality – e.g., [ ] → [coronal] – such redundancy rules are not found to be in use.[30]

Mester and Itô (1989) discuss the phonology of Japanese mimetics and the behavior of palatality within this lexical class. Mimetics are characterized by a morpheme-level feature of palatalization which associates to the rightmost coronal consonant within the mimetic word (37ii). If the mimetic lacks a coronal, palatality associates to the initial consonant, including the arguably placeless *h* (37iii). Palatality does not associate to the liquid *r*, even when this sound is properly placed to receive it (37iv). It should be noted, however, that *r*-initial mimetics are not encountered.

(37) (i) Japanese consonant inventory: {p, t, k, b, d, g, s, z, m, n, r, h, y, w}

(ii)	potya-potya	(*pyota-pyota)	"dripping in large quantities"
	kasya-kasya	(*kyasa-kyasa)	"noisy rustling of dry object"
	dosya-dosya	(*dyosa-dyosa)	"in large amounts"
(iii)	pyoko-pyoko	(*pokyo-pokyo)	"jumping around imprudently"
	hyoko-hyoko	(*hokyo-hokyo)	"lightly, nimbly"
(iv)	nyoro-nyoro	(*noryo-noryo)	"slow wriggly movement"
	hyoro-hyoro	(*horyo-horyo)	"looking thin and weak"

The analysis of mimetic palatality proposed by Mester and Itô goes as follows: Palatality is a floating feature which seeks a coronal anchor, starting from the right edge of the stem. When it finds none (37iii–iv), it defaults on any consonant, seeking association in unmarked left-to-right fashion. The coronal liquid *r* is not chosen as an anchor for palatality in forms like *nyoro-nyoro* because *r* is place-unspecified; its place features may be assigned by a redundancy rule of the form [+cont, +cons, +son] → [coronal]. This placeless liquid is contrasted with the coronal stops and fricatives {t, d, s, z}, which must be

assumed to carry at least [coronal] values. Mester and Itô point out (p. 276) that one cannot recast the analysis in terms of saying that the palatality feature seeks a coronal qua placeless consonant. If that were so, there would be no reason to skip *r* in (37iv). It appears then that Japanese invokes only one means to derive place-underspecified representations: place features are omitted only when they are predictable from stricture values.

Two unanswered questions remain, however. Japanese has only one variety of fricatives; the sibilants {s, z} do not contrast with fricatives at other points of articulation. The same logic that predicts the place-underspecified status of *r*, might be expected to give us placeless sibilants, since a redundancy rule such as [+cont, –son] → [coronal] can be formulated on the basis of the co-occurence condition *[+cont, –son, labial or dorsal] and is consistent with the markedness facts involving the preferred point of articulation in fricatives. Yet forms like *kasya-kasya* establish clearly that the sibilants are fully-specified coronals. It appears then that Japanese does not exploit systematically the strategy of eliminating place features predictable from stricture specifications. This observation leads us to conclude that no general principle predicting underspecification emerges from – or is consistent with – Mester and Itô's analysis.

Behind the question just raised lurks a more serious difficulty. Although *s* is frequently the only fricative of a language, it never patterns as placeless. There are, for instance, no V-to-V assimilation rules that can skip *s*, but not other consonants. A variety of such rules, in which some subset of coronal sonorants pattern as transparent, has been documented by Kaun (1993a), who reanalyzes evidence first provided by Paradis and Prunet (1989), and by Parkinson (1993). The typical list of consonants transparent in V-to-V assimilations is *r*, *l* and occasionally *n*. The latter can be viewed in such cases as a nasalized flap, the nasal counterpart of *r*. Obstruents, whether or not possessing a distinctive point of articulation at their level of stricture, are never transparent. This observation suggests that the syndrome of liquid placelessness identified by Mester and Itô has no connection to issues of distinctiveness: the liquid is transparent not because its place features are predictable from its stricture, but for different reasons, which remain still unclear.

I illustrate this point – which will soon become significant in a different context – with the example of Cochabamba Quechua (Wallace 1988; McEachern 1993), where high vowels lower to mid before a uvular (38i), or before a cluster of a liquid + uvular (38ii). Before other C + uvular clusters, including those where C is an alveolar obstruent, the high vowels remain unaffected (38iii). The only apparent exception is that of (orthographic) *nq* clusters, which should be analyzed as homorganic sequences [Nq] (38iv). Vowels are lowered in this context, but this case can be assimilated to that involving high vowels strictly adjacent to a uvular. A distinct process lowers a high vowel immediately after a uvular (*qu* → *qo*): this process allows no segment to intervene between the factors of the rule.

(38) (i) pisi-qa → pise-qa "little" (topic)
(ii) riku-rqa → rikorqa "he/she saw it"
chilquy → chelqoy "to strip bark"
(iii) riku-sqa → rikusqa "he/she had seen it"
(iv) riku-nqa → rikoNqa "he/she will see it"

The regressive assimilation shown above could be viewed as the spreading of the Retracted Tongue Root specification from *q* onto a preceding vowel, subject to the condition that *q* and the vowel should have adjacent place nodes. A placeless segment may therefore intervene. If this is the right analysis, the transparent consonants are *r* and *l*, not coronal obstruents like *s*. This conclusion is reminiscent of Mester and Itô's analysis of the Japanese data, but with the disturbing twist that two distinct liquids must count as placeless. More significantly we observe that, exactly as in Japanese, *s* – whose point of articulation is equally nondistinctive in Quechua – does not behave as transparent.

We must draw from this case several related conclusions. First, the phenomenon of liquid transparency cannot be attributed to any form of temporary underspecification: it is highly implausible that Quechua *r* and *l* are placeless simultaneously. This conclusion is strengthened by Parkinson's (1993) results, according to which all coronal sonorants of Rwaili Arabic {r, l, n} are transparent in V-to-V assimilations. Second, we infer from the Japanese and Quechua data, as well as from Kaun's (1993a) survey results, that no matter how nondistinctive its point of articulation may be in a given language, *s* is never placeless: there are no redundancy rules of the form [–son, +cont] → [coronal, +anterior]. This is another way of recording the failure of Lexical Minimality, but, in this case, the problem identified casts doubt on Contrastive Underspecification itself, since the unattested redundancy rules are in fact consistent with (33).

2.3.3 *Which Features are Privative?*

We return now to the discussion of Lexical Minimality begun in section 2.3.1. If this principle is under attack, we must consider once again the evidence suggesting that binary features have only one value, the marked one, represented underlyingly. The omission of the unmarked value of [ATR], [voice], [round], etc., had been attributed by Kiparsky, Archangeli, Pulleyblank, and others to the requirements of Lexical Minimality. Once we abandon that principle, the analyses illustrating it must be reconsidered. The overall conclusion emerging from this review is that in most cases where it is justified to omit a feature value from underlying structure there is also considerable evidence that the value is permanently missing; the feature is privative.

2.3.3.1 *Privative [round]*

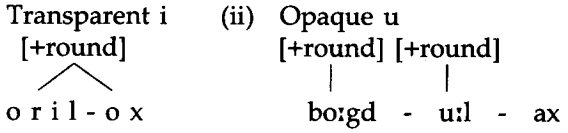
Consider first the behavior of [round], a feature whose unmarked value is [–round]. Unrounded vowels are, in all well-understood cases, non-triggers of

Round Harmony. In one case at least, that of Khalkha Mongolian, distinctively unrounded vowels are also non-blockers: they allow Round Harmony to proceed across them. The relevant facts, discussed in this context by Steriade (1987a), are summarized below: Mongolian spreads [+round] from an initial [–high] vowel onto a following [–high] vowel (39i). The high front vowel *i* – in any numbers or lengths – may intervene between the target and the trigger of harmony (39ii); the round high vowels *u* and *ü* may not (39iii).

(39) (i) Mongolian vowels: {a, e, o, ö, i, u, ü}
(ii) sons-ogd-ox "to be heard", örg-ögd-öx "to be raised" (cf. nee-gd-ex "to be opened")
(iii) oril-ox "to weep", oril-ogd-ox "to be wept" (*oril-ax)
(iv) boogd-uul-ax "to hinder" (*boogd-uul-ox)

The difference between neutral *i* and blockers *u*, ü, is clearly due to the fact that the latter possess the spreading feature [+round]. Since *i* does not block harmony, Locality requires us to assume that *i* does not possess [–round]:

(40) (i) Transparent i (ii) Opaque u



The problem arises when we note that the [–round] value of *i* is distinctive in Mongolian: *i* contrasts with the rounded front *ü*. Why then would *i* fail to be specified as [–round]?

A possible answer is that [–round] does not exist: the feature is universally and permanently privative. The chief predictions of this approach are that [–round] will never give rise to assimilation or dissimilation. We *can* refer to the absence of an autosegment – and unrounded vowels will a form natural class on the basis of the absence of [round] – but absence cannot spread and repeated absence does not violate the OCP and cannot lead to dissimilation. These predictions are largely correct.[31] It should be emphasized that, if the facts are correctly characterized as above, [round] must be privative regardless of the fate of Contrastive or Radical Underspecification; nothing else will explain the absence of assimilatory or dissimilatory [–round] effects. Let's grant, then, that [round] is privative. The speculation suggested by this is that all cases requiring underlying absence of a distinctive F value involve features that are in fact single-valued. We review briefly the relevant cases, with the aim to provide a general characterization of the features that may pattern as privative.

