must be diacritically (i.e. arbitrarily) marked as being unstressable. Thus, Crowley analyzes a word such as *tahósi*, with penultimate stress, as having an unstressable vowel in the first syllable, forcing stress to fall on the penultimate. If the antepenultimate is unstressable in a longer word (and there are quite a few words of this sort), stress will fall on the pre-antepenultimate, as in *mólatine* 'man', or *táripenge* 'lazy'. When suffixes are added to words, stress continues to fall on the antepenultimate vowel of the entire word, as we see in (65a); that the same kind of stress shift appears in the forms with 'unstressable' antepenultimate vowels (cf. (65b)) supports the notion that stress is assigned in these forms as with other words. This set of facts suggests that antepenultimate vowels can be marked extrametrical in Paamese, violating the Peripherality Condition.

(65)   (a) ínau   'I'
            ináu-lii   'oh, me'
            inau-lií-risi   'oh, me again'

        (b) mólatine   'man'
            molatíne-se   'only the man'

In this chapter, we have seen some of the basic motivations for, and techniques of, metrical phonology, and have seen how they interact with the analysis of syllable structure. Many important issues have been ignored, or just barely touched upon, including phrase-level stress and intonation, the effect on segmental phonology of the establishment of the foot as a unit of constitutent structure, and the effect of metrical structure on other prosodic systems, such as tone and vowel harmony. Some of these will be dealt with, albeit briefly, in the following chapters; others must, unfortunately, be left for readers to pursue in the literature.[23]

5
Lexical Phonology

5.1   INTRODUCTION

I suggested in the introductory chapter that the theories of autosegmental and metrical phonology are a direct outgrowth of the generative research program developed in *The Sound Pattern of English* (Chomsky and Halle 1968). Another theoretical development of this classical generative enterprise which has been influential in the last several years also deserves our attention: *lexical phonology*.

In the late 1970s and early 1980s, a number of theoretical proposals concerning the relationship between what in pregenerative years would have been called *morphophonemic* rules and *purely phonological* rules were synthesized into a framework called lexical phonology.[1] Some of these proposals are independent of the theories of phonological representation that we have been discussing in this book, but the approaches – lexical phonology on the one hand, and autosegmental and metrical phonology on the other – have had considerable influence on each other in the last few years. Of these interactions, perhaps the most active and fruitful has been that between *underspecification* theory within lexical phonology and the most recent work on autosegmental representation.

Lexical phonology can be divided into two distinct, but related, theories, a theory of phonology and a theory of morphology. We will begin with a discussion of the theory of phonology, focusing on the issues of the Strict Cycle Condition, the Elsewhere Condition, structure preservation, and underspecification theory. Since so much of the literature on the subject assumes a knowledge of English morphology, I present a brief overview of the principles of English stress and segmental phonology that bears on the morphological questions. Then we will look at the motivation for the various divisions of English morphology that have been made, focusing on Kiparsky's conception of the organization of English morphology as outlined in Kiparsky (1982a), and some modifications made more recently. Finally, we shall consider the knotty problem of cyclic derivations, focusing our attention on English, and concluding that the word-based morphology of English establishes a relationship of an

interesting sort between the morphological structure and the phonological structure.[2]

5.1.1   The theory of lexical phonology

Lexical phonology begins with a division of phonology into a *lexical component* and a *post-lexical* component. (These sectors have also been called 'cyclic' and 'post-cyclic,' but the infelicitousness of the latter terms has become apparent more recently, since the lexical phonology may, on most views, contain non-cyclic parts as well.) The post-lexical component has also been called the *phrasal phonology*.

The post-lexical phonology involves two major sorts of rule applications: (i) those operating crucially across word-boundaries or making crucial use of phrasal or syntactic structure, and (ii) those that fill in, specify, or refer to non-distinctive features – the 'sub-phonemic' rules, we might say. For example, the principles discussed in chapter 2 that govern the appearance of the stops and spirants in Spanish have both of these properties: they specify a difference that is not contrastive in Spanish, and they do so on the basis of a phonological (or phonetic) principle that is not sensitive to whether the context material is in the same word or a different word. Thus, we find the stop version of *b*, *d*, or *g* when a homorganic non-continuant sonorant precedes, whether it is in the same word or the preceding word.

The class of lexical rules is also composed of two subtypes. Lexical rules involve, first, those phonological adjustments that are fundamentally occasioned, or triggered, by the juxtaposition of morphemes, such as the velar softening of /k/ in *electri*/k/ when the suffix /-ity/ is added, forming *elektri[s]ity*; cf. (1). We may also include here the shortening that occurs to the stem vowel of strong verbs when the past-tense suffix is added (e.g. *plead/pled*, *feed/fed*, *feel/felt*, perhaps *say/said*).[3] Second, there are those lexical phonological rules that perform the modifications in the segmental structure required when the underlying form fails to satisfy the phonotactic conditions that make a string a well-formed word, such as the condition that the segments all be assigned to well-formed syllables.[4] In this second group we may include all the types of epenthesis rules discussed in chapter 3 above, such as in Turkish or Selayarese.

> (1)   Velar softening
>   k → s/ — i

Lexical rules, as their name suggests, apply within the lexicon, and hence before all post-lexical rules and without reference to any phonological material in neighboring words (see (3) below). The representations
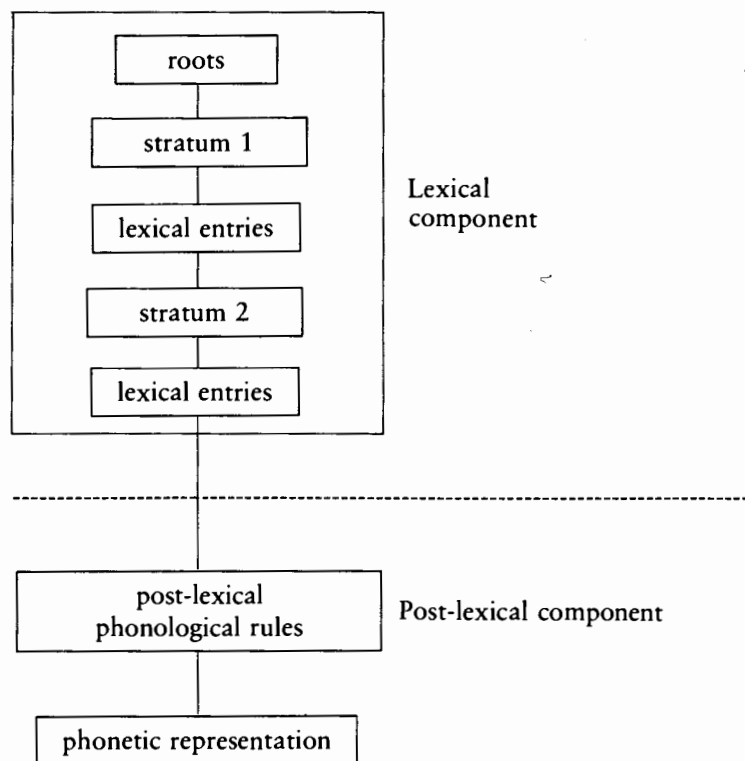
in the lexicon are described using only the *distinctive*, or *contrastive*, features of the language, and hence lexical phonological rules can make reference only to the distinctive features of the language, and to none of its redundant or predictable phonological features, such as the aspiration of stops in English. Similarly, since vowels in English, as in most languages, are non-contrastively voiced, vowels are not marked for voicing in the lexical phonology. It then follows that voicing cannot spread in an assimilatory way from vowels to neighboring segments in the lexical phonology, as has been suggested in various places in accounts that have been offered for post-vocalic voicing of consonants. The principles that lie behind this general move will be discussed further below under the rubric of the 'eliminate redundant features' principle of underspecification theory; we will refer to this as the *first principle of underspecification theory*.

A feature that is simply predictable within a language is barred from the lexical phonology by what we may call *feature filters*, as in (2). In principle, these filters could take in more than a single segment. A post-lexical rule will fill in such predictable features (here, voicing).[5]

> (2)   Conditions on the feature *voice*
>   feature filter: no voice specification with sonorants
>   * $\begin{bmatrix} \alpha\text{voice} \\ +\text{sonorant} \end{bmatrix}$

Features may also be left unspecified underlyingly in another way, and for another reason, only to find the feature specification filled in during the course of the lexical derivation. That is, it may be possible to rule out *underlying* specification of certain features in certain positions, though not by filters of the sort in (2), which are in effect throughout the lexical derivation. Certain kinds of underlying specification may be, and should be, filled in by the effects of lexical phonological rule. Thus a lexical phonological rule, which will typically act across morpheme boundary for reasons we turn to in just a moment, will also have the function of filling in a value of certain unspecified features. Thus in Zoque, for example,[6] there is a rule that voices non-continuant obstruents after nasals (see (4a)). To express the *naturalness* of finding voiced non-continuants after nasals, we can leave the voicing underlyingly unspecified in such an obstruent, getting a free ride on (4a) to do the work of filling in the voicing specification. In fact, this gets at the heart of an important conception: changes that accompany the juxtaposition of morphemes created by the morphology (so to speak) are normally in the direction of what is somehow felt to be a simpler structure even as far as monomorphemic forms are concerned. However, this result would not

(3)   Model of lexical phonology

```
┌─────────────────────────────────┐
│        ┌──────────────┐         │
│        │    roots      │         │
│        └──────┬───────┘         │
│        ┌──────┴───────┐         │
│        │  stratum 1    │         │      Lexical
│        └──────┬───────┘         │      component
│        ┌──────┴───────┐         │
│        │ lexical entries │       │
│        └──────┬───────┘         │
│        ┌──────┴───────┐         │
│        │  stratum 2    │         │
│        └──────┬───────┘         │
│        ┌──────┴───────┐         │
│        │ lexical entries │       │
│        └──────┬───────┘         │
└───────────────┼─────────────────┘
- - - - - - - - ┼ - - - - - - - - - - - - - -
        ┌───────┴────────────┐
        │   post-lexical      │          Post-lexical component
        │ phonological rules  │
        └───────┬────────────┘
        ┌───────┴────────────┐
        │ phonetic representation │
        └────────────────────┘
```

force the post-nasal consonant to be underlyingly unmarked for the feature [voice], something that we might actually want to do.

There is a good deal more to say about this, but in any event, the situation I have just described represents one current point of view regarding underspecification within lexical phonology, and we will return to the matter in section 5.1.5 below.[7] In the meantime, we will use the phrase *the second principle of underspecification theory* to refer to the view that features should be left unspecified if a lexical phonological rule would be capable of filling them in.

Consider the case of the tense vowels in English, which typically (and most clearly in the case of mid vowels) have an offglide. The presence of the *y*-glide in [ey], or the *w*-glide in [ow], is obligatory, and in no way contrasts within English with its absence; there is no tense [e] or [o] without the glide. Because no contrast is possible, the presence or absence is not phonologically distinctive (or contrastive), and hence it is not part of the lexical phonology. The rule or generalization responsible for the

(4)   (a) Voicing after a nasal
$$\begin{bmatrix} -\text{sonorant} \\ -\text{continuant} \end{bmatrix} \rightarrow [+\text{voice}] \ / \ [+\text{nasal}] \ —$$

    (b) Examples from Zoque (cf. (66), ch. 3) (Wonderly 1951: 120)

min 'come' + pa → minba 'he comes'
nʌm 'say' + pa → nʌmba 'he says'
pʌn 'man' + čʌki figure → pʌnǰʌki 'figure of a man'

presence of these offglides is thus a part of the post-lexical phonology, and it follows that no such diphthongs can be created (or found, for that matter) within the lexical phonology.