2.3.3.2 Privative [nasal], [spread] and [constricted]

The features of nasality, aspiration, and glottalization form a class by themselves, in that all assimilatory and dissimilatory processes involving them refer

to [+nasal], [+spread], [+constricted], never to the opposite values. To my knowledge, no explicit case for the binarity of aspiration and glottalization has been – or could be – made. The case against [–spread] and [–constricted glottis] as phonological values was presented by Lombardi (1991) and Steriade (1992). Nasality represents a slightly different case, as there exist processes possessing the appearance of local [–nasal] assimilation. Local postoralization (ma → m^ba) and preoralization (am → a^bm) have been discussed in terms of spreading orality (S. Anderson 1976; Kiparsky 1985). These phenomena, as well as other possible lines of argument for a [–nasal] value, have been reanalyzed as consistent with the idea that nasality is privative (Steriade 1993a, 1993b; Trigo 1993). There is virtually no evidence left suggesting that orality is represented phonologically, in any language.[32] The conclusion that nasality, aspiration, and glottalization are privative features helps explain, in the present context, frequent asymmetries in the patterning of nasal vs. oral sounds, or aspirated/ glottalized vs. plain consonants. Thus aspiration and glottalization are subject to frequent dissimilatory constraints, of which Grassmann's Law (Collinge 1985) is the best known: only one aspirated stop is allowed within a given root, in Indo-European, Sanskrit, and Greek. If aspiration is a binary feature, the hypothesis of Contrastive Underspecification (33) dictates that distinctively nonaspirated segments will be underlyingly marked [–spread]. Thus, in a consonant inventory like that of Sanskrit, where {p, t, k} contrasts with {ph, th, kh}, the unaspirated series will be marked [–spread]. Any reasonable extension of the principle of Generality (11) will lead us to expect then that the dissimilating feature may be [–spread] in addition to [+spread]. If it is, then roots like /p . . . k/ or /t . . . k/ will be disallowed, since they would contain two [–spread] values. But, as noted, unaspirated unglottalized stops are never disallowed from co-occurring with each other, in any domain. It is only the assumption that aspiration is a privative feature that allows us to analyze Grassmann's Law and comparable phenomena in ways compatible with (33); the plain series {p, t, k} possesses no feature value for either aspiration or glottalization, at any derivational stage. What is distinctive in this case is the permanent *absence* of laryngeal features, not the presence of a minus-value.

2.3.3.3 *Equipollent Features: ATR and [back]*

The behavior of ATR, [high] and [back] is significantly different from that of the features reviewed so far. There exist good examples of processes engaging either value of these features. If, for instance, [–ATR] is needed for Yoruba harmony (Archangeli and Pulleyblank 1989), [+ATR] is needed for Vata and Akan (Kiparsky 1985), and both [–ATR] and [+ATR] are required for Kalenjin (Ringen 1989), then it is impossible to claim that ATR is permanently and universally a privative feature with a cross-linguistically constant phonetic implementation. We might suggest that ATR's privativity is language specific,[33] but this hypothesis leads to a notational variant of the claim that the feature is binary but possesses reversible markedness (Archangeli 1988 and

Pulleyblank 1992). Let us accept then that the tongue root feature is equipollent. The problem here is, as pointed out by Archangeli and Pulleyblank (1989), that ATR gives rise to dominant/recessive harmony patterns, that are best analyzed in terms of saying that only one ATR value is present in a given system. We reviewed earlier the case of Yoruba ATR harmony, in connection with the discussion of the RROC. The Yoruba vowel system is {a, ɛ, ɔ, e, o, i, u} and it appears that only the [–ATR] values are active. Only [–ATR] triggers harmony (cf. (24) above) and only one distinctive [–ATR] value is allowed underlyingly within a root; the mid vowels ɛ, ɔ may co-occur within a root only as the result of [–ATR] harmony. All these observations suggest strongly that the recessive value [+ATR] is absent, at least in the underlying representations of Yoruba. The problem for Contrastive Underspecification is that this value is absent in {e, o}, the *distinctively* [+ATR] vowels. As discussed above, one proposed resolution of this difficulty has been to parametrize markedness, to assume that languages are free to choose which ATR value to count as marked, and thus, which ATR value to allow in underlying representations. Several problems with this line of thinking were pointed out earlier, in the discussion of Radical Underspecification. A distinct question may be raised now, after we have identified a class of genuinely privative features: if the choice between (underlyingly) privative [+ATR] and privative [–ATR] is determined on a language-specific basis, why is the choice between privative [–nasal] and privative [+nasal] not determined in comparable ways? If markedness is reversible in ATR, why is it cross-linguistically constant in [nasal], [round], etc.?

The answer I propose is that equipollent features like ATR correspond in fact to two distinct privative features defined as two opposing gestures on the same or related articulatory dimensions (see the parallel discussion in chap. 14, this volume). There are several ways to implement this idea, which can be enumerated here, but not explored. One possibility is that the privative [–ATR] of languages like Yoruba (whose vowel inventories contain {i, u, e, o, ɛ, ɔ, a}) should be identified as [+low] instead, as suggested by Goldsmith (1985) à propos of Yaka, by Goad (1992) and by Casali (1993). [ATR] is hardly needed in languages with limited vowel inventories, such as Yaka and Yoruba, where [low] is sufficient to distinguish {a, ɛ, ɔ} from {i, u}. Moreover, Yoruba-like systems never give rise to [+ATR] harmony rules, according to Casali, a surprising fact in the context of a markedness reversal theory. Casali also observes that nine-vowel systems (i.e., {i, ɪ, u, ʊ, e, ɛ, o, ə, a}), where the distinctive presence of ATR is undeniable, do not display harmony types in which [–ATR] is the active value. This too is unexpected on Archangeli and Pulleyblank's theory of parametrized markedness, but follows from Goad's and Casali's hypothesis: the set {i, u} does not differ from {ɪ, ʊ} in terms of [low], but in terms of advancement, and therefore ATR must be phonologically active in such languages. An alternative possibility, appropriate for languages like Azerbaijani Aramaic (Hoberman 1988), is that what Archangeli and Pulleyblank might call dominant [–ATR] in such systems is in fact [Retracted Tongue Root] (or [Constricted Pharynx]), a feature with distinct articulatory

properties from both the [low] of Yoruba and the privative [+ATR] of Igbo or Akan. The feature active in Aramaic harmony involves active tongue root retraction relative to the neutral position, while the privative [+ATR] of languages like Igbo or Akan would involve active tongue root protraction.[34] The parameter distinguishing Yoruba from Aramaic and Igbo would then involve choice of distinctive features rather than choice of marked value. An analysis along these lines makes a number of different predictions from that proposed by Archangeli and Pulleyblank, quite aside from its very distinct implications for the theory of markedness. Most important is the prediction that the articulatory correlates of so-called [–ATR] vowels should differ, depending on whether [–ATR] is the dominant or recessive value in harmony: the dominant [–ATR] (i.e., [low]) of languages like Yoruba should involve no active tongue root retraction, whereas that of Aramaic should.[35] Further, the dominant [+ATR] of Igbo is predicted to be phonetically distinct from the recessive [+ATR] (i.e., [–low]) of Yoruba. The latter should be phonetically equivalent to the recessive [–ATR] of Igbo, whereas the former should involve active advancement of the tongue root. A different prediction is that [RTR], [ATR] and/or [low] may coexist in a language. In the case of [ATR] and [RTR], this prediction may be confirmed in Kalenjin (Hall et al. 1974; Ringen 1989) and Chilcotin (Goad 1991).

Like [ATR], the feature [back] displays no clear markedness difference between its two poles, front and back. No implicational data supports the notion that one value of [back] is, taken in isolation, more or less marked than the other. The phonological behavior of [back] is also ambiguous: some [back] harmonies, such as that of Finnish (Kiparsky 1981; Steriade 1987b) are best analyzed in terms of privative [+back] while others, such as Chamorro (Chung 1983), clearly involve [–back] as the active value. It is conceivable that both palatality (i.e., [–back]) and velarity (i.e., [+back]) represent independent privative features, which may co-occur (as in Russian, cf. Keating 1985 and 1988, and Barra Gaelic, cf. Clements 1986b), but need not. One positive consequence of such a view is that phonetically central vowels (i.e., [ä], [ə], [ɨ], [ʉ]) could be analyzed more plausibly as segments lacking either palatality or velarity. Choi's (1992) acoustic analysis of the Marshallese central vowels strongly supports this interpretation.

Space does not permit a review of the privative or equipollent behaviors of other features, in particular the tonal features, as well as nontonal [high], [voice], and [anterior]. The goal of this section was not so much to provide an inventory of the features patterning like ATR as it was to support the idea that each articulatory dimension should be studied independently in determining whether it corresponds to one or several privative features. Conclusions about privativity, markedness, and, hence, underspecification based on the study of [nasal] or [round] are not directly applicable to features like ATR. Our discussion led to a proposal from which *all* features emerge as privative. In this respect, we come to partially agree with writers like van der Hulst and Smith (1985), Schane (1984, 1987), Kaye, Lowenstamm, and Vergnaud (1985),

Anderson, Ewen, and Staun (1985), Goldsmith (1985), Rice and Avery (1989), and Rice (1990), who have proposed inventories of single-valued features.[36] We have seen, however, that a uniformly privative theory requires a considerable expansion in the feature inventory; the question that should arise now is how to constrain this expansion. This issue is discussed further below.

To summarize, then, we have proposed that seemingly equipollent features like tongue root position and backness are in fact sets of two or more privative features defined on the same or perceptually related articulatory dimensions. A language may choose to utilize just one of the features drawn from a given set: Akan chooses ATR, not RTR, whereas Yoruba chooses [low] not RTR or ATR. The features proposed are universally privative, and no appeal to reversible markedness is being made in their treatment. The recessive segments, in any given harmony system involving such features, are recessive because they permanently lack the relevant values.

2.3.3.4 *A conjecture about the neutral position, markedness and feature privativity*[37] If [ATR] represents two distinct privative features of advancement and retraction, why is it then that nasality cannot be split into privative [raise velum] and privative [lower velum]? Why is rounding not the sum of two privative features, [purse lips] and [spread lips]? Why, in other words, is it that only certain articulatory parameters can be split into two privative features while others appear better suited for analysis in terms of a single privative feature? This is a fundamental question which should arise quite independently of the debate over the parametrization of markedness.

A possible answer to it is to be sought in Chomsky and Halle's (1968, p. 300) notion of *neutral position*, the speech-specific rest position of articulatory organs.[38] Certain articulatory dimensions, such as the vertical movement of the soft palate, may have a built-in asymmetry between their two extremes: one pole represents the rest position of the relevant organ, relative to which the other pole must count as a deviation. Such articulatory parameters will give rise to the standard analysis in terms of privative features. These features are privative because only one deviation is possible, along the relevant dimension, from the neutral position, or, alternatively, only one deviation has significant enough acoustic consequences.[39] Other dimensions, which perhaps include the horizontal movement of the tongue root, may have distinct and salient acoustic consequences regardless of the direction of displacement. Both advancement and retraction of the tongue root may, for instance, be analyzed as displacements relative to the neutral position of this organ. This interpretation of the facts allows the distinction between marked and unmarked feature values to be encoded directly into the representations. Context-free markedness in a sound reflects the fact that the sound involves a gesture that deviates from the neutral position. Corresponding to this articulatory deviation there is a linguistic mark, i.e., a feature specification. Absence of deviation along some dimension results in absence of a specification for the corresponding feature. Sounds which are unmarked, in the articulatory sense defined, are thus

phonologically unspecified. This proposal encodes directly only one aspect of segmental markedness, the articulatory effort involved in a deviation from the neutral position. Other sources of markedness – other properties that make a sound phonetically nonoptimal – appear to have either no consequences at all for its representation or may have different consequences from the ones discussed here.

Our speculation about the role of the neutral position in defining markedness and the range of possible feature values acknowledges the fact that different features have different markedness properties; nasality and tongue root position pattern differently and should not be used interchangeably in discussions of markedness. I suggested that differences in the markedness properties of different features should be linked to inherent asymmetries between the perceptual consequences of different articulatory gestures.

2.3.3.5 The impact of a privative feature system in theories of underspecification

The preceding sections led to the conclusion that the hypothesis of Contrastive Underspecification (33) can be maintained in the context of a feature system in which most, perhaps all, features are privative. Since many of the arguments originally presented as supporting (33) were cast in a framework where most features are binary, we must consider what is left of that evidence. This is necessary not only in order to assess the correctness of (33), but also to answer a core question of any underspecification theory: how extensive is the evidence that features present on the surface may be underlyingly missing? This question comes up as soon as we realize that a privative feature framework eliminates the need for most forms of nontrivial underspecification. Sonorants, for instance, are presented (in analyses such as Kiparsky 1985) as nontrivially underspecified for [voice]; they are claimed to lack underlying voice values, because such values are predictable, but they acquire them on the surface.[40] It is the existence of this residual class of nontrivial instances of underspecification that is at issue, for if such cases vanish as well, we will have to conclude that (33a) is false as stated, since it implies that some nontrivial underspecification is possible and attested.

The candidate analyses requiring nontrivial underspecification fall into two categories: those that make crucial use of binary features and those that are unaffected by the transition from a binary to a privative feature system. We briefly consider both.

(1) [Lateral] In the first class falls the case of lateral underspecification presented in Steriade (1987b) and amplified by Cohn (1993). In Latin, sequences of *l . . . l* dissimilate to *l . . . r*, presumably a reaction against the OCP violation involved in the [+lateral] . . . [+lateral] sequence. The only nonlateral whose intervention blocks this process is *r*. Thus *milit-a:lis* becomes *milit-a:ris* "soldierly" but *flo:r-a:lis* remains unchanged. This suggests that the *l . . . r . . . l* sequence of *flo:r-a:lis* contains the feature values [+lateral] . . . [–lateral] . . .

[+lateral] and thus no OCP violation. The idea then is that the liquids *l* and *r* are distinctively specified for laterality, since they differ only in that feature, whereas the nonliquids are redundantly nonlateral, hence subject to underspecification. This conclusion is compatible with Cohn's (1992a) findings for Sundanese, where it is the [–lateral] sequences /r . . . r/ that induce dissimilation. (Both analyses, it will be noted, violate Generality, since both assume that only one value of [lateral] generates OCP violations: *r . . . r* is well-formed in Latin and *l . . . l* is well-formed in Sundanese.)

The force of this argument for nontrivial underspecification resides exclusively, as Kenstowicz (1993) has noted, in the belief that nonliquid coronals like *t* acquire surface [–lateral] values. If they do not, for whatever reason, there are no grounds for distinguishing between an underspecified stage where the *t* in *militːaris* lacks a [–lateral] value and a surface stage, where it possesses one. There is in fact no reason to assume that [–lateral] is ever assigned to nonliquids like *t*. As for what distinguishes *r* from *l*, binary laterality is only one possibility. We could, alternatively, say that *r* is a retroflex rhotic (marked [retroflex] or [rhotic]) unmarked for [lateral]. Sundanese then dissimilates sequences of [retroflex] values, Latin dissimilates sequences of [lateral]. Kenstowicz suggests that the blocking effect of *r* in *floːr-aːlis* could be attributed to the fact that the dissimilated */floːr-aːris/ violates the OCP too, this time on the [retroflex] tier.[41] There is empirical evidence that an analysis along the lines suggested by Kenstowicz must be the right approach to lateral disharmony. Crowhurst and Hewitt (1993) cite a Yidiny lateral dissimilation similar to Latin (*l* → *r* / . . . ____ *l*) which is blocked in *r . . . l . . . l* strings, that is when a rhotic *precedes* the *l . . .l* string to which dissimilation might apply. It is clear that the blocking effect of the rhotic here could not be analyzed in terms of locality: the *r* does not come in between the potential factors of the rule and does not prevent them from being adjacent on their tier. Rather, the rhotic blocks the rule because dissimilation will turn *r . . . l . . . l* into *r . . . r . . . l*, a string violating the OCP on the retroflex tier.

(2) [Anterior] or [laminal] The case of [anterior] as a feature subject to nontrivial underspecification has been repeatedly discussed: the feature is taken to be binary in Steriade (1987b), where the Chumash sibilant harmony of the form s . . . š → š . . . š and š . . . s → s . . . s is analyzed. The nontrivially underspecified segment in such cases is *t*, the anterior coronal stop, which fails to trigger, undergo, or block the sibilant assimilation. The suggestion here too is to treat *t* as predictably [+anterior]. The Chumash consonant system does not oppose a [+anterior] nonsibilant to a [–anterior] one, and thus *t*'s anteriority can be considered nondistinctive. A phonetician might look at this case in very different terms, however: *š*-type fricatives differ from *s*-type fricatives in terms of the shape (slit-like vs. groove-like) and length of the constriction (Brosnahan and Malmberg 1970, p. 103). Stops contrast with each other in terms of the active and passive articulators involved in the constriction, not in terms of the shape and length of the constriction. Thus, one might well treat such cases as

involving featural distinctions that are permanently unavailable in stops, not instances of nontrivial underspecification. The other possibility, suggested to me by P. Ladefoged, is that the feature involved in Chumash and Chumash-like sibilant harmonies is laminality, a feature that is typically subject to phonetic (i.e., trivial) underspecification in the nonsibilant stops of languages, like English, French, or Chumash, where only one nonsibilant coronal stop exists. Dart (1990) demonstrates that in the absence of a *t*:*t̪* distinction, the unique coronal stop of English and French is free to fluctuate between apicality and laminality. In contrast, the sibilants *s* and *š*, which contrast, assume fixed values for this feature. We may speculate that the same holds for Chumash. If so, *t* fails to participate in harmony because it *permanently* lacks a value for [laminal]. Such phenomena argue then against Full Specification, not for (33).[42]

The issue of distinctiveness is, however, not irrelevant in the analysis of such cases. The permanent underspecification of *t* for [laminal] stems precisely from the absence of phonological contrast. But we see that what follows from the lack of contrast is not the temporary variety of underspecification that leads to the construction of derivational theories, but the permanent variety.