The notion that only contrastive features play a role in the lexical phonology, and that feature combinations that are not permissible underlyingly continue to be illegal throughout the lexical phonology, may suggest – quite correctly – that the output of the lexical phonology, the last stage at which these generalizations are in force, is roughly comparable to the phonemic level of traditional structuralism. This is in several ways a useful parallelism to bear in mind, though in at least one major respect the structuralist conception fails to match up: the lexical phonological representation is very much *word-bound*. It cannot, in principle or in practice, include in its domain of generalization any material outside the word. Thus, as rules of morphophonemics in pregenerative structuralist phonology could assign a given underlying morphophoneme to any of several distinct phonemes, but not to any particular allophone, so too lexical phonological rules can have as their output an element that satisfies the constraints of the lexical phonology, even though in its actual phonetic form it may have to undergo some post-lexical rules that modify it in further ways.[8]

5.1.2   The Elsewhere Condition

Work on lexical phonology, following Kiparsky (1982a,b), has put to considerable use a general principle which he has called the *Elsewhere Condition* (cf. Kiparsky 1973).[9] The Elsewhere Condition states, in informal terms, that, when two principles of operation are in conflict at a certain point in the derivation, then the one whose domain of operation is more restricted has priority of action. For example, given the statement that all obstruents are voiceless in a given language, and the statement that all affricates are voiced, the latter will have priority over the former in the case of a particular alveopalatal affricate. Even though the affricate is indeed an obstruent, the effect or application of the first principle is

overridden by the second, more specific, statement, since it concerns only affricates. Thus the more general statement must content itself with applying only 'elsewhere' – elsewhere from where more specific statements hold sway.

In some cases, the two principles that may come into conflict in this way – a conflict that the Elsewhere Condition serves to resolve – may be simply rules of the language. Typical cases of this sort are pairs (or sets) of mutual bleeding rules. For example, Lardil (see Hale 1973) has a rule deleting a vowel after another vowel (e.g. *wanka* + *uṟ* becomes *wankaṟ*), and a rule inserting *w* between *i* and *u* (e.g., *keṉṯi-uṟ* becomes *keṉṯiwuṟ*). The latter rule applies in a more specific environment, and hence takes precedence over the former, more general, rule.[10]

Kiparsky (1982a,b) suggests that lexical entries can enter into an Elsewhere relationship with phonological rules. This means that when there is a conflict between what is specified in a given lexical entry and what a rule wants to do to that lexical entry, then the more specific of the two will win out, and the more general will cede. Clearly, suggests Kiparsky, the lexical entry is the more specific of the two. That may sound, at first blush, like saying that no phonological rule ever gets to apply to anything, but that is not so. It is true that the rule of velar softening in (1) will not be able to apply to a word like *king*, by this reasoning; but that is just the effect that Kiparsky is trying to derive – the effect that lexical rules will not undo segment combinations inside of morphemes. Velar softening will not be blocked from applying to cases where the *k* and the *i* are in separate morphemes, as in *elektrik-ity*, since, according to Kiparsky, *elektrik-ity* is not, in the relevant sense, a lexical entry. *elektrik* is a lexical entry, yes; we might say that the suffix *-ity* is a lexical entry (though in Kiparsky (1982a), he chooses not to say that); it may even be the case that the combined production *elektri[s]ity* is a lexical entry; but *elektrik-ity* is not, never has been, and never will be a lexical entry. Hence even augmenting the Elsewhere Condition to the point where lexical entries can block the effects of lexical phonological rules will nonetheless leave those phonological rules free to apply across morpheme boundaries (or, for that matter, any time the form has been crucially modified with respect to what its underlying, or lexical, form was).

Lexical phonological rules, then, will not be able to win in a conflict with a specification in a lexical entry, but nothing will bar them from applying to *fill in* a feature specification that was left unspecified for one reason or another. If the lexical entry is silent with respect to the feature specification of a given segment, then the lexical rule can apply even strictly within a morpheme, for the Elsewhere Condition is not going to block it from doing so – nothing will block it. Hence lexical rules should

have the property of being able to change feature specifications when they apply across morpheme boundaries, but they may fill in feature specifications of segments left unspecified even within single morphemes; for example, the voicing of a stop after a nasal in the same morpheme in Zoque can be filled in by the rule in (4a). These points are often summarized in the following way: lexical rules are feature-changing in derived environments, but have only a filling-in function in non-derived environments. The first part of that statement is also known as the *Strict Cycle Condition*. The name is a bit cryptic, but it alludes to a conception of these facts as resulting from a principle that a rule must apply to phonological material at the first chance – the first cycle – or else for ever hold its peace, and must never return to that earlier cycle to have an effect. We will return to the notion of the cycle below, in our discussion of strata.

Kiparsky proposes, then, that lexical rules and lexical entries enter into an 'Elsewhere relationship' during the lexical phonology – more specifically, within each stratum of the lexical phonology. The Elsewhere Condition would be dropped, he suggests, as a principle relating the final lexical entries with the post-lexical rules. Thus a post-lexical rule would be just as capable of *changing* a feature it found while trying to apply strictly within a word or morpheme as it would be to make a change across a word or morpheme boundary.

5.1.3 Structure preservation

The notion of *structure preservation* is an important one within lexical phonology. It is based on the idea that there are constraints on possible underlying segments in the inventory of a given language, and constraints on possible autosegmental associations, and that the same constraints that apply to underlying representations hold throughout the derivation during the lexical phonology (and that these constraints are then dropped during the post-lexical phonology).

The particular conception of structure preservation presented in Kiparsky (1985) is one whereby rules whose output would otherwise violate a constraint on permitted phonological structures in the lexicon are prevented from applying. If there were a constraint against voiced labials in a language as in (5), for example, with voicing on a separate tier, then a voicing assimilation rule as in (6) would be blocked from applying to a sequence *p-d*, for example.

(5) $*\begin{bmatrix} +\text{anterior} \\ -\text{coronal} \\ +\text{voice} \end{bmatrix}$

(6)   [−son]   [−son]

                    [+voice]

We may refer to this interpretation of structure preservation as *change-inhibiting* structure preservation, in the sense that, by including this principle in our grammar, we will insure that fewer featural changes, or changes of autosegmental structure, will occur than would have been the case without it. Any rule that would create a violation will be blocked from applying. This is not the only possible interpretation of the basic notion of structure preservation in the lexical phonology, and we will briefly explore some alternatives below, whereby rules that create changes must ensure that additional changes be made so as to guarantee, as far as is possible, that representations respect the various positive and negative conditions placed on them within the lexical phonology.

It is, in fact, difficult to provide clear examples of the change-inhibiting interpretation of structure preservation, in part because it becomes more important than ever to justify the precise formulation of the phonological rule(s) involved. To illustrate this, we will look in some detail at a case study that has been offered in support of change-inhibiting structure preservation, and suggest that alternative views of the data are preferable, views that leave no work for this version of structure preservation. The point is an important one, bearing directly on the question as to precisely how rules apply.

Kiparsky (1985) presents the following analysis of Catalan in detail, based on work by Mascaró (1976).[11] He would like to show that a *lexical* rule of nasal assimilation is responsible for the homorganicity of nasal+consonant sequences, such as *bi*[n̪t]*ε* 'twenty', where both consonants are coronal. However, the [n̪] there is not just coronal: it is also dental, a non-distinctive feature of all Catalan *t*s, which the *n* has taken on derivatively, because nasals are thoroughly homorganic with a following consonant in Catalan. Liquids, on the other hand, are alveolar but again, not distinctively so. Thus, a phonetic contrast between alveolar coronals and dental coronals exists, but not distinctively or contrastively. Since the dentality of *t* is not distinctive, Kiparsky suggests, that feature could not have been assigned by the *lexical* assimilation rule: the feature that distinguishes dentals from alveolars is not yet permitted in the lexical phonology, since it is not distinctive. There must be two stages of nasal assimilation in our derivation. One assimilation rule is arguably lexical, on Kiparsky's view, because it applies before a word-level rule that deletes certain word-final consonants in clusters; this sequence of rules, applying in a feeding order, creates the forms shown in (7).

(7)   (a) /bint/   [bin]   *vint*   'twenty'
       (b) /bɛnk/   [bɛŋ]   *venc*   'I sell'
       (c) /kamp/   [kam]   *camp*   'field'

The fact that the [n] in (7a) is alveolar rather than dental is a crucial aspect of Kiparsky's discussion. The word-final *t* has been deleted at word-level, leaving behind a nasal that has been specified for *part* of its point of articulation, the part that is distinctive within the Catalan lexicon. The nasal in *bint* is thus coronal and non-high; it is distinct from *m* and *ñ*, not to mention from *ŋ*. But it is not yet specified as to whether it is dental or alveolar. This is not specified until the post-lexical default rules have a chance to apply. They assign coronal obstruent stops to a dental point of articulation, an assignment shared with a preceding nasal; but a coronal nasal (other than /ñ/) that is not part of a complex segment is always alveolar.

So far, while this is an interesting treatment, it is an argument for underspecification, not for (change-inhibiting) structure preservation. No rule has failed to apply because its output violated a lexical constraint. This is the next, and for our current purposes, crucial, part of the argument.

Kiparsky proposes that the rule of nasal assimilation in Catalan is as in (8).[12] In these structures, as on several occasions before, we have put the point-of-articulation features on a separate tier, and as before, we note that we will return to this matter in chapter 6. Rule (8), a lexical rule, is

(8)   nasal   [nasal]
                       |
       skeleton   C_____C
                              |
       P of A   [α P of A]

intended to account for the obligatory homorganicity of nasal-consonant sequences. Kiparsky explains that he understands his formalism *not* to apply in case the skeletal position on the left is already associated to any features or autosegments that are incompatible with the change (by disagreeing, for example), a matter that becomes crucial in the post-lexical functioning of this rule, as we shall see.

One significant problem to focus on, however, is that (8) does not function to rule out lexical sequences of non-homorganic nasal+consonant sequences, since they would be unaffected by the rule.[13] To the contrary, Catalan nasals simply cannot bear a distinctive point of articulation underlyingly when preceding a consonant word-internally. What we *really* need is two things: a restriction against assigning an

underlying point of articulation to a nasal that precedes a consonant,[14] and a rule spreading point of articulation to just such nasals. We will return to this reanalysis in a moment.

Kiparsky suggests that there is a lexical restriction on the point of articulation that can associate with a nasal, disallowing any velar nasal ŋ that is not in a nasal-consonant cluster, i.e. that is not immediately followed by *k* or *g*. He proposes the formula given in (9). This is the place

(9)   * [−back]
           ‡
           [+high]
           |
           [+nasal]
           |
           C   in an unlinked matrix

where (change-inhibiting) structure preservation as such comes into the picture. The rule of nasal assimilation (8) will be tested to see whether its application may be blocked – to see if its output violates the condition in (9). In some cases, rule (8) applies at the phrase level (see (10) below); in other cases, involving *ñ* and *m*, it does not. The question arises as to whether preservation of the structure in (9) is the reason for its non-application in this latter class of cases.

But is (9), indeed, a proper statement of a structural property that must be maintained in the lexical phonology? Is it the proper way to indicate that ŋ can appear only when followed by a velar consonant? I will suggest a more appropriate alternative below, but it is fair to note that there has been little serious effort, in either the generative tradition or any other, to develop a theory of natural phonological constraints on representations. If Stanley (1967), for example, offers certain formal ways of representing dependencies, it can hardly be said that an effort was made in the formulation to provide a means to represent commonly found constraints in an especially simple and direct fashion. In this book, we have attempted to take some steps toward providing such a formulation, with the {minimum,maximum} specifications of autosegmental association discussed in chapter 1, and, more importantly, the concept of autosegmental licensing.

We shall suggest that the important structural properties maintained in the Catalan representations derive from licensing properties of the Catalan syllable. If this is correct, and if we are correct in taking these licensing conditions to be phonotactics of the word-level representation, then we do not have here a case of a condition that must hold pervasively throughout a derivation, serving as the basis for the inhibition of a rule's

application. These syllable phonotactics will, instead, hold at the word level, and will only indirectly influence what is likely to arise as an underlying form.

We may briefly review the structure given in (9) that Kiparsky suggests is preserved through the derivation. He indicates that the filter should block structures in which there is no [−back] autosegment associated with a [+high] nasal consonant. It is the second principle of underspecification, the one eliminating the unmarked value of a distinctive feature from underlying representations, that necessitates this backwards statement, i.e. prevents a positive statement that such [+high] nasal consonants must be [+back], i.e. velar. Put another way, the second principle of underspecification prevents us from saying directly that velar nasals must be in a doubly linked matrix, which is what Kiparksy suggests is the crucial factor. We return to a licensing approach to the matter below.[15]

In any event, Kiparsky proposes two lexical rules: (8), an assimilation rule, and (9), a filter.

There is now the phrase-level nasal assimilation process to explore. The facts, briefly, are these, following Kiparksy (1985: 95): an *m* hardly assimilates – it assimilates only to a following *f* (10a); an *n* assimilates to any following point of articulation (10b); *ñ* does not assimilate at all, nor does ŋ.

(10)   (a) assimilation of *m*
               so[m] amics   'we are friends'
               so[m̟] feliços   'we are happy'
               so[m] dos   'we are two'
          (b) assimilation of /n/
               so[n] amics   'they are friends'
               so[m] pocs   'they are few'
               so[m̟] feliços   'they are happy'
               so[ŋ] grans   'they are big'

Now, not all conceivable combinations of points of articulation of a nasal and a consonant can arise across a word boundary. The data available suggest that the generalization is that only the *n* – the nasal literally unmarked for point of articulation, thoughout the lexical phonology[16] – assimilates with respect to the contrastive points of articulation, as Kiparsky argues. The only other post-lexical assimilation is 'subphonemic', i.e. not involving the lexically contrastive features of the language (the difference between a labial and a labio-dental, in this case).

However, the analysis employing change-inhibiting structure preservation achieves this end only indirectly. Kiparsky assumes (and I agree with

him here as well) that rule (8) applies in both the lexical and the post-lexical phonology; but if we adopt such an approach, we may need to account for various *non*-occurrences of nasal assimilation at phrase level. Kiparsky suggests, for example, that the reason the labial *m* does not assimilate to a coronal at phrase level in Catalan (we find [som dos] 'we are two', for example, with no assimilation of the *m*) is that nasal assimilation is ordered before the rule that fills in the point of articulation of *d*; hence there is no point of articulation that *can* spread to the *m*. *m* does not associate to a following velar across word boundary, either, but under the analysis in question this would be for a different reason, not because of the extrinsic counterfeeding order just mentioned. In this case, the structure *m#k* would be as in (11), to which nasal assimilation (8) will not apply, Kiparsky suggests, 'because the configuration [+high, +labial] is banned by a marking condition both lexically and postlexically'. Though we have seen a suggestion here that a rule is blocked because its output violates a condition, this is not yet structure preservation, since the condition is not one that can be (or at least, has been) motivated as part of the lexical phonology of Catalan.

(11)   [+nasal]
             | blocked
             C..........×.......... C
             | |
          [+labial]   [+high]

          /m/   /k/
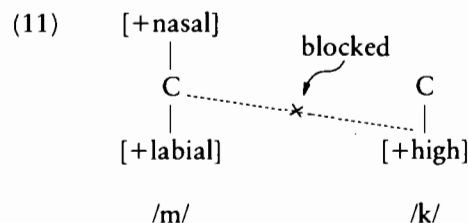
Kiparsky suggests that filter (9) can be appealed to account for the last remaining example of non-application of nasal assimilation at the phrase level, that which does not occur when ŋ is followed by a palatal, such as λ. Here Kiparsky suggests that the reason such assimilation fails to take place is quite different. On this analysis, if lexical filter (9) did not exist, then at the point in the phrase-level phonology when nasal assimilation was about to apply, we would find the structure illustrated in (12). Although nothing within this account would block rule (9) from applying, the suggestion is made that the post-lexical, phrase-level rules are

(12)   [+nasal]
             |
             C..................... C
             | |
          [+high]   [+high]
                                  [−back]
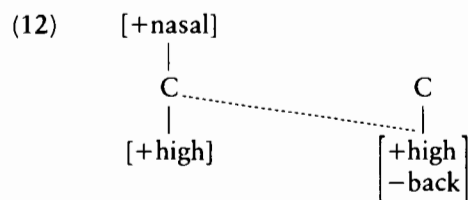
extrinsically ordered in such a way that nasal assimilation precedes (and is counterfed by) the rule that deletes a word-final velar. That is, any word that appears to end in a velar nasal when pronounced in isolation (such as [tiŋ] 'I have', or [baŋ] 'bank') must actually end in a stop (/k/ or /g/) through to the final level of the lexical phonology by structure preservation, and hence at least part way into the post-lexical phonology. If we assume that the rule that finally gets rid of it follows nasal assimilation, then the velar stop can be viewed as a phantom segment intervening between the nasal that is almost, but not quite, at the end of one word, and the consonant at the beginning of the next.

When all is said and done, then, it is reasonably clear that this particular version of structure preservation does not do a great deal of work that could not be done better another way. In chapter 3, we investigated how restrictions on point of articulation should be treated, and concluded that striking and basic patterns could be understood in terms of the ways in which languages can license a point of articulation autosegment. In a quite a few languages, the coda can never license a point-of-articulation; in Catalan, the coda can license a point of articulation, but the restriction appears to be that a sonorant position (either onset or coda) can never license a velar point of articulation.[17] The rhyme, like the onset, can license only one point-of-articulation autosegment, and thus a sequence of a nasal and a consonant with distinct points of articulation will not be permitted.

This restriction, since it derives from licensing, holds at the word level. Thus (13a) is ill-formed there, since one of the point-of-articulation autosegments would not be licensed; (13b) is well-formed.

We suggest that the crucially operative principles at work in Catalan are the following: (i) a rule of nasal assimilation that provides a point of articulation only to nasals that are unspecified for a point of articulation, formulated much as in rule (8) above, though applying in a fashion discussed below; (ii) a rule that specifies that adjacent consonants with identical points of articulation *share* a point-of-articulation autosegment;[18] (iii) post-lexical specification of non-contrastive features. All three of these points are already present, though in slightly different form, in the account Kiparsky offers.

The filter in (9), however, must be replaced by conditions on licensing in a coda. Quite generally, licensing considerations at word level will block any nasal that is specified for a point of articulation if there is another consonant specified for point of articulation in the same coda, since only one point-of-articulation specification may appear licensed by any given licenser. All other things being equal, a nasal that is alone in its coda will take on any point of articulation that the coda licenses, in Catalan just as in English, via rule (8).

(13)   (a)

(b)
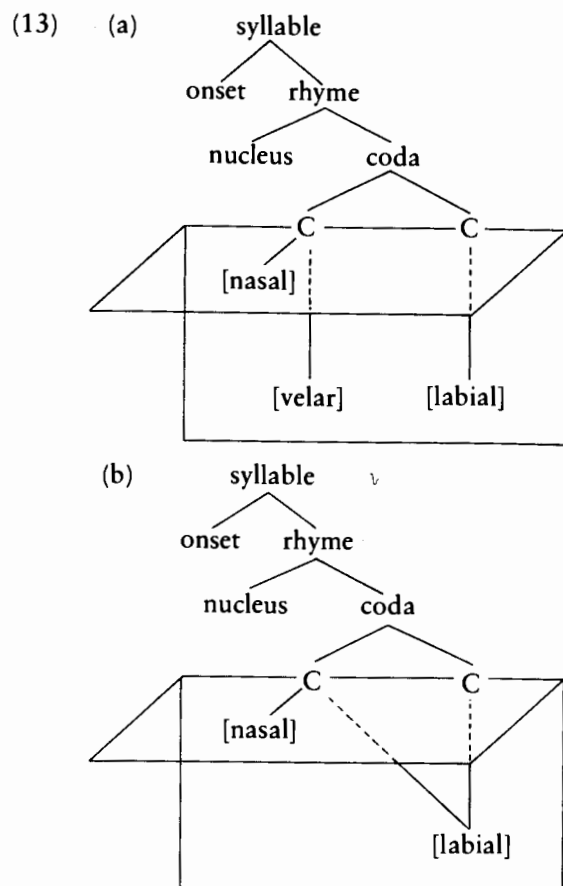
In Catalan, as in English, there is no underlying segment ŋ. That is to say, the structure in (14a) is possible (where the broken line associating [velar] with the nasal position is not present underlyingly, and is filled in only later), though that in (14b) is not. As we have seen, and as (14) illustrates again, the notion of licensing explicates how segments that are not allowed as underlying segments may arise by assimilation – even within the lexical phonology.

The use of (9) to block nasal assimilation via a change-inhibiting view of structure preservation involved explanations of why (8) did not apply to the *m*, the *ñ*, and the *ŋ* that were produced through the lexical application of nasal assimilation. Kiparsky suggests that the proper conception of rule application is one in which an assimilation rule, of the sort we have seen in (8), will apply equally to underspecified segments and fully specified segments, being blocked only by the sorts of conditions as in (9).