(3) [Voice] The example of voicing in sonorants has been used, here and elsewhere, as typical of the reasoning which leads to derivational approaches to underspecification. As noted at the beginning of this chapter, sonorants act, depending on the language, as fully specified [+voice] or as unspecified. The voice-unspecified sonorants of Russian (Kiparsky 1985) fail to trigger or block voicing assimilation; those of Japanese (Itô and Mester 1986) fail to trigger or block a voicing dissimilation. There is no doubt that, on the surface, the "unspecified" sonorants of Russian and Japanese are realized with a vibrating glottis, and in that sense, they are clearly voiced. But it is improbable that the sonorants share with the voiced obstruents all relevant articulatory adjustments that lead to vocal cord vibration. In producing voiced obstruents, the supralaryngeal cavity may have to be actively expanded, either by continuous larynx lowering or by tongue root advancement or by both strategies.[43] That such adjustments are not phonologically irrelevant has been shown in extensive detail by Trigo (1991); tongue root advancement in the production of the voiced stops of Madurese (Trigo 1991) and certain Akan dialects (Stewart 1967, cited in Casali 1993b) results in predictably ATR vowels. Such voiced obstruents must be specified, therefore, as both voiced in the sense of [vibrating vocal cords] and voiced in the sense of being pharyngeally expanded qua [ATR]. We may conjecture, then, that voiced obstruents share in permanent exclusivity some active process of pharyngeal expansion, which, according to the remarks on markedness sketched above, would have to correspond to a phonological feature. It may well be then that what propagates in Russian and dissimilates in Japanese is this element of pharyngeal expansion, a feature which sonorants will never acquire. Underspecification would therefore be permanent in such cases as well. On the other hand, in languages like English, where sonorants and voiced obstruents pattern alike in voicing assimilation,

the propagating feature would have to be the distinct element identified here as [vibrating vocal cords].[44] Our position is then that voiced (non-flap) obstruents are always specified as both [pharyngeally expanded] and [vibrating vocal cords]. Sonorants are permanently specified for the latter, permanently unspecified for the former.

(4) [High] We must return now to the pattern of Bantu vocalic assimilation discussed earlier. Recall that mid vowels propagate [–high] onto suffixal high vowels, whereas the phonetically non-high *a* does not. The facts are repeated below:

(41)

Past	Neuter	Applied	Gloss
tul-a	tul-ika	tul-ila	"dig"
fis-a	fiš-ika	fiš-ila	"hide"
kos-a	kos-eka	kos-ela	"be strong"
sek-a	sek-eka	sek-ela	"laugh at"
pat-a	pat-ika	pat-ila	"scold"

Two related issues arise here. One is the binarity of [high], whose plus-value is the only one to spread in Romance (Calabrese 1988; Vago 1988; Flemming 1993) and Menomini (Cole and Trigo 1989). Only [+high] represents a significant deviation from the neutral position, hence we would expect only [+high] to be active. The facts of Bantu are problematic in that [–high] is clearly active here. Second, if [–high] was phonologically available, the existence of the Bantu pattern would appear to support (33), in the sense of showing that a surface-present feature, the [–high] of *a*, is phonologically inert.

I believe that the problems identified here stem from our poor understanding of markedness. The key observation in this case is that mid vowels are underlyingly disallowed in Bantu suffixes, as well as most Bantu prefixes (Guthrie 1970); affixes have the reduced vowel inventory {a, i, u}. To understand what spreads in Bantu we must understand, even in preliminary terms, what motivates the limitation of affix vowels to this particular set {a, i, u}. The same preference for this maximally dispersed set of vowels is displayed by Chumash affixes (Applegate 1971) and by many instances of vowel reduction (Mascaró 1976; Calabrese 1988; Kamprath 1989). The high vowels may be preferred in such cases because they are maximally distinct from each other, and from the third vowel a of the triangle.[45] Articulatorily speaking, however, they are nonoptimal when compared to the non-high vowels, since they involve a greater deviation from the rest position of the tongue body. Conversely, the articulatorily optimal mid vowels have the perceptual disadvantage of poorer discriminability. The point that emerges from these remarks is that there are several dimensions of phonetic optimality which may well conflict with each other; the perceptually optimal vowel may not be the articulatorily optimal one. To fully account for the Bantu data considered here we must assume that both articulatory and perceptual nonoptimal qualities of a sound are encoded as distinct phonological features. In the case of height, the

articulatory mark encoding a deviation from the neutral position is [high], the feature giving rise to assimilations in Romance and Menomini. The perceptual mark, encoding say nonoptimal discriminability, may be provisionally referred to as [nonperipheral], a property shared by mid vowels and high central ones. We may say then that the Bantu pattern of harmony shown in (41) involves spreading the marked property [nonperipheral] from root to suffixes.[46] The low vowel *a* is peripheral, hence it is expected not to participate in this process. Here too it turns out that no appeal need be made to nontrivial underspecification.

The reader might object, at this point, that we have succeeded in avoiding several cases of nontrivial underspecification, at the considerable cost of introducing new features. Note, however, that the descriptive problems addressed here had not received solutions in any other feature system or theoretical framework. The Bantu problem cannot be solved if we simply decide to accept [–high] and spread it from root to affix; such a solution does not explain the connection between the Bantu harmony and the facts of markedness reviewed above. Nor will it explain why the low *a* always fails to spread its [–high] value: several Bantu dialects (reviewed in Goad 1992) have the harmony shown in (41) and *a* is active in none of them. This is not what we would expect if our account will consist of simply ordering the [–high] assimilation relative to the redundancy rule [+low] → [–high]. Why aren't there languages where the redundancy rule comes *before* [–high] assimilation? The existence of a [nonperipheral] class of vowels is further supported by the facts of Chumash, where the set {e, o, i̱} is allowed in roots but excluded from affixes. Chumash affixes, like Bantu ones, can only be drawn from the system {a, i, u} in underlying representation. It is clear that Chumash does not simply exclude [–high] from its affixes, since that would not account for the restricted status of i̱. Rather, both Chumash and Bantu limit the occurrence of the least discriminable members of their vowel inventory {e, o, i̱} to the most salient positions, that is, the root syllables. We continue the discussion of these issues in section 3.

2.3.4 *Conclusion on Contrastive Underspecification*

There are two sides to Contrastive Underspecification. On the one hand, it presents, as Radical Underspecification does, a derivational view of feature specification, sharing the belief that features present on the surface might be missing underlyingly. On the other hand, Contrastive Underspecification limits the cases of potential underspecification to features that are predictable in a syntagmatic context or in virtue of co-occurrence filters. In this second sense, in which Contrastive Underspecification departs from Lexical Minimality, there is some support for this view which was reviewed in section 2.3.1. We have observed, however, that most other proposed instances of nontrivial underspecification consistent with (33) can be reanalyzed. There are no truly convincing examples of a three-step scenario like (3), where we must start with contrastively underspecified representations, then apply a P rule, and

then let the redundancy rules take effect. In most cases, the evidence mistakenly interpreted as support for (33) yields to a better understanding of the feature involved in the process and reveals that underspecification is trivial and permanent. In other instances – such as the case of liquid transparency – we had to admit that no reasonable account exists at present, and therefore that no theoretical conclusions can be drawn.

Looking again at the list of potential examples of temporary underspecification reviewed, we note that the only solid ones are those discussed in section 2.3.1, all of which involve underspecification induced by positional neutralization. Thus Gaagudju apicals are underspecified for [anterior] word-initially, because no contrast is allowed in that position betweeen [+anterior] and [–anterior] apicals. Is a derivational approach necessary in this case at least? Must we start by leaving [anterior] unspecified in initial coronals, then apply [anterior] harmony, and then let a default redundancy rule apply which inserts, if harmony fails, [+anterior] in the position of neutralization? We see next that more interesting options exist.

3 A Nonderivational Approach to Positional Neutralization: Harmony Without Temporary Underspecification

I consider now in more detail the ways of formalizing positional neutralization in a grammar and its consequences for underspecification. The issue turns out to have wide ramifications and can only be sketched here. Throughout the discussion we must bear in mind the fact that the feature complexes disallowed from the positions of neutralization are generaly marked: they are the same elements that many languages bar absolutely. Any statement of positional neutralization must reflect this. A second preliminary remark is that some yet-to-be-formulated theory of prominent positions is assumed here. They include prosodically defined positions such as the mora (i.e., the syllabic peak), the stressed syllable, and the onset. They also include morphologically defined positions, such as the root or the content word, or the edge – typically the beginning – of the word. What determines membership in this class of prominent positions is still unclear. Many instances of positional neutralization discussed below refer to contexts that are prominent on more than one count: nasality, for instance, is allowed to associate underlyingly to Guaraní continuants only if they are moraic (i.e., vowels) as well as stressed. How to understand such cumulative effects is also unclear.

3.1 Direct and Indirect Licensing

Consider first the case of Greek aspirated vowels, which are allowed only word-initially. We identify the relevant feature combination as [+spread (glottis),

+sonorant]. In languages that prohibit it absolutely, a filter such as *[+spread, +sonorant] obtains. We unpack the contents of this filter as follows, using licensing language inspired by Goldsmith (1990).

(42) A [+spread] autosegment must be licensed, in every associated segment, by the presence of [–sonorant] in that segment.

This requires that every segment associated to the [+spread] feature possess the value [–sonorant]. Greek may be said to obey (42), with the diference that one class of prominent positions is also permitted to license aspiration: the word-initial vowel. The appropriate statement is (43), a version of (42) to which an escape clause, italicized below, has been added:

(43) [+spread] must be licensed, in every associated segment, by the presence of [–sonorant]; *or by an association to a word-initial mora.*

Consider now the distribution of nasalized continuants. Most languages prohibit such segments (cf. Cohn 1993 and references there). The corresponding statement is (44):

(44) [+nasal] must be licensed, in every associated segment, by the presence of [–continuant].

Many languages allow nasal vowels, but no other nasalized continuants. This suggests (45):

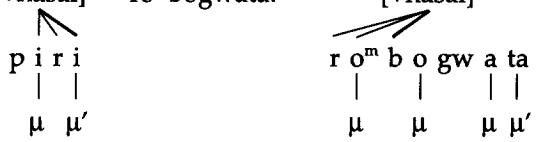
(45) [+nasal] must be licensed, in every associated segment, by the presence of [–continuant]; *or by association to a mora.*

In Copala Trique, nasal vowels – but no other continuants – are allowed to occur, but only in the stressed syllable (Hollenbach 1977). Trique obeys a modified version of (45), (46):

(46) [+nasal] must be licensed, in every associated segment, by the presence of [–continuant]; or by asociation to a mora, *in a stressed syllable.*

Next comes the case of Guaraní, the language where nasal vowels are allowed underlyingly under stress, whereas voiced noncontinuant nasals occur regardless of stress. Guaraní is exactly like Trique, with the difference that it also possesses nasal harmony: nasality spreads from the original segment that licenses it to all other segments of the stress foot. The consequence of harmony is the occurrence of strings like pĩr̃ĩ́ "to shiver" or r̃õᵐbogwatá "I made you walk" – represented autosegmentally below – where [nasal] is licensed in one of its associated positions, though not in all. Association of [nasal] to the licensed position is indicated below with a bold line. The vowel of the stressed syllable is designated as μ́.

(47) Guaraní pĩrĩ: [+nasal] rõ^{m}bogwatá: [+nasal]



This case has been analyzed (Kiparsky 1985) in terms of an initial stage of the grammar where the equivalent of (46) is upheld, and a later one, where it is violated. The structures in (47) would then illustrate the later stage. We note however that the basic facts can also be dealt with statically: Guaraní might differ from Trique in allowing a given [+nasal] value to be licensed by at least one associated segment, *not necessarily by all.* If so, we would write (48) as the positional neutralization statement valid in Guaraní. The italics identify the difference between Trique and Guaraní.

(48) [+nasal] must be licensed *in at least one associated segment* by the presence of [−continuant], or by asociation to a m̃ora, in a stressed syllable.

This statement permits the [+nasal] values in (47) to be viewed as licit, since they do possess an association to a segment – the stressed í or the m – where they are licensed.[47] We may refer to conditions like (46) as requiring *direct licensing,* and to (48) as permitting *indirect licensing* (Goldsmith's conception of licensing corresponds to what we call indirect licensing). The association of [nasal] to the Guaraní stressless continuants is indirectly licensed by the fact that the same [+nasal] autosegment possesses an association to a stressed vowel or to a noncontinuant.[48]

Morphological categories may also be involved in direct, as well as indirect licensing. Thus Mazahua Otomí (Spotts 1953, p. 254) has a category of affixes (called stem formatives by Spotts) which consist of one or more consonants followed by a non-epenthetic vowel of schwa-like quality. The full vowels of Mazahua – a, ʌ, ə, ɔ, e, o, i, u, oral or nasal – may not occur in the stem formative morphemes, except in virtue of root-controlled harmony:

(49) khù-ʔu "sister of a man" ɲi-ʔi "head"
kə̂-sʔã "to be on top of" sʌ́-tʔʌ "arrive here" tê-sʔe "climb a tree"

Simplifying somewhat, we may state that the category of stem formatives identified by Spotts is coextensive with the derivational affixes of Mazahua. In this case, the licensing statement will be the one below:

(50) [αF] , where F is any vowel feature, must be licensed, in at least one associated segment, by a root or inflectional morpheme.

Interestingly enough, the rounded vowels of Mazahua do not induce harmony across the nonlaryngeal consonants: *khù-ʔu* "sister of a man" is a special

case, encountered only when *h* and *ʔ* come between the stem-formative schwa and the root vowel. Aside from this case, there are no rounded vowels in the stem formatives: a rounded stem vowel is followed by ə in the stem formative.

(51) tʔo-šə "nest" tʔɔ-šə "while"
ngu-mə "house" tʔõ-rə̃ "firewood"

We may attribute the absence of round-triggered harmony to the existence of a separate, stronger, licensing condition that governs [round]:

(52) [+round] must be licensed, in every associated segment, by membership in a root or inflectional morpheme.

(52) states that every segment possessing [round] must belong to some nonderivational morpheme. And this bars it from the stem formatives, even in the case where it is harmonically linked to the root.[49]

A similar situation is encountered in Chumash, the language where roots may contain peripheral vowels {a, i, u} as well as central ones {e, o, i̵} (Applegate 1971). Affixes, however, may contain only peripheral vowels. Chumash too has affixal harmony: prefix *a* assimilate to a non-high vowel in the root; and prefix *u* assimilates to root *i̵*.

(53) aqpala-woyoc → oqpolo-woyoc "to wear down crookedly" (woyoc "to be crooked")
qal-wala-tepet → qel-wele-tepet "to roll up and tie bundle" (tepet "to roll")
yul-cʼi̵ → yi̵ci̵ "to be sharp of heat (= to be hot)" (cʼi̵ "to be sharp")

We are here in the presence of a system which licenses marked (nonperipheral) vowel qualities only in the root morphemes. Indirect licensing is permitted, as in Mazahua.

(54) Nonperipheral vowels must be licensed, in at least one associated segment, by membership in the root morpheme.

Proto-Bantu appears to display an exactly identical restriction (Guthrie 1970): the vowel system in roots includes {i, u, ɪ, ʊ, e, o, a}, whereas affixes are drawn from the peripheral set {i, u, a}, with only a few prefixes belonging to the high lax category {ɪ, ʊ}. No affixes contain mid vowels. Indirect licensing is permitted, as in Chumash; the suffixal vowels are subject to a height harmony whereby *i, u* become *e, o* by assimilation to a preceding mid vowel.[50] The specifics of this process were discussed in section 2.3.3.5(4) above.

An interesting comparison between direct and indirect licensing is occasioned by the analysis of [round]. Several Altaic languages – Vogul (Kálmán 1965), Bashkir (Poppe 1962), and Ostyak (Trubetzkoy 1939) – allow round

vowels only in the initial syllable of the word. These languages require the direct licensing statement in (55):

(55) [+round] must be licensed, in every associated segment, by membership in the initial syllable.

This states that [+round] may be linked only to segments belonging to an initial syllable. It thus rules out both disharmonic strings like *iCu* and harmonic ones like *uCu*.

Better known, and probably more widespread, are languages where [+round] is said to originate only in the initial syllable, but where it is able to spread beyond its directly licensed position. This is the case of Yokuts (Newman 1949), a language where only the first vowel may be distinctively round. Yokuts harmony spreads the feature further into the word, from [α high] to [α high] vowel. Kuroda (1967, p. 43) describes the distribution of [round] in Yokuts as follows: "if the first vowel of the word is non-round, all the other vowels in the word are non-round; if the first vowel is round the round vowels follow until a vowel with an opposite value for diffuseness [i.e., height, D.S.] is reached, after which all vowels in the word are non-round." This indicates that [+round] may occur initially, as well as in contexts where it can be attributed to harmony, but nowhere else. Indirect licensing is clearly at work here.

(56) [+round] must be licensed, in at least some associated segment, by membership in the initial syllable.

The effects of (56) can be inspected below: [+round] is directly licensed in (57a), indirectly so in (57b). It is unlicensed and, hence, ill-formed, in (57c).

(57) (a) gophin (b) ʔopottow *(c) paxattow (unattested)

[+round] [+round] [+round]

Root controlled harmony systems, such as that of Akan (Clements 1978) are very similar to Yokuts. In Akan, the active [+ATR] value is licensed in the root. It may be indirectly licensed, and this generates the effects of harmony.

There are obvious extensions of the idea of indirect licensing to the cases of local assimilation in which onsets spread place or laryngeal features onto adjacent coda consonants. These cases have been dealt with in terms comparable to those proposed here by Itô (1986) and Lombardi (1991).[51] Thus, a coda obstruent may not possess an independent [voice] value in languages like Russian, German, Polish, or Catalan (Mascaró 1976 and Lombardi 1991). In a subset of these languages, onsets spread [+voice] onto preceding codas. Such cases fall under the statements in (58); the difference between languages like Russian, where there is onset-to-coda assimilation and languages like German, where there is not, concerns only the possibility of indirect licensing.

(58) Licensing [+voice] in German (a) and Russian (b)
  (a) [+voice] must be licensed, in all associated segments, by the presence of [+sonorant]; or by membership in the onset.
  (b) [+voice] must be licensed, in at least some associated segments, by the presence of [+sonorant]; or by membership in the onset.[52]

The same approach handles the cases of place feature licensing in coda consonants that were discussed earlier in Steriade 1982 and Itô 1986: languages in which codas must be homorganic to following onset consonants. In such cases, we would say that consonantal point of articulation features are directly licensed in the onset, indirectly so in the coda. A statement like (59) will handle the relevant facts of languages like Japanese or Diola Fogny:

(59) [aF], where F is a consonantal point of articulation feature must be licensed, *in at least one associated segment*, by membership in the onset.

Languages where no supraglotally articulated codas are allowed (e.g., Western Popoloca, described by Williams and Pike 1973), but only *ʔ* or *h*, impose a stricter condition of direct licensing:

(60) [αF], where F is a consonantal point of articulation feature must be licensed, *in all associated segments*, by membership in the onset.

One further comment on the connection between indirect licensing and assimilation. Indirect licensing conditions are necessary for the existence of multiply linked structures such as (47) and (57), but they are not sufficient to predict the occurrence of assimilation. To fully characterize individual cases of assimilation one must define further conditions – having to do with the locality of the rule, its bounding domain, etc. – which were left out of the picture here, because they are not related to our argument.