(14)   (a)

(b)
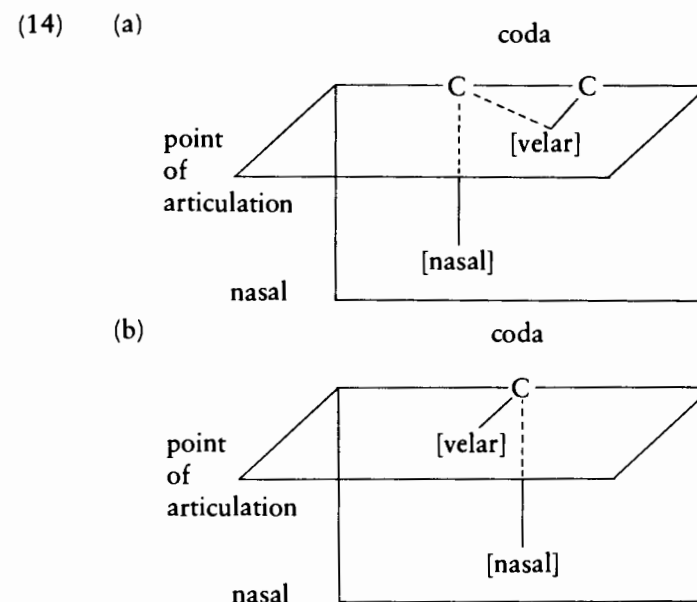
Our view, however, is somewhat different. As we will explain in more detail below, we suggest that a large range of phonological rules are best understood when formulated so as to apply just in case their application serves to create a well-formed representation out of one that was ill-formed.[19] If we distinguish here simply between an underlying representation, a level representing essentially the output of the lexical phonology (the word level), and the phrase level, we have a picture as in (15), with phonotactics of the sort in (16) defined at the latter two levels of representation, the word level and the phrase level. The establishment of phonotactics as in (16) makes it possible for the phonotactics to then trigger the application of such a rule as in (8). (16) makes explicit certain properties of the skeleton/point-of-articulation tier chart, the chart on the horizontal plane in (14). This phonotactic indicates that all [+consonantal] segments on the skeleton are freely associating segments in the point-of-articulation chart, and that at the phonetic level they must associate with a point-of-articulation autosegment. Put simply, (16) says that all consonants must be specified for point of articulation; rule (8) helps nasals that are in violation of that phonotactic to come into line with it.

(15)   Underlying representation
        ‖ ← rule (8) may apply
        Word level: phonotactic in (16) is applicable
        ‖ ← rule (8) may apply
        Phrase level: phonotactic in (16) is applicable

(16)   Word-level and phrase-level phonotactic
Skeleton/nasal chart
Freely-associating segments: {+consonantal, ∅}
Minimum/maximum association: skeletal tier: (1,1)

Finally, we assume that, when a word-final *m* (a labial nasal) is followed by a word-initial labial consonant, as in the second example of (10a), the two consecutive, identical point-of-articulation autosegments merge to form a single autosegment, as illustrated in (17b). When the non-contrastive feature specifications are filled in, marking the continuant as a labio-dental, this information regarding specific point of articulation is shared by both consonants, the nasal and the continuant.[20]

(17)   (a) Sequence of *m–f* across word boundary
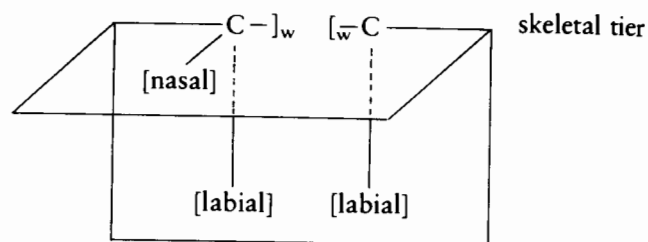
(b) Merger of two adjacent, identical consonants
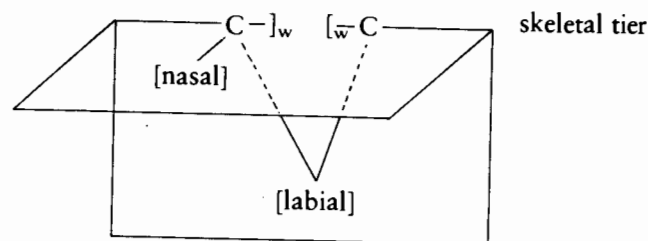
(c) Post-lexical specification of labio-dental
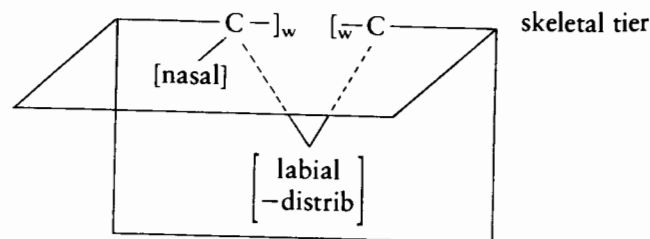
It is worth pointing out that there is a close relationship between underspecification and the conception of structure preservation. In a framework in which segments are fully specified, if any single feature is changed, then the chances are good that the result will not be a permissible underlying segment; phonological systems are rarely so symmetric. For example, if, when voiced, an *s* becomes not *z* but *r*, then all the other features other than just voicing that distinguish an *r* from a *z* (e.g. [sonorant]) will have to be changed as well, if structure preservation is to be maintained. On the other hand, if only a small number of distinctive features are specified, then rules that take away feature specifications will always (or almost always) be structure-preserving (since even the totally unspecified segment is going to become something, either a *t* or a glottal stop, or whatever, as the language chooses). Rules that add features will also tend more frequently to be structure-preserving, especially if they are applying to a highly underspecified segment already.

Consider the case of Klingenheben's Law in Hausa, for example, according to which a *p* or *k* in coda position becomes a round glide *w*.[21] If *p* and *k* are underlyingly specified as [-continuant] and [-voice], then these features will have to be modified in the process; if the features are underlyingly unspecified, and filled in only post-lexically, then the deletion of the point-of-articulation features, and their replacement by the feature [+round] (assuming that that is the correct statement of the change) is all that needs to be said, as in (18). Note that the change here involves actual replacement of the features on the left.[22]

(18)   [−coronal] → [+round]

Consider the more complex case of Kuman, spoken in Papua New Guinea (Lynch 1983). If we look closer at this system, we find reason to conclude that, in the contrast between a plain *l* and a pre-velarized *gl*, the latter is, surprisingly, the unmarked lateral in the language, and with a circumspect use of underspecification, the alternation between *gl* and *t* can be analyzed as a change of a single feature, [voice].

(19)   p t k
            b d g
               s
            m n
               l gl
            w y

The segment transcribed as *gl* deserves special attention. It is a

velarized liquid, arguably a single segment, since it may occur in in syllable rhyme where no clusters otherwise appear. An underlying *t*, when in syllable onset position, is realized on the surface as [r] except word-initially, where it is realized as [t]; a *t* in coda position is realized as a [t] when an *n* immediately follows, and as an [r] otherwise.

There is a phonological contrast between *l* and *gl*, but there are relatively few *l*s at all in the language. Some are created by rule, from a merger of *gl* and *i*; others occur in borrowings. Lynch suggests that at an earlier stage of the language there was only one liquid, which presumably changed (perhaps in all positions except before *i*) to *gl*. Subsequently, in a relatively small number of words, a contrast has arisen between *gl* and *l*; the dominant majority of liquids are of the velarized form. As we noted, Lynch suggests that, synchronically, some *l*s are created by a fusion of *gl* and *i*. Observe the imperative paradigms in (20) for the stems *yu* 'bring', *pr* 'hear', and *yagl* 'plant'. The singular, dual, and plural suffixes are *o*, *iro*, and *io*, and Lynch suggests that *gl+i* generally becomes surface [l].

(20)

	bring	hear	plant
singular	yuo	pro	yaglo
dual	yuiro	priro	yaltro/yagltro
plural	yuio	prio	yalo

Let us assume that there is presently in Kuman a lexical contrast between the velarized *gl* and the plain *l*, and that an unmarked liquid will be specified as [+prevelarized] post-lexically. An *l* that does not surface as [gl] is lexically specified as [−prevelarized]; that is, that feature *is* lexically contrastive, but the unmarked value is positive, and is relevant only to the liquids.

There are two rules in Kuman that devoice the *gl*, making it a derived *t*. In the case of the output of rule (21a), a quite transparent rule, this surfaces as a [t] (see (22)); in the case of (21b), a much more opaque but still plausible rule, its surfaces as an [r], as *t* will do, as mentioned above (see (22)).

(21)   (a) gl → t / — n
          (b) gl → t / — V gl

The following data illustrate both rule (21b) and a rule deleting *gl* before *b*, as well as several other processes that we will not focus upon: several epenthesis processes, including that of *u* between a *b* and a stop; and voicing of the *k* in the indicative suffix *ka* in an environment that is not entirely clear – after voiced consonants, and apparently after vowels,

(22)   (a) stems with final *gl* with possessive suffixes ('my leg', etc.)

	kagl 'leg'	*yobugl* 'bone'	*mabugl* 'forehead'	*siragl* 'thing'
1 sg.	kat-na	yobut-na	mabut-na	sirat-na
2 sg.	kat-n	yobut-n	mabut-n	sirat-n
3 sg.	kagl-e	yobugl-o	mabugl-o	siragl-mo
non-sg.	kat-no	yobut-no	mabut-no	sirat-no

(b) stems with final *t*

	bit 'head'	*piut* 'chest'	*kobt* 'navel'
1 sg.	bit-na	piut-na	kobut-na
2 sg.	bit-n	piut-n	kobut-n
3 sg.	br-e	piur-ie	kobr-ie
non-sg.	bit-no	piut-no	kobut-no

but before epenthesis. The first-person singular subject marker is *i* with vowel-final stems, such as *yu* 'bring'. There is a thematic *n* in the future conjugations of vowel-final stems. The third-person singular *w* also metathesizes to the right, as we see in (23). In these data, we see several examples of *gl* becoming *t* before *n*, by rule (21a); we also see the stem final *gl* of *yagl* becoming *t* (which surfaces as [r] in onset position) before the future morpheme *-agl*, as when *yagl-agl-ka* surfaces as [yaratnga].

(23)   (a) Singular subject markers

	1st	2nd	3rd
aorist	ø/i	n	w
future	ø	n	b

(b)

		yu 'bring'	*pr* 'hear'	*yagl* 'plant'
Aorist	1st	yuiga	prika	yaglka
	2nd	yunga	pitnga	yatnga
	3rd	yugwa	prukwa	yaglkwa
Future	1st	yunaglka	praglka	yaraglka
	2nd	yunatnga	pratnga	yaratnga
	3rd	yunabuka	prabuka	yarabuka

The question we may now return to is just what rule (21a) really looks like. The change from the phonetic segment *gl* to *t* seems like it comprises a goodly number of featural changes: *gl* is velarized, while *t* is not; *gl* is voiced, while *t* is not; *gl* is a sonorant, while *t* is not. However, if we

specify only the marked values of the distinctive features that specify *gl*, we can simply describe it as [+consonantal, +sonorant]. It is, in the context of the consonant system in (19), the unmarked sonorant, just as *t* is, in that context, the unmarked obstruent: unmarked both for point of articulation and for all other secondary markings, such as nasality and voicing. (Recall that *gl* is taken to be unmarked for velarization, while *l* is specifically marked as [−prevelarized].) Thus the change from *gl* to *t* is minimal: it is the change from sonorant to obstruent. The 'structure-preserving' character of the change followed, more or less automatically, from the underspecification approach, since, at the point in the derivation where the change takes place, the phonological properties of the *gl* which would have to have been eliminated in order to make a *t* out of a *gl* are simply not specified.

Regardless of whether or not rules may be blocked in their application by the principles of structure preservation, other examples do support the conclusion that the output of lexical phonological rules will normally negotiate in order to produce an output that is consistent with the word-level phonotactics of the language. Let us consider two kinds of examples that illustrate a this point.