3.2 *Significance of Indirect Licensing Conditions for the Theory of Underspecification*

It is important to emphasize, in connection with the licensing conditions just studied, a point made, I believe, by Itô: the homorganicity conditions governing coda consonants are frequently static, in the sense that they may obtain only morpheme-internally or may not be supported by synchronic alternations. There are frequently no reasonable grounds for a derivational analysis in which an earlier, place-unspecified coda becomes a later, place-specified segment, through assimilation to an onset. It is significant then that our indirect licensing approach, like Itô's earlier analysis, is static; it does not distinguish between an initial stage where place features are allowed only in the onset and a later stage where they are allowed in the coda too. The distinction we draw

between the place features of coda and onset segments is only that between directly and indirectly licensed positions, not that between derivational stages. This is worth mentioning, because the very same device of indirect licensing, upon which we base a static analysis of coda conditions, was shown to apply equally well to phenomena such as Guaraní continuant nasalization or Yokuts rounding harmony. These rules had in the past received only derivational treatments based upon the distinction between early and late stages. The reader will recall, for instance, that Guaraní was analyzed by Kiparsky (1985) as involving a first stage where nasality is allowed only on stressed vowels and stops (our directly licensed nasal segments) and a later stage, where it is permitted on all continuants. Yokuts and similar systems had also been analyzed derivationally (by Steriade 1979 and Archangeli 1984) in terms of an initial stage where [+round] is well formed only in initial syllables and a later stage where it is permitted across the board.[53] We have suggested that the obviously static phenomena – such as the prohibition of independent place features in coda – and the less obviously static ones – such as the prohibition of independent nasality on Guaraní stressless continuants – should receive the same sort of treatment, one based on indirect licensing conditions. The static analysis presented here brings out the property shared by this entire class of phenomena, the fact that some features may surface only when licensed by an associated slot.

This conclusion has a direct consequence for the hypothesis of underspecification. One of the reasons for postulating segments that are underlyingly unspecified for some feature F, is that these segments, and only they, are targeted by assimilation rules that spread F. Specified segments, it is widely believed, do not undergo assimilation. But the largest class of unspecified segments subject to "later" assimilation for the unspecified feature is precisely the class of segments identified here as containing features subject to indirect licensing. Our analysis of indirect licensing phenomena did not appeal to underspecification. We found no reason to distinguish a first stage, where some segments lack values for F, from a second stage, where they acquire F values.

The reader may wonder whether a distinction between an initial stage and a later one must not also be built into our indirect licensing account in the case of morphemes that do not contain a proper licenser for some feature *within their underlying representations*. For instance, Yokuts suffixes do not contain a proper licenser for the feature [+round] internally to their lexical entries; only an association to the initial syllable licenses [+round] in Yokuts, and the suffixes are necessarily non-initial. Only after the suffixes have been attached to a stem possessing a licensed [+round] value can they acquire a proper association to that feature. The question then is what we declare the underlying [round] value of these suffixal vowels to be. And the only answer that appears to make sense is that they have no value for [round] at all: [+round] is impossible, [–round] is idle and probably nonexistent. My proposal here would be that [round] values in the suffixes of Yokuts, and comparable cases elsewhere,

are not underlyingly underspecified. Rather, they are *unlearnable*, whether or not they are specified. They are unlearnable because they will never be in a position to manifest their existence.

That this suggestion is on the right track is indicated by the rare case of a class of morphemes which can occur in both direct and indirect licensing positions. These are the morphemes *nal* "at" and *töl* "from" of Hungarian, which can be used both as stems and as suffixes (cf. Vago 1976, p. 244ff.). In Hungarian only a stem, and in native stems only an initial syllable, can license [back] values. As stems, *nal* and *töl* display unpredictable [back] values which determine harmony onto suffixal morphemes: *nal-am* "at me" versus *töl-em* "from me". As suffixes, these morphemes cannot license their underlying [back] features and thus become subject to harmony: *ház-nal* "at the house", *ház-tol* "from the house"; *kép-nel* "at the picture", *kép-töl* "from the picture". The [back] value of these morphemes happens to be learnable only because they can also occur as stems. Any account of Hungarian harmony (e.g., Ringen 1988a) that rests on the idea that vowels must be *underlyingly* unspecified in order to undergo vowel harmony will necessarily have to leave the behavior of these morphemes unexplained.

4 Conclusions

At the beginning of this survey, I divided the circumstances in which a feature value might be predictable, and hence unspecified, into three classes: features predictable from co-occurrence conditions, feature values identified as unmarked by a context-free statement, and features subject to positional neutralization.

The first category yields no clear cases of nontrivial underspecification.[54] Let us return to the original example of sonorant voicing, a feature apparently predictable from the co-occurrence condition *[+sonorant, –voice]. Perhaps sonorants act as [0 voice] in Russian because their voicing is predictable, but an equally good explanation for the facts is that the feature involved in the Russian process is inherently a feature characterizing obstruents alone. Pursuing this second approach has two important benefits. First, it allows us to provide a one-step account of the Russian assimilation facts – and other similar paradigms – without proliferating derivational stages. All we need to state is that onset obstruents propagate onto coda obstruents the feature of pharyngeal expansion that makes vocal cord vibration possible. On this account, we need not distinguish what sonorants look like underlyingly and on the surface, and we need not invent ordering principles that will predict when they are allowed to acquire their surface [+voice] value. The second benefit of abandoning the hypothesis that predictability rooted in co-occurrence filters yields temporary underspecification is this: we need not wonder any longer why certain patterns of predictability never lead to underspecification. Why, for

instance, does the predictable coronality of Japanese or Quechua *s* not lead to temporary placelessness? This is a question that does not arise once we admit that nothing follows from a feature's predictable status.

The second category of potential underspecification is that rooted in context-free markedness. I illustrate this case with the example of nasality. The unmarked state of speech sounds is to be oral. If that means that they are specified [–nasal], we might have to ask ourselves whether a redundancy rule like [ ] → [–nasal] is appropriate, and when. But, as it turns out, where context-free statements of markedness are justified, the corresponding features are best viewed as privative. Underspecification in such cases is genuine, but permanent. No redundancy rules are needed. The reader is reminded that this conclusion spares us the considerable difficulties encountered by context-free redundancy rules (these were reviewed in sections 2.2.4–2.2.5). I do not debate the existence of specification rules, rules which provide features for epenthetic vowels in specific contexts, but it seems impossible to equate these with redundancy rules affecting all the segments of a language.

We expected to encounter a third variety of underspecification in the study of positional neutralization. Neutralization means predictability and predictability may lead to temporary underspecification. But here too, we saw that it is at least possible to view the facts in purely declarative terms, without distinguishing between the underspecified lexical representations and the fully specified surface structures. Given the simplicity of the declarative analyses presented, I feel that the burden of proof belongs in this case with the defenders of a derivational approach to positional neutralization. The simplest assumption is that no cases of temporary underspecification are to be found here either.

The evidence, although not always clear, does not appear to support the hypothesis of nontrivial underspecification. We may recall now the principles that led phonologists to this hypothesis, Lexical Minimality and Full Specification. These are ideas that presuppose a derivationally organized phonological component. Neither seems likely to survive.

NOTES

This chapter originated as lecture notes written in 1991 for a UCLA seminar on underspecification theory. The contribution of the seminar participants is gratefully ackowledged. Special thanks to Abby Kaun, Jongho Jun, and Edward Flemming, all of UCLA, and to Andrea Calabrese, Morris Halle, Phil Hamilton, and Keren Rice, whose own research on related topics has clarified my thoughts on many of the issues addressed here. Dani Byrd and Ian Maddieson provided extremely useful comments and references for the sections dealing with markedness. Finally, I am grateful to Elan Dresher for reminding me that, if underspecified representations exist, that ought to follow from something.

1 Other formulations of Lexical

Minimality appear in Archangeli (1984, 1988).

2 This principle is explicitly discussed by Kenstowicz (1993).

3 See Goldsmith (1992), Prince and Smolensky (1993), Scobbie (1992), and references there.

4 The terms, and the distinction made by them, appear in Steriade (1987b) and Archangeli (1988).

5 Although the history of derivational underspecification cannot be retold here, we should mention works like Halle (1959), Ringen (1975), Clements (1977), and Kiparsky (1981) where this practice was first explicitly defended and where efforts were made to support the assumption of underspecification with empirical evidence. Trubetzkoy (1939) is also sometimes cited in this connection (cf. Clements 1987 and Archangeli 1988). He did not, however, subscribe to the version of *derivational* underspecification we discuss in the text. Trubetzkoy formulated criteria to determine which features are phonologically irrelevant in a system, but he was assuming mono-stratal grammars in the context of which it is impossible to believe that irrelevant features are specified at some late stage.

6 See Calabrese (1988) for a defense of the filter-based formulation; see also Myers (1991) for an alternative. Further discussion of this issue appears in Mohanan (1991).

7 Although the connection between markedness and underspecification is seldom made explicit, one gathers that universally valid co-occurrence statements are considered more plausible sources of underspecification than the parochial co-occurrence statements holding within one language only. Thus Calabrese (1987, p. 290ff.) cites the case of the Russian affricates *ts*, *tš*, which are predictably voiceless, unlike the other Russian obstruents, because they happen to lack voiced counterparts. Although predictably voiceless, these affricates do not pattern as unspecified for [voice], unlike the predictably voiced sonorants. Calabrese suggests that the fully specified behavior of Russian *ts* and *tš* should be attributed to the fact that no universal statement bars voiced coronal affricates. In contrast, a universal statement – formulated by Calabrese as the *[+sonorant, –voice] – renders the voicing of sonorants predictable. Calabrese's proposal leaves one to wonder why language-specific feature co-occurrence conditions – like the one involved in Russian *ts* – are ignored in determining how much phonological information to leave unspecified.

8 This problem was first noted by Stanley (1967), who invokes it to motivate in part his rejection of underspecification. Attempts have been made at resolving this difficulty while accepting some version of underspecification (Christdas 1988; Steriade 1987b) but both proposals violate Lexical Minimality and offer no alternative principles as the basis of underspecification patterns. Goldsmith (1990, p. 244) also notes the problem of arbitrariness and appears to suggest, in line with proposals by Archangeli (1984, 1988), that the particulars of a phonological system will always help determine which feature to leave out of the underlying forms. On this line of reasoning, see Calabrese (1988, chap. 2).

9 This point is made by van der Hulst and Smith (1985) and Goldsmith (1990, p. 245).

10 This section and its continuation (section 3) were written in answer to questions raised by Edward Flemming, who is hereby thanked.

11 Odden (1993) provides a useful review of the issue. See also Archangeli and Pulleyblank (1987), Steriade (1987a), and Myers (1987).

12 This notion, which turns out to be considerably richer, is explored under the general name of Faithfulness by Prince and Smolensky (1993).

13 The notion that some constraints are more important than others is explored at length by Prince and Smolensky (1993).

14 The term is due to Mester and Itô (1989).

15 Unlike Chomsky and Halle, however, Kiparsky appears to assume that some markedness difference between values can be observed for every feature. Thus the marked value for [back] is assumed to be [+back] (1981); and the marked value for [ATR] is assumed to be [+ATR] (1981, 1985). The 1981 MS suggests that the markedness difference between feature values should be interpreted in privative terms: only the marked value exists, at all derivational stages. This suggestion seems to have been abandoned in the 1985 article; it did, however, reemerge in the more recent work of other phonologists, notably van der Hulst and Smith (1985 and later). We discuss it in section 2.3.3.

16 I cannot find specific statements to this effect in Kiparsky's papers, but the analyses presented all appear to rely on this assumption. This aspect of the theory of markedness was made fully explicit by Calabrese (1988). Calabrese adds the hypothesis that filters are organized in a partial ranking such that the violation of a low-ranked filter in a language entails the violation of a more highly ranked filter.

17 It is worth noting that the ordering of assimilation before (17) cannot be attributed to a general principle (such as the Elsewhere Condition: Kiparsky 1973a). In other languages, such as English, the redundant voicing of sonorants does trigger voicing assimilation.

18 If voicing assimilation must apply lexically as well as postlexically, then we need either (15)–(16) or alternative mechanisms in order to insure that sonorants will not become voiced in time to trigger the postlexical voicing assimilation. But no clear arguments establish the need for a *lexical* application of voicing assimilation. If so all we need is the stipulation about rule ordering coupled with the assumption that voicing assimilation applies postlexically both within and across word boundaries.

19 Further facts about the typology of voicing assimilation suggest that the connection to Structure Preservation should be downplayed in any case. Thus voicing spreads from sonorants in languages like English (Mascaró 1987), Yakut (Kenstowicz 1993), Greek and Latin (Steriade 1982), and many others. These cases involve word-internal assimilations that fail to apply in the phrasal contexts where postlexical processes are expected to act. These are then lexical rules applying *after* (17); only the need for extrinsic ordering appears to survive this comparison.

20 See Kaun (1993a) for a careful discussion of this case. Kaun demonstrates that the Catalan coronals must in fact be fully specified already within the lexical

component and that the asymmetry between alveolars and other consonants is unrelated to underspecification.

21 A distinct source of difficulty for Kiparsky's model involves the predictions of Structure Preservation. Although (15) expresses a widespread tendency to uphold within lexical derivations the set of filters obeyed underlyingly, the facts uncovered so far suggest that lexical rule applications are not invariably structure preserving. There is extensive discussion of this point in Calabrese (1988).

22 Goldsmith's (1990, p. 243ff.) presentation of underspecification within lexical phonology makes this point very clear.

23 Cf. also Abaglo and Archangeli (1989).

24 Archangeli (1988) appears to suggest, following Christdas (1988), that certain features are always fully specified. Stricture features are among the fully specified set. If so, [sonorant] will always be present in an underlying segment, and all other features could be left out. But this is a departure from Lexical Minimality – since we are specifying more features than we absolutely must – and we need to understand what motivates it. It is not sufficient to say that there are no markedness considerations on the basis of which to write a redundancy rule like [ ] → [α sonorant]. The model advocated by Archangeli and Pulleyblank is committed to override markedness, whenever this suits Lexical Minimality; otherwise the very hypothesis of markedness reversal would be unintelligible.

25 Levin (1987) has proposed that, at a certain point in the derivation, roughly at the onset of the postlexical phonology, redundancy rules cease to apply. Levin suggests that the difference between epenthetic vowels and "excrescent" vowels – predictable brief vowels lacking a definite quality – follows from the fact that the latter are generated *after* the last redundancy rules have stopped applying. Pulleyblank's (1988a) analysis of Yoruba *i* as a null segment is incompatible with this suggestion. Yoruba *i* must be placeless, according to Pulleyblank, well into the postlexical component, which means that it undergoes the redundancy rules at a point where Levin's theory predicts no redundancy rule application. In any case, the proposal that late-inserted segments fail to undergo redundancy rules is irrelevant to the discussion in the text, since the permanently placeless segments we discuss are either underlying (as in the case of most *h*'s and French schwa) or are derived lexically (English schwa).

26 The terms *contrastive* and *distinctive* are used interchangeably here.

27 Hamilton's analyses of Gooniyandi and Gaagudju diverge considerably from those presented here. An entirely different approach to these patterns is discussed in section 3, when we take up again the subject of positional neutralization.

28 Kaun (1993b) and Casali (1993b) have discovered numerous phonological patterns comparable with that of Tigre, in that a three-way underlying contrast between front, central, and back/rounded vowels gives rise to harmony rules that affect only the central vowels: *uCɨ* → *uCu*, but *uCi* remains *uCi*. In the cases considered by Casali and Kaun, only the feature [round] spreads, but it is clearly the case

that front vowels *i* and *e* are specified as [–back], since they do not undergo rounding harmony. It is their palatality that blocks the application of rounding: in the relevant languages (S. Paiute, Chemehuevi, Miwok, and Nawuri), front rounded vowels are disallowed. The harmony undergoers are phonetically central, rather than back unrounded, and this fact represents a violation of principle of Full Specification entirely comparable to that found in Tigre. On the phonetic side, Choi (1992) has presented extensive acoustic evidence for the surface lack of [back] articulatory targets in the central vowels of Marshallese.

29 Further arguments of this type are reviewed in Steriade (1987b).

30 A very different conclusion is presented in Paradis and Prunet (1991) and the volume in which it is found, a collection of studies based almost entirely on the assumption that alveolar coronals are underlyingly placeless. A critical review of most arguments for coronal placelessness presented there has appeared in Kaun (1993a), who demonstrates that they are invalid as formulated. See also McCarthy and Taub (1992).

31 Ringen (1988a) assumes that [–round] spreads in Hungarian. The appearance of [–round] harmony in Hungarian stems from alternations like *bab-hoz, hold-hoz, tüz-höz* versus *hit-hez, fej-hez* (discussed in extensive detail by Kornai 1987). It appears that suffixes like *-hoz* are basically round and undergo derounding when preceded by the unrouded front vowels. The obvious alternative to derounding is to claim that the vowel in *-hoz* is invariable as to height – a mid vowel which must stay mid – but carries no specifications for backness and rounding. This vowel will surface as *o* when preceded by *a*, because it undergoes backness harmony. Any mid back vowel must be rounded in Hungarian: there is no ʌ. In this case, therefore, the [+round] value is due to the vowel's backness and to the fact that it must stay mid. In the other cases (*hold-hoz, tüz-höz* versus *hit-hez, fej-hez*) rounding and/or backness come from harmony. There is no need for assimilatory derounding here, or elsewhere. The Yaka facts described by Goldsmith (1985) as supporting [–round] spreading are similar to those of Lamba cited in (12) above. The phenomenon is discussed below in section 2.3.3.5(4); the assimilatory effect of [–round] is by no means obvious. I am not aware of other candidates for [–round] harmony. No disharmony effect known to me mentions [–round].

32 There exists, however, evidence – carefully assembled by Cohn (1990) – that oral segments may possess specified articulatory targets for a raised velum position. Cohn asserts, on the basis of this evidence, that [–nasal] values are needed. But the connection between phonological features and articulatory targets is far from obvious, as Keating (1990) points out. While we must conclude from the absence of a target that the corresponding feature value is also missing, presence of a phonetic target can be interpreted in ways consistent with phonological underspecification.

33 One may interpret Clements (1988) and Calabrese (1988) as invoking this general possibility. Language specific privativity was also suggested by Goldsmith (1987) for [round] and, more lamely, by

Steriade (1987b) for [voice] and [back]. The case of [voice] has since been dealt with by Mester and Itô (1989) and by Lombardi (1991).

34 This is in effect the view, defended by Czaikowska-Higgins (1987), Halle (1989), and Goad (1991). [ATR] and [R(etracted) TR] represent two distinct privative features, rather than two values of the same feature.