The first is illustrated by the discussion surrounding example (25) in chapter 1, where it was pointed out that a language (Sukuma, in the discussion there) may have a restriction on the number of association lines associated with a given vowel position, for example. In a language where the maximum is one, if a rule should reassign a tone to the vowel in question, the rule is not blocked from applying by structure preservation or anything of the sort (as Kiparsky's view would predict); as we saw in several tonal examples in chapter 1, the association line added by rule is preserved, and the association line that had previously been present is eliminated in order to maintain the constraint on permissible structures.[23]

Consider another example of a similar sort, this time one involving elements that are less prosodic in character. Shaw (1980) reports a rule of vowel coalescence in Dakota, where the underlying inventory of vowels consists of the canonical five oral vowel system {i, e, u, o, a} and the canonical three nasal vowel system {į, ų, ą}. When a nasalized /ą/ is followed by a front vowel, the result is a high front nasal vowel, and when it is followed by an /o/ (no examples of /u/ are given), a high back nasal vowel is formed. (I simplify the matter of boundaries, which is irrelevant to the point at hand.)

As Shaw points out, this process is remarkably similar to another rule that she motivates, simply deleting an oral /a/ in the same context. In that case, the mid vowels that remain behind, so to speak, do not need to be raised (e.g. /ka + epča/ > [kepča]). But there are no nasal mid vowels in the inventory of the language; so when the /a/ is lost, and its nasality is

(24)   (a) ą + i > į
             cʰą + i + čoγa > [cʰįčoγa]
        (b) ą + e > į
             ptą + ehą > [ptįhą]
        (c) ą + o > ų
             xʔą + o + tʔa > [xʔų tʔa]

preserved (an instance of what we called 'stability' in chapter 1), the result is a nasal mid vowel. But there are no nasal mid vowels in the inventory of underlying or word-level vowels in the language, and so it is not surprising that the result that would independently be expected – a nasal mid vowel – is modified immediately to form a nasal high vowel, which is a possible word-level vowel of the language. If, in line with the discussion of vowel features presented in chapter 6, we take mid vowels to be specified for the feature [low] and high vowels not to be, then the result described here can be viewed as quite similar to the tonal case sketched above. Dakota has vowels that are specified [+round] and [+ low], i.e. /o/, and also those that are [−round] and [+low], i.e. /e/. It also has vowels that are specified as [−round] and [+nasal] /į/; but it has no vowel that is [−round] and [+low] and [+nasal], i.e. a front, mid, nasal vowel. When [+nasal] is associated to a vowel that is [−round] and [+ low], the combination is not allowed, and one of the previously associated elements (in this case, the feature [+low]) is removed, allowing the resulting combination ([−round] and [+nasal]) to remain.[24]

5.1.4 Strata

In addition to dividing the rules of phonology into the two major classes of lexical and post-lexical rules, lexical phonology allows for the lexicon to be divided into what are called *strata* (or, interchangeably, *levels* or *layers*), as we had already seen in the basic architecture given in (3) above. This division affects three major and distinct points. Through the formal use of strata, lexical phonology makes claims about (i) the reanalysis of what once were called phonological boundaries; (ii) the ordering of morphological elements, i.e. affixes; and (iii) the cycle in phonology. We will return to the notion of the cycle in section 5.2 below, and so will not discuss it here, except in passing and to define the basic organization of strata.

Strata are small compartments in which affixation processes and phonological rules come packaged together. They are linearly arranged, so that the first stratum has as its potential input (or domain on which to operate) the monomorphemic roots of the language. Each such root may undergo affixation of one of the stratum 1 affixes, at which point it will

have the opportunity to pass through all the stratum 1 phonological rules, some of which it may chance to trigger.

Now, clearly, the idea that morphological affixation simply 'happens' to a form as it passes through the stratum, as if by chance, is not a very appealing metaphor; but since this is more a theory of the interaction of morphology and phonology than of morphology *per se*, we shall have to bear with it. There certainly is no very sensible way to think of the affixation of various derivational suffixes as being triggered, as it were, by some cloud of meaning that is hanging over the stem waiting to be 'realized' by some affixation process. We cannot attach a meaning to a morphological root, and pass it through a morphological flow chart, expecting it to pick up just those suffixes that allow the expression of that meaning – certainly not in the case of lexical meaning expressed by derivational morphology. In any event, the point is simply to acknowledge that this is not a good way of thinking about derivational suffixation as far as its connection to meaning is concerned, but it will have to do for our expository purposes.

Strata themselves may be either cyclic or non-cyclic, though the former has often been assumed to be the normal case. In a cyclic stratum, as in (25), as each affix is in turn added to the base by morphological processes, the entire set of rules of that layer will have an opportunity to apply. In a non-cyclic layer, the phonological rules of the layer will not apply until all the affixes of the layer have been attached.

As a form passes through a stratum, it is bracketed in a way that shows its history within that stratum (see (26)). These brackets are not labeled,

(25)   Internal structure of a stratum

---
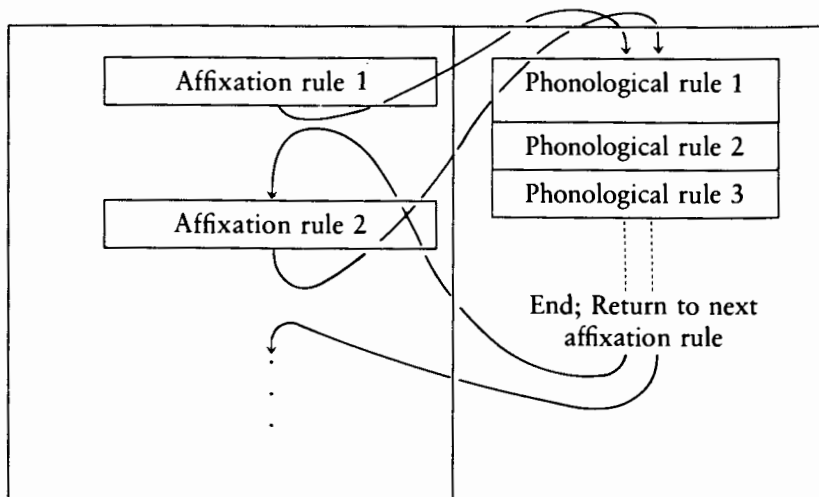
(26)   [[[instrument] al] ity]

except perhaps insofar as a left bracket differs from a right bracket. Phonological rules within a stratum can make reference to brackets and bracketing. (But frequently they will not need to when it might have been thought necessary, because the Strict Cycle Condition will guarantee that most lexical rules apply only across morpheme boundaries.)

At the end of each stratum, all the internal bracketings are dropped, so that rules of a given layer can never make reference to any internal structure that was created on an earlier stratum. To use traditional terminology, if in a given language the stem is created by the first stratum, and the word is created by the second stratum, then the morphology and phonology of the second stratum can make no reference to any morphological boundaries strictly inside the stem. Likewise, the post-lexical phonology can make no reference to any word-internal boundaries. This restriction is called the *Bracketing Erasure Principle* in the literature.

This leads to an important position within lexical phonology: phonological representations do not contain labeled boundaries, such as the + and # of Harris (1951) or Chomsky and Halle (1968). Instead, there is a nesting of domains defined entirely by the morphology, from the root outwards, as in (27), for example, and the rules of the phonology will be

(27)   [prefix [[root] suffix]]

able to look at larger and larger domains, starting with the smallest. The functions served by boundaries in earlier theories are translated into other functions, in other ways. (And of course, the translation is not utterly perfect, in the sense that it is not the case that anything that can be done the one way can be done the other: if this were so, why should it matter which way we did it?)

Let us first classify the principal functions of boundaries in phonological rules. Using the boundary symbols of classical generative phonology, we may list four. (i) Certain rules apply only across morpheme boundaries, and we can express that by writing $X \rightarrow Y / - + Z$. (ii) Certain rules apply only across word boundaries, and we can express this by writing $X \rightarrow Y / - \#\# Z$. (iii) Certain rules apply only at word boundary (typically, word-finally). (iv) Certain rules cannot apply across a word boundary. This last function can include two kinds of cases: one where the rule applies strictly inside words, but not into material that would be introduced with the # boundary, and the other where the rule strictly refuses to apply outside of the phonological word.

These various boundary-oriented phonological effects are treated in different ways in lexical phonology. If we take the morphemes intro-

duced with a + boundary to be introduced on the first stratum, and those introduced with a # on the second, then rules that must apply across a + boundary become lexical rules of the first stratum, and the rule's failure to apply to stratum 1 segments that are not separated by a + boundary becomes the responsibility of the Strict Cycle Condition, discussed above. Rules that apply only across true word boundaries have to have those boundaries explicitly stated; but they will be, necessarily, post-lexical rules, and there is a bracketing in the phonological string that indicates true word bracketing. Thus we could write a rule whose structural description is as in (28), as we have already seen in KiHunde; if applying post-lexically, this means that the forms are separated by word-level bracketings.

(28)   V]_w   CV
         | |
         H   H

On the third point, rules that apply only at word boundaries can be treated in more than one way. They can be post-lexical, with a word-level bracket indicated, just as in (28). However, they can also be rules of the last stratum of the lexical phonology. If the rule in some way involves a feature that is not lexically distinctive, then it must necessarily be post-lexical; however, if it does not, then it may be either post-lexical or lexical.

Finally, a rule that cannot apply across a # boundary (in the older terminology) may be reinterpreted in one of two ways. If it is a rule that applies to stem material, though not across a #, then it may be a stratum 1 phonological rule, as with the rule deleting a schwa before vowel in English. This rule deletes the schwa in *buddha+ism*, but not across a #, as in *Indiana#ism*. On the other hand, the rule may be a stratum 2 rule, if it applies more broadly inside a word, but not outside the phonological word more generally.

One point that on occasion is not clearly made in the literature is that the development of strata has been to some extent a way of formalizing an increasing dependence of phonological rules on particular morphemes without making the point explicitly, and this can be suspiciously, even dangerously, misleading. An extreme case may make that clear. If there were a grammar with as many strata as there were morphemes, then (all other things being equal) whatever phonological rules the affixation of a particular morpheme required could be assigned to the stratum in question. So if, for example, there is a strong past-tense suffix *-t* in English which triggers shortening of vowels (*keep/kept*, perhaps *say/said*, and so on), and we call 'stratum 17' the stratum that contains just that affixation process, then the shortening rule would be made a phonologi-

cal rule of stratum 17, and would apply only to forms derived in that stratum. In this extreme case, phonological rules would be directly linked to particular morphological processes with no explicit and formal indication of the fact.[25]

Lexical phonology embodies the hypothesis that the morphemes that trigger particular phonological processes form a discrete and identifiable set, and that the affixation processes corresponding to them can be ordered together as an uninterrupted series of rules.

The bulk of classical generative phonology is concerned with how morphemes are modified when juxtaposed, and we have already seen, during our discussion of the Elsewhere Condition above, that lexical phonology is designed in order to integrate the proposal that a lexical rule may restrict itself to applying only in a derived environment, which is to say that a rule will refrain from applying in a context where all the conditions for the rule's application are found strictly within a single morpheme (or, more accurately, where all the conditions for the rule's application were already found within that single morpheme in its underlying form).