35 Languages like Ateso (Trigo 1991) and Akan-Twi (Painter 1973; Pike 1967) support this interpretation in that they display a phonologically active [+ATR] feature and have [–ATR] vowels articulated with the tongue root in neutral rather than retracted position. In other languages with ATR harmony, such as Igbo (Trigo 1991), all articulatory elements involved in the ATR distinction assume polar values rather than extreme vs. neutral. We may atribute this type of phonetic realization to a language-specific enhancement strategy: the phonologically recessive value [–ATR] is realized with active retraction in order to maintain a clearer distinction from the active [+ATR] value. In yet other languages, such as Anum (Painter 1971), the [–ATR] vowels are realized with active retraction, while the [+ATR] set is realized with the tongue root in neutral position. For such cases, however, there are no phonological arguments establishing that [+ATR] is phonologically active. The facts reported by Painter are consistent with the interpretation that Anum has a harmony based on [Retracted Tongue Root]. The hypothesis advanced here – that the activity of the tongue root is reflected by two independent and privative features [ATR] and [RTR] – finds an interesting parallel in Maddieson and Ladefoged's (1985) analysis of "tense/lax" distinctions in some Lolo-Burmese languages spoken in China. Maddieson and Ladefoged show that for these languages the relevant distinctions involve two sorts of contrasts: lax (breathy) phonation vs. modal phonation and tense (laryngealized) phonation vs. modal phonation.

36 Pulleyblank (1992) mentions the following consideration as an argument against replacing one binary feature like ATR with two privative features of retraction and advancement. The two values of one feature F may not cooccur within a segment as a matter of logic: one entity cannot be simultaneously assigned to two opposite categories, such as [+ATR] and [–ATR]. But two distinct privative features should be able to freely cooccur: therefore we must find some alternative way to explain why segments cannot be simultaneously specified as ATR and RTR. This difficulty disappears when we note that many feature specifications are mutually incompatible, mostly for reasons of articulatory mechanics: one clear example is that of aspiration and glottalization, distinct features on any account, which never reside within the same segment.

37 The following discussion appears in abbreviated form in Steriade (1993b).

38 My colleagues in the UCLA Phonetics Laboratory inform me that little, if any, experimental evidence bears on Chomsky and Halle's claim (1968, p. 300) that "just prior to speaking the subject positions his vocal tract in a certain characteristic manner" which differs "from the configuration of the vocal

tract during quiet breathing," i.e., on the claim that the neutral position is a directly observable articulatory state. This does not invalidate however the usefulness of this notion. It is possible that the speaker may conceptualize his articulatory activity by reference to an abstract neutral position. It remains, of course, to understand why this reference point, whether directly observable or not, is chosen. I should note that phoneticians make extensive use of the assumption that some neutral position can be identified, for at least certain articulatory parameters; for instance, Lindblom's (1986) or Maddieson's (MS) discussion of articulatory effort presupposes some notion of rest or modal position.

39 In the case of [round], the neutral position is clearly intermediate between spreading and pursing the lips, however it may be the case that only pursing has, independently of any other articulatory gesture, a sufficiently salient acoustic effect.

40 The distinction between trivial and non-trivial underspecification is made in Steriade (1987b) and further discussed, under the name of inherent underspecification, by Archangeli (1988).

41 To complete the analysis suggested by Kenstowicz we would have to say, in the terms of Prince and Smolensky's (1993) Optimality Theory, that Latin ranks more highly the OCP as it applies to [lateral] than that applying to [rhotic]: lateral sequences dissimilate, whereas rhotic sequences do not. The full constraint sequence required for Latin under such an analysis is Parse [rhotic] ≫ OCP [rhotic] ≫ OCP [lateral] ≫ Parse [lateral]. The rhotic OCP effects are visible in Latin only when we determine whether or not to apply the [lateral] dissimilation. The idea to consider Parse as a separate function for each feature derives from work by Kirchner (1993).

42 Kaun (1993a) discusses other analyses, primarily Shaw's (1991) account of Tahltan sibilant harmony, from which *t* emerges as place-unspecified. Kaun demonstrates that it is not possible to maintain such an analysis consistently and that the only necessary assumption is that *t* is unspecified for anteriority. It is still unclear, given the lack of phonetic information on the articulation of Tahltan coronals, how these facts can, on Shaw's or Kaun's analysis, be made consistent with the suggestions made in the text.

43 See Westbury (1979) and Trigo (1991) for a review of this issue.

44 Rice and Avery (1989) as well as Piggott (1992) have recognized that voicing in sonorants is a phonologically different phenomenon from obstruent voicing. They represent the two features as, respectively, Sonorant Voicing (SV: a class node dominating [nasal], [lateral], and other sonorancy inducing-features) and [voice]. This proposal, whose spirit is close to that presented in the text, has the drawback that it requires implausible adjustments for the cases in which sonorants and voiced obstruents pattern together. Rice and Avery propose to attribute to the voiced obstruents of languages like English an SV node, despite the fact that these segments are genuine obstruents and lack any of the SV dependents that would justify postulating this class node in the first place. I do

not question their observation that obstruents with reduced closures – flaps – are necessarily voiced, in virtue of their manner of articulation. But the nonflapped voiced obstruents of English do not fall in this class.

45 Stevens (1972, p. 56) appears to suggest that the three vowels {a, i, u} have quantal properties: each is characterized by acoustic attributes that remain stable over a range of distinct articulatory configurations. Stevens's data unfortunately does not include a comparison between the peripheral vowels {a, i, u} and nonperipheral ones. A distinct hypothesis about what makes {a, i, u} optimal is Lindblom's (1986) notion of maximal perceptual distance alluded to in the text.

46 See Suomi (1983) and Kaun (1993b) on reasons why it might be advantageous to extend through harmony the temporal span of vowels with relatively poor discriminability.

47 The complete account of Guaraní must mention the fact that voicing is also involved in the licensing of nasality. An additional statement must express the fact that [+nasal] must be licensed *in every segment* by the presence of [+voice]. This condition must be satisfied over and above the condition discussed in the text: voiceless segments do not become nasal through harmony.

48 Languages such as Sundanese (Cohn 1990) where [nasal] originates in a stop and spreads only onto vowels, without creating nasalized nonvocalic continuants, can also be handled rather naturally by combining a direct and an indirect licensing condition: [nasal] must be licensed, in at least some associated segment, by [–continuant] and, in all associated segments, by [–continuant] or by μ. This condition will block spreading onto nonsyllabic continuants. The Sundanese case may also indicate that equating the notion vowel with a moraic position may not be quite right, although it is irrelevant for our purposes. Sundanese *h*, which can be analyzed as an aspirated, nonmoraic vowel, also undergoes nasalization.

49 The fact that rounded vowels are allowed to occur in the stem formative when it crosses a laryngeal suggests strongly that the general case of harmony in stem formatives involves separate spreading of individual features (as argued by Sagey 1986, for Barra Gaelic); whereas the case of translaryngeal harmony exemplified by *khù-ʔu* involves spreading of the entire vocalic segment *onto* the laryngeal consonant and beyond it. Thus in *khù-ʔu* we are dealing with a single *u*, and [round] in this *u* is licensed, because *u* is a root segment.

50 Thanks to Tom Hinnebusch for his help in verifying this generalization.

51 One difference between our approach and that of Itô (1986) stems from the latter's reliance on the Linking Constraint (Hayes 1986). Itô assumes that place features are ill-formed in the coda consonant, even in the case where they are licensed by association to a following onset. Her analysis relies on the idea that such violations are undetectable, in virtue of the fact that they involve multiply linked features. There are technical problems with this assumption. If the Linking Constraint could have the effect of concealing filter violations, it should be possible to spread nasality onto voiceless sounds in Guaraní (cf. n. 48) or

rounding onto stem formatives in Mazahua. The licensing approach presented in the text makes no such assumption; the feature governed by a statement of indirect licensing must be licensed in some position, in which case it is well-formed wherever else it may be associated.

52 This formulation may have to be adjusted if voicing in obstruents involves two distinct aspects, as suggested above, the presence of vocal cord vibrations and the process of pharyngeal expansion.

53 More recent treatments of Yokuts-like harmony systems (e.g., Archangeli and Pulleyblank 1993) distinguish an initial stage where [+round] is floating, a second stage where it associates to the initial syllable, and a third stage where it spreads onto the remaining syllables. Although an improvement over earlier treatments, this kind of analysis is also based on the proliferation of derivational steps which remain unnecessary.

54 This is not to say that there are no unclear cases left unsolved, but simply that the best understood examples of temporary underspecification turn out to involve altogether different principles. I have left unsettled the account of liquid transparency cases discussed in 2.3.2, beyond suggesting that Mester and Itô's account of it should be rejected. I have also not attempted to analyze here the much disputed cases of Hungarian and Finnish vowel harmony (cf. Vago 1976; Kiparsky 1981; Goldsmith 1985; Steriade 1987b; Ringen 1988a; and many others), because they raise a large number of issues that remain unsolved. Finally, I have not even attempted to touch on the question of tonal underspecification, relying in this case on Clements's (1988) suggestion that prosodically functioning features may have distinct behaviors that differ from those of features whose normal span is one segment.

Note added in proofs: Since writing this chapter, I have read Michael Broe's 1983 dissertation from the University of Edinburgh, which contains an extensive critique of the central idea of underspecification theory – that predictable feature values are underlyingly absent – and an alternative proposal for representing dependencies among feature values.