Interesting and intricate arguments have been made for and against this Strict Cycle Condition. In any event, one of the central principles of lexical phonology has been this limitation on lexical rules to a kind of application that can change a feature specification only when it is explicitly the addition of morphological material that has satisfied the rule's structural description. But this general principle also follows, it should be clear, from a rather different interpretation of the significance of many lexical rules, in particular, the interpretation according to which these rules are explicitly linked to morphological processes.[26]

Let us take a brief look at the structure of the Bantu verb from the point of view of a theory of lexical phonology. The first set of questions involves dividing the phonology into the lexical and the post-lexical phonology. Post-lexical rules in a Bantu language such as KiRundi include the rule that affects adjacent vowels across word boundaries, reducing them to a single, short vowel with the quality of the vowel on the right, a rule whose effects were noted in chapter 1, in connection with example (46) there. Within the lexical phonology, it is necessary to consider whether the phonology or the morphology motivates the division of the lexicon into two or more strata. Bantu languages support a distinction between the processes that build a stem out of a root (or 'radical'), on the one hand, and the processes that build an inflected word out of a stem. These two aspects are described within a lexical phonology with two distinct strata, one for each series of steps.
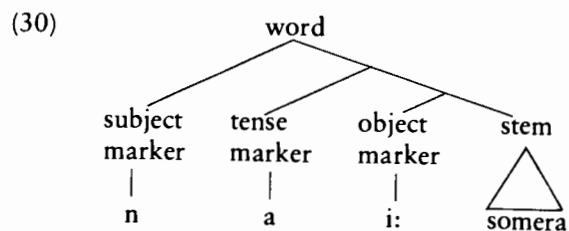
The first stratum consists of the suffixation of derivational suffixes of the form V(C) (called 'extensions', in the Bantu literature) to a radical

that is typically of the form CV(V)C. This series of affixations, with each step the formal expression of the morphosyntactic features associated with the suffix in question, leads to a verbal stem with the structure shown in (29). The domain thus defined is also the crucial domain for the assignment of tense-specific tonal patterns in most Bantu languages; in KiRundi, this includes a L–H (Low–High) melody used in subordinate
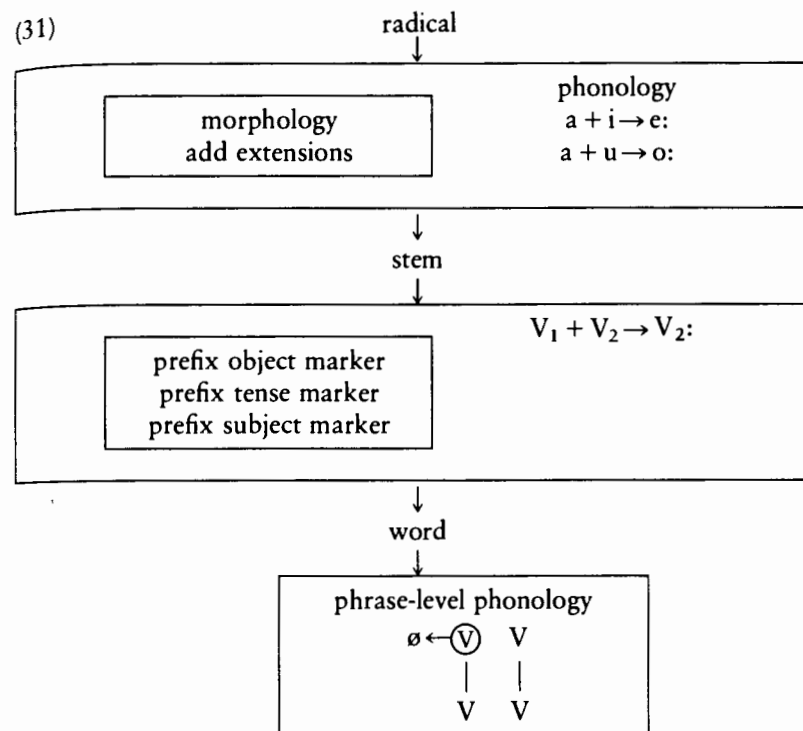
(29)

```
                    stem
              _____/|_____
             /     /    \      \
          root  extension  extension  final
                                      extension
```
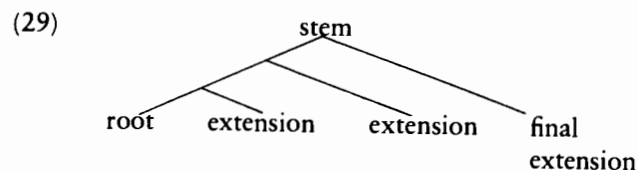
and negative clauses. In addition, when two vowels are juxtaposed within this domain, rather than being affected as they would be in the post-lexical component, they obey a different principle. The result of juxtaposing two vowels is a long vowel here, and if the first is the low vowel *a* and the second is a high vowel (i.e. *u* or *i*), then the quality of the composite vowel is a mid vowel otherwise like the second vowel; thus *a+i* becomes *e*, and *a+u* becomes *o*. For example, the root *ha*, meaning 'to give', plus the applicative suffix *ir*, forms *heer*.[27]

The verb, in turn, is formed by prefixing inflectional elements on a second stratum. When vowels are juxtaposed on this second stratum, a distinct rule is invoked which again creates a long vowel, but one whose quality is fully determined by the vowel on the right of the input string. Thus a sequence *a+i* creates *i:* as its output, as in the sequence *n+a+i:+ som+er+a* [niisomera] 'I read for myself', precisely as in Luganda, a closely related language that we looked at in some detail in chapter 2. The morphological structure of the output of this second stratum will look like (30). By indicating no structure below the node marked 'stem' – i.e. the output of the first stratum – I intend to emphasize the fact that bracketing erasure has made the internal structure of that sequence no longer linguistically visible. Thus a layered picture of the KiRundi verb might look like (31).

(30)

```
                    word
              _____/ _____
             /      |    \      \
       subject   tense  object  stem
       marker   marker  marker   /\
          |        |       |    /  \
          n        a       i:  somera
```

(31)

```
                    radical
                       ↓
        ┌─────────────────────────────────────┐
        │  ┌────────────────┐      phonology   │
        │  │  morphology     │     a + i → e:   │
        │  │  add extensions │     a + u → o:   │
        │  └────────────────┘                  │
        └─────────────────────────────────────┘
                       ↓
                    stem
                       ↓
        ┌─────────────────────────────────────┐
        │  ┌──────────────────────┐   V₁ + V₂ → V₂:
        │  │  prefix object marker │            │
        │  │  prefix tense marker  │            │
        │  │  prefix subject marker│            │
        │  └──────────────────────┘            │
        └─────────────────────────────────────┘
                       ↓
                    word
                       ↓
        ┌─────────────────────────────────────┐
        │     phrase-level phonology            │
        │          ø ← Ⓥ    V                  │
        │              |     |                  │
        │              V     V                  │
        └─────────────────────────────────────┘
```

5.1.5 Underspecification theory

Underspecification theory involves several notions which have appeared in generative and non-generative theory, and which can in principle be accepted or rejected independently of the other ideas of lexical phonology, but which fit comfortably with the central ideas of the theory.[28,29] I have already summarized the two core ideas of underspecification theory with the following principles. (i) First, *eliminate redundant features from the lexical phonology*: features that are not distinctive in a subpart of the the inventory of underlying segments must not be expressed with segments in that subpart – they are ruled out by what we have called feature filters. (ii) Second, *eliminate 'unmarked' feature specification from underlying forms*: in any given phonological context, each remaining (i.e. non-redundant) feature has an expected, or unmarked, value and a less expected, or marked, value; only the latter may be explicitly present in underlying forms. These two principles differ considerably, and represent two different types of underspecification theory. The line between them is on occasion a hard one to find, but if we

step back slightly, we should be able to get a clear grasp of the difference between them.[30]

The first principle is in effect a set of restrictions on feature co-occurrence. We saw in the case of Digo, the Bantu language discussed in section 1.6, that there was a Low tone on all voiced obstruents on the surface. This generalization is so strong and straightforward that we can draw the conclusion that this Low tone is not in any way constrastive: it is totally predictable. The 'eliminate redundant features' principle would then force the conclusion (which we adopted anyway in the analysis of Digo) that the Low tone is not present underlyingly. In a model that distinguishes between a lexical phonology and a post-lexical phonology, it would furthermore be predicted that this non-contrastive Low tone is not inserted until the post-lexical phonology. Similarly, since voicing is not contrastive among the sonorants in English, sonorants could not be marked for voicing according to this principle.

It must be acknowledged that this principle (which seems in many ways just like phonological good sense turned into an explicit principle) is often difficult to apply in practice because one not infrequently finds oneself in a situation where two distinctions (or features) are redundant in a given description, but it is hard to see which one is logically dependent (and thus predictable, and thus to be eliminated from the description in the lexical phonology), and which is logically prior (and thus to be maintained in the lexical phonology). For example, if a language had a vowel system with just {i, e, u, o}, the front/back contrast and the round/unround contrast would be difficult to untangle: which one is contrastive, and which predictable? Is a vowel in such a language round *because* it is back, or is it back because it is round? Typically, one must look deeper into the system to determine the right way to analyze it.

Many cases present no difficulty at all. Voicing, for example, is rarely distinctive among the sonorant segments of a language (the vowels, glides, nasals, and liquids), and it is generally distinctive among the obstruents, or a significant subset of the obstruents – as in English, for example. Among the sonorants in English, then, voicing is non-contrastive, and by the 'eliminate redundant features' principle it would not be specified in the lexical phonology. It follows, then, that no rule of lexical phonology in English can refer to the voicing of a sonorant in English, either in the structural change or in the structural description of a rule; an obstruent could not assimilate in voicing to a sonorant within the lexical phonology precisely because the sonorants are not voiced within the lexical phonology: they are redundantly voiced later, within the post-lexical phonology. Generalizations of this sort may be stated in terms of the filter given in (2) above, eliminating as ill-formed any

structure within the lexical phonology which is both [+sonorant] and [+voiced] – or [−voiced], for that matter. The rule that fills in the feature [+voiced] for the sonorants in English must be a post-lexical rule, since no structure within the lexical phonology would allow such a rule to do its job, given the filter in (2). This situation will contrast with what we find for the second kind of underspecification, to which we turn now.

The second core idea of underspecification theory goes one large step further in reducing the amount of information stated in the deepest representation. It proceeds along the following line of reasoning. Consider a feature such as voicing which has been restricted by our first principle from co-occurring with any [+sonorant] segment in a language such as English. This leaves only the [−sonorant] segments, the obstruents, for the feature voicing to co-occur with. A [-sonorant] segment can be thus be either [+voice] or [−voice]. But therein too lies a redundancy; for the obstruent in English *must* be the one or the other – it *must* be either [+voice] or [−voice]. The second principle of underspecification theory says that we must eliminate this redundancy from the underlying form as well, and one of the two feature values ([+voice] or [−voice]) must be eliminated from the underlying forms, and instead be assigned by rule. This rule will be – and must be – a lexical rule, however; we will call it a *lexical default rule*. The main reason it must be a lexical rule is that, most simply stated, it will be written just to assign the default value of the feature (say, [−voice], in this case). This assignment must not apply, of course, to segments that are already (e.g. underlyingly) marked as [+voiced] – that was the whole point of this default rule, which allowed us to conceive of the +/−voice contrast as being underlyingly privative. As we have seen, the Elsewhere Condition will give us this result, if we take this rule to be a lexical rule, for the Strict Cycle Condition says that a lexical rule will not *change* a feature value in a non-derived context. Thus we see, even within the theory itself, there is a deep and fundamental distinction between the two kinds of underspecification.

There is another, more traditional, vocabulary that can be conveniently used to express this second principle. It can be viewed as an attempt to view all feature oppositions as being underlyingly *privative*, in Trubetzkoy's (1967) terms, which is to say, involving an opposition between a positive something, on the one hand, and nothing, on the other. Trubetzkoy distinguished between such an opposition and what he called an *equipollent* opposition, in which both feature values play an equal role in the phonology. The second principle of underspecification theory, the 'eliminate unmarked feature specifications' principle, can be restated thusly: all features are underlyingly privative.

It remains an open question precisely to what extent this second

position is correct, whether in its strongest version or in some weaker version, such as that there is a strong tendency to use features privatively in the lexical phonology. The issue is a complex one, interacting with virtually every other theoretical issue current today, and, as I have indicated, the matter remains unsettled. The position adopted in this book is that, except for major class features, there is a strong tendency for features to be privative underlyingly; but under certain circumstances, when a language can reduce two privative features to one equipollent feature, the language will opt for the use of a single equipollent feature. On this point, see further discussion in section 6.3.1.

There is a second way in which the lexical default rules may be evaluated from the perspective of lexical phonology. We mentioned earlier that within lexical phonology, there is no special set of phonological redundancy rules apart from other rules; rather, all lexical phonological rules will function sometimes as lexical default rules, spelling out explicitly what is an unmarked 'expected' phonological sequence or structure in a language, and at other times as 'true' phonological rules, actually *changing* phonological features or structure.

A lexical default rule such as (32), which says that obstruents will be voiceless unless otherwise specified, looks especially like a default rule and not very much like a rule that will apply across morpheme boundaries, but only because the structural description of the rule is so small. Another example, though, is considered by Kiparsky (1982a), who suggests that rule (33) of trisyllabic shortening (cf. Chomsky and Halle 1968) expresses a generalization about morphologically simple forms.

(32) $[-\text{sonorant}] \rightarrow [-\text{voice}]$

(33)   Trisyllabic shortening
        (a) $V \rightarrow [-\text{long}] / \_\_ C_o \begin{bmatrix} V \\ -\text{stress} \end{bmatrix} C_o V$

        (b)

Long vowel	Short vowel
divine	divinity
serene	serenity
opaque	opacity
omen	ominous
declare	declarative
derive	derivative

Trisyllabic shortening, understood as a generalization across non-derived lexical forms in the language, suggests that it is much more common for a vowel to be short if followed by an unstressed vowel, if

that unstressed vowel itself is not word-final. *Canada* fits this description; *stevedore* or *nightingale* does not. If we count relevant entries in a list of monomorphemic English words, we will find, Kiparsky suggests (and I believe he is correct), that the majority will fit this pattern. Thus, short vowels, like the first one in *Canada*, may underlyingly be unspecified for length; this literal 'unmarkedness' is interpreted as a statement regarding inherent complexity of any given underlying form.

Any phonological generalization, like this one concerning the distribution of vowel length, which becomes a redundancy rule of the language[31] is thus a rule of the lexical phonology, and will by the same token come to apply to morphologically derived items in a new way. Now they may not simply specify an underlyingly unspecified feature: they may change an existing feature. Thus a long vowel, marked [+long] underlyingly, will become short ([−long]) when the suffixation of a morpheme places it in the appropriate environment for trisyllabic shortening, as when *divine* is suffixed to form *divinity* – the second vowel of the word becomes short.

We may summarize the general trend in the following way. For any given feature, the language 'prefers' one feature value in a given context, and it is the function of the lexical rules to spell out which this preferred feature value is. This preference will show up both in terms of statistical predominance in monomorphemic words, and in the changes that occur when affixes are attached. The prediction is, then, that the alternations that are produced when affixes are attached always make the resultant word 'simpler' or 'less marked' from the phonological point of view of the language – they never produce a more marked segment. That certainly is an interesting claim, and it is one worth discussing further; for, while it makes correct predictions in a good number of cases, one cannot say that the analysis of English given within the framework of lexical phonology has looked at the point closely. It is certainly not obvious that the presence of /s/ in front of /i/ is more natural and less costly in the underlying representation of an English word, as rule (1), velar softening, proposes. *King* is not obviously more unwieldy than *sing*, though that is ultimately the position of lexical phonology. But it should be clear now that ultimately the claim of lexical phonology, and what puts it at odds with almost every traditional view of morphophonemics in the broadest sense is this: lexical phonology takes entirely seriously the idea that the principles that create systematic phonological changes in the sound shape of a base and an affix when the two are juxtaposed are *no different in kind* from the principles that determine the markedness or complexity for monomorphemic structures.[32] In the case of velar softening, for example, this seems unlikely. This unlikeliness is increased, it seems to me, by the fact that the vowel or vowels that trigger the rule are

vowels which, though we may casually write them as *i*, mirroring the orthography, are typically schwa-alike in quality. In fact, the *locus classicus* on the matter is Harris (1951: 221), who observes that:

all morphemes which have members ending in /k/ when not before *ity*, have members ending in /s/ instead when before *ity*: *opaque/opacity*; {*ic*} in *electric-electricity*. . . . In such cases, it is possible to say that all morphemes which occur before *ity* will in that position have members differing in certain phonemes from the other members of the respective morpheme. . . . This statement has now become a statement about *-ity* rather than about *electric*, *sane*, etc., since the alternation does not occur before other morphemes which can be considered phonemically similar to *ity* (e.g. we have no alternation before *al*, *er*, which also begin with /ə/: *electric-al*, *saner*).

It is by no means obvious that the governing factor in the *k/s* alternation of 'electric'/'electricity' is the high vowel quality of the following vowel in close juncture. We will return to this question below.

Let us summarize our discussion in (34).

(34)   (a) Absolute constraints on lexical representations
        1. They hold throughout the lexical phonology, from start to finish.
        2. They eliminate the possibility of using either value of a given feature.
        3. They define what is a possible and what is an impossible phonological representation in a language.
        4. They are stated as static filters.
        5. They express the 'eliminate redundant features' claims of underspecification theory.
    (b) Markedness statements about underlying forms
        1. They gradually fill in unspecified values throughout the lexical derivation, operating as ordered rules; before their operation, the unmarked value of their feature must be unspecified; after their operation, the unmarked value is specified.
        2. That is, they turn a privative opposition into an equipollent opposition.
        3. They define what is a 'better' (less marked) representation, and what is 'worse' (more marked), though they do not define what is 'good' ('possible') or 'bad' ('impossible').
        4. They are no more and no less than the lexical rules of the phonology; hence they may also serve to change

feature specifications if they chance upon a derived representation.
    5. They express the 'eliminate unmarked feature specifications' part of underspecification theory.

5.2   CYCLIC DERIVATIONS IN ENGLISH

5.2.1 *The cycle*

The cycle is one of those notions behind which lies a very solid and reasonable conceptual core, but which has been used, often thoughtlessly, in such a fashion as to make well-minded citizens highly suspicious of its functioning in good society. We have already seen its formal side; we have seen how strata can be organized in a cyclic fashion, so that phonological rules have a chance to apply after each bout of affixation. Let us turn from that to the other, more practical, side of the notion.

Every language has processes for making new words, typically by affixing prefixes and suffixes, and if the base to which the affixal material has already undergone the phonological modifications necessary to make a well-formed word out of it, there will be a kind of *benign pattern of cyclicity*, in that the same principles that might have been necessary to make the base into a word will now have to come into play in order to make the combined structure a well-formed word; cf. (35).

(35)   Morpheme in its underlying representation
            | ←————— Phonological rules (e.g. stress)
        [word-1]
            | Affixation process
        [word-1]   affix
            | ←————— Phonological rules (same as above)
        [[word-1] affix] word-2

The cycle played an important role in the formal system of Chomsky and Halle's *Sound Pattern of English* (1968), but it was Brame (1974) who first put his finger on a crucial point: that the *word* is the unit that functions recursively, and that allows for the benign cyclicity that is illustrated in (35). This notion was taken up by Aronoff (1976), and served as the basis for what he referred to as a *word-based theory* of morphology, a point that was then crucial to the development of Strauss (1979) and Kiparsky (1982a).

Harris (1982) provides an excellent and precise illustration of the

conceptual core of the notion of a *word*-based approach. He observes that in Spanish, the root *desdeñ-* 'disdain' is not a word, and therefore is not subject to the word-level restriction that forbids word-final palatal nasals /ñ/ in Spanish. *desdeñ* is a morpheme; words, not morphemes, are the targets of such constraints as are embodied in rules such as (36).[33] The root *desdeñ* appears in the verb *desdeñar* and its related inflected forms, including the second-person singular subjunctive *desdeñes*, which derives from a stem *desdeñ+a* (including the thematic vowel *a*) and an inflectional suffix *-es*. The verbal derivation may perhaps be best analyzed in two cycles; that is, it may be that the stem to which inflectional material (as in KiRundi, above) is added is best analyzed as a word, in the relevant sense. Be that as it may, such a stem still includes the thematic vowel *a*, and therefore the palatal nasal is not in word-final position, and hence is not subject to rule (36). On the other hand, the noun *desden* 'disdain' (and the plural formed from it, *desdenes'* is derived from the *root* without the thematic vowel, and hence rule (36) applies to the underlying form /desdeñ/ which is trying to become a *word*, forming the surface form *desden*.

(36)   ñ → n / —] _word_

We have, then, the result of *benign cyclicity*: the theoretical result of claiming that (at least some) morphology is word-based, which is to say that it is constructed out of already formed words. Other notions of cyclicity would arise from a claim that rules applied cyclically to domains that were not independently motivatable as words. We shall not explore this apparently unnecessary notion here.

When correctly understood, then, the notion of cyclic derivations derives from the idea that certain morphological processes that *produce* words may themselves take as their input certain objects that have already been turned into well-formed words. Certain generalizations that involve word-level units may then hold both of the larger unit and the smaller unit, but ultimately this should be no more surprising or controversial than the proposition that main clauses may contain subordinate clauses.

As we shall see in some detail below, English contains many highly productive morphological processes that very clearly – one is tempted to say, ostentatiously – mark this cyclic character, with very distinctive indications that a word is contained within a word. Compounding is just one obvious example of this, but (as we shall discuss below) so is a certain kind of productive word formation based on the suffix *-ism*, as in *Indianaism* [ɪndiænəɪzm] 'a way of speaking peculiar to people from Indiana'. The larger word is clearly derived from the internal word by suffixation: [[Indiana]ism].

Where a good deal of uncertainty has remained concerns those aspects of word formation in English where there is good reason to believe that two words are related, but the boundary (or 'juncture') between the base and the suffix is not clearly marked phonologically. For example, it certainly seems *reasonable* to believe that the words *secrecy* and *presidency* are derived from the words *secret* and *president*, but at first blush one could hardly say that the case for such a derivation is transparent. The point is that, after all, even if *presidency* is formed from *president*, we need evidence that the base form *president* is being treated as a word, and is thus undergoing phonological rules as a word, before the affixation of the suffix *y*. Put another way, what is needed is a reason to treat *presidency* differently from *fraternal*, for though most of us would be willing to agree that *fraternal* is derived by the affixation of *-al* to *fratern-*, there certainly is no evidence that the base *fratern-* is analyzed and treated phonologically as a word or anything else *prior to* the affixation.

There are grounds for believing, however, that certain kinds of affixation in English are productive word-based processes which produce words that have no obvious word-type juncture inside them. For example, the suffix *-ic* can be productively added to many classes of words (e.g. *capitalist/capitalistic*), including even some monomorphemic proper names; I encountered *Agamemnonic* recently in a novel. Some of the word-level phonological effects that the base word would be expected to undergo have been either undone or undercut by the processes of *-ic* affixation, clearly. For example, the *m* of *orgasmic* is not syllabic, though it is in the base form *orgasm*, which is clearly the result of the syllabification process mentioned in chapter 3 in reference to words like *meter*. Perhaps similarly, the stress structure of the base words *Agamemnon* is modified in the nonce-formation *Agamemnonic*, and the stress form is modified even more in the case of other *-ic* suffixations, such as *metallic* or *atomic*, where syllables that otherwise would have no stress at all are now stressed. Similarly, segmental material can be lost, in formations like *tantra/tantric*, in order to avoid certain vowel sequences that are not found word-internally in English.

This kind of word formation (which, looking ahead, we may call *close-juncture* formation) seems to do its best to hide the independent wordhood of the base – covering up, so to speak, or hiding the right-hand boundary of the word that serves as the base of the affixation. In this kind of word-formation, the right-hand part of the word is modified along with the suffixation – modified, interestingly enough, not just in any way, but by and large in a way that makes the word look more like a non-derived word, phonologically speaking. But what would support a cyclic approach would be evidence over in the left-hand side of the word

that supports a view that the word was processed first *as a word* before the affix was attached. In addition, as a highly related question, we may consider the case of whether, when two of these close-juncture suffixes are added successively, there is evidence that each stage is treated phonologically as a word, or whether we wait until both suffixes are added before applying the word-level rules and constraints. We return to this question with regard to English below, when we have looked at a few processes in more detail. We will see then that there is positive evidence in support of the cyclic view.

5.2.2 The vowel system

In this section we will look at English stress and vowel quality. Some aspects of the stress patterns of English words are quite uncontroversial once a few grounds rules have been made explicit. Other aspects of the problem require a deeper look at the complex morphophonemic alternations that arise when derived words are formed. One area where a simplification of these alternations can be achieved involves the sub-area of vowel alternations, and to understand that we must discuss the analysis of the English vowel shift and vowel reduction as proposed in *The Sound Pattern of English* (Chomsky and Halle 1968).

We must first distinguish between *main stress* (or equivalently, *primary stress* or *1 stress*) and *non-main stress* (or *subsidiary*, or *secondary/ tertiary stress*). All lexical items of a major grammatical category (noun, adjective, non-auxiliary verb) have a syllable that bears main stress. Native judgments are sharp and reliable concerning which syllable bears the primary stress; the only shakiness concerns a handful of words where the primary stress may appear on either of two syllables. These words are typically trisyllabic, with the main stress falling either on the first or the last syllable, as in *cigarette* or *magazine*. Primary stress can be identified by the high pitch associated with it and the low pitch found immediately following it.

Most words of more than one syllable also have syllables bearing a stress that is not primary. There has raged for decades a controversy as to whether there are distinctions among the stresses of non-primary stressed syllables. Some say that all non-primary stressed syllables bear the same degree of stress ('secondary stress'); others, following the Trager and Smith (1951) tradition, assign several distinct levels of stress. Virtually all of the work in metrical phonology (as well as in classical generative phonology) has followed the second assumption, though evidence in its favor is remarkably slim. The issue necessarily involves the treatment of long words, as a moment's thought makes clear, and judgments are extremely subtle. For example, it has been observed that there are two

different possible pronunciations of the word *Ticonderoga*. In both, the main stress falls on the penultimate syllable, and in both the third and fifth syllables are unstressed. In both, the first and second syllables are stressed; but the two pronunciations differ with respect to which of the first two syllables has higher stress. In the Smith–Trager tradition, these two pronunciations are represented

$$3\,4\quad0\,1\,0\qquad4\,3\quad0\,1\,0$$
Ticonderoga   and   *Ticonderoga,*

where '1' marks primary stress '3' a non-primary stress, and '4' a non-primary stress lighter than '3'. It is not entirely clear what status should be given to this observation. The difference between these levels of 4-stress and 3-stress in such words is clearly not contrastive, in the sense of representing possible lexical or morphemic contrasts. Some speakers apparently report preferring one pronunciation; others report free variation. One analysis (Kiparksy 1979) bases an argument for a certain theoretical position on the fact that unanalyzed words, such as *Ticonderoga*, have two possible pronunciations (those just indicated), whereas superficially similar words that are derived from other, smaller words have only one possible pronunciation (for example, *sensationality*, where the first syllable may never be perceived as more prominent than the second; cf. also *categorization*, based on *category*, or *iconoclastic*, *anticipation*, or *superiority*).

Fortunately, most of the issues we will be concerned with do not involve the question as to whether there is more than one level of non-primary stress in English, since the very nature of the facts themselves is unclear. What is most important is whether a syllable is stressed or not, for once it has been determined which syllables are stressed, it is generally easy to predict mechanically which syllable will be assigned the primary stress (Schane 1979b). Holding aside some special cases,[34] the general principle is this: the rightmost stressed syllable that is not in the final syllable will bear the main stress. For example, the first and third syllables are stressed in both *telephone* and *telephonic*, but in the first case, main stress falls on the first syllable because the last syllable is not a possible candidate for bearing primary stress. Once a syllable is added, as in *telephonic*, the third syllable can bear the main stress, because it is the rightmost stress and it is not in the final syllable. This operation of main-stress assignment can hardly be a clear example of a cyclic operation, of course, since the choice of main stress must await all suffixal syllables (or, more accurately, all suffixal syllables of a certain large class). Better to apply the operation just once, after all the relevant syllables are in place.

The distinction between a stressed syllable (which means, here and

elsewhere, a syllable that has either primary stress or non-primary stress) and an unstressed syllable is, generally speaking, not difficult to make. It has generally been assumed that any syllable with a reduced vowel, or schwa, is unstressed; while this has on occasion been questioned (Schane 1979a), I shall maintain this assumption as well. Similarly, the diphthongs of English appear only in stressed syllables.

Thus, the first syllable of *atomic* is unstressed, because it is a schwa, though we can see that it is underlyingly the vowel [æ], as the form *atom*, with stress on the first syllable, brings out. The main (and only) stress in *divine* is on the second syllable, and there we find a diphthong [ay].

However, there do remain a good number of cases where there is some uncertainty as to whether or not a lax vowel is stressed, as, for example, in the second syllable of the word *indignation*.[35] In general, the literature has appealed in these difficult cases to theoretically neutral observers, usually the pronouncing dictionary of Kenyon and Knott (1944), but this is not the most satisfactory arrangement imaginable. Appealing to authority, no matter whose dictionary, is only a stopgap measure.

As we noted above, all of the work done within the framework of lexical phonology has assumed the theory of English phonology proposed in Chomsky and Halle's (1968) *Sound Pattern of English* (*SPE*) concerning the treatment of the vowel system, and the problematic long vowels (or diphthongs). At a superficial level (roughly, a systematic phonetic level), American English has a vowel system like that in (37), which is my own; inasmuch as I am from New York, the details concerning the low vowels differ from those of many other speakers of what might otherwise be considered standard American English. (Whether [a] is short remains controversial.)

(37)   (a) Short vowels

[ɪ] kit	[ʌ] putt	[ʊ] put
[ɛ] pet		
[æ] cat	[a] cot	[ɔ] caught

   (b) Long vowels

[iy] key	[uw] boot	
[ey] pay	[ow] go	[ɔy] toy
[ay] tie	[æw] cow	

Morphophonemic alternations relate pairs of vowels – generally, pairs consisting of a long vowel and a short vowel – that are phonetically quite different, for a historical reason. The long vowels of English underwent a systematic change called the Great English Vowel Shift. For example, what five hundred years ago was a long front mid vowel (e:) was raised, to become a long front high vowel. This rule was a late rule in the phonological system at the time, however, and it applied only to long

vowels, so an underlying /e:/ could escape this raising if the addition of a suffix to the morpheme containing a long vowel caused the vowel to shorten. For example, the second vowel of *serene* /sere:n/ was underlyingly long, and so the pronunciation [sere:n] shifted to the more modern [seri:n] (ignoring here the offglide). But the derivation of the word *serenity* involved affixation of the suffix *-ity*, and this suffixation in turn occasioned the shortening of the preceding vowel, producing the intermediate form /seren+ity/. Here the second vowel has shorted, and so will fail to undergo the late rule of vowel shift (e: → i:). We will thus find, in the grammar of this stage of English, alternations between short [e] and long [i:]. And this is still found in modern-day English, as illustrated by the alternations in (38), modified from a table in Halle

(38)   Alternations involving vowel shift

(a) ay ~ ɪ   divine/divinity
crucify/crucifixion
satire/satiric
Christ/Christmas/christian
five/fifth/fifteen (but cf. nine/ninth, and eight/eighth)
wide/width
Palestine/Palestinian

(b) iy ~ ɛ   serene/serenity
intervene/intervention
hygiene/hygienic
deep/depth, heal/health, steal/stealth
plead/pled, bleed/bled, flee/fled, feel/felt
mean/meant, meet/met, keep/kept

(c) ey ~ æ   sane/sanity
volcano/volcanic
marginal/marginalia

(d) æw ~ ʌ   profound/profundity
pronounce/pronunciation (and other words with *-nounce*, like *denounce*)
south/southern
abound/abundant
flower/flourish
tower/turret

(e) ow ~ a   verbose/verbosity
telephone/telephonic
cone/conic
harmonious/harmonic

(f) uw ~ ʌ   reduce/reduction

(1977). It is the position of *The Sound Pattern of English* (a position, as noted above, that is maintained by lexical phonology) that this is an accurate synchronic picture of English today as well.

The alternation among the pairs of front vowels in (37a–c) is reasonably regular and productive, though the same can hardly be said for the back vowels (see especially McCawley 1986). As I indicated above, the *SPE* analysis posits a regular phonological rule that recapitulates the historical change. Thus, the front mid vowel [e] may be long or short, in the lexical phonology. If it is short it is also lax, and surfaces as [ɛ]; if it is long, then it undergoes the vowel shift rule, and is raised to [i:]. Eventually all long front vowels get a *y*- offglide, by a late post-lexical rule. Finally, it is worth bearing in mind that, within the logic of the system, a long vowel that undergoes vowel shift may either be long underlyingly, never being shortened during the derivation, or be short underlyingly, and become lengthened during the derivation (e.g. *margin-al*, *margin-a:l-ia*).

The point of this discussion is to see how the *SPE* description encourages us to try to account for the phonological differences between many pairs of derivationally related words in a purely phonological way. The vowel quality differences noted in (38) can be abstracted away from if we can provide a reasonably small number of phonological rules that shorten or lengthen the vowels in question, leaving it for a much later rule to effect the vowel quality changes.

But derivationally related forms in English differ not just in vowel quality, but also, quite frequently, in their stress patterns (e.g. *compete/ competition*, *atom/atomic*). Lexical phonology, following a number of workers in this area,[36] has defended the position that there exist phonological rules that shorten and lengthen vowels, and assign metrical structure after such length-affecting rules; further length-affecting rules will also apply after metrical structure has been established. The stress pattern of a morphologically derived English word, then, is claimed to be predictable according to the regular patterns established for monomorphemic words, except that some shortening and lengthening rules apply before the stress rules apply, and others apply also after the stress rules. We will explore this by first presenting a brief overview of the principles of stress assignment for monomorphemic words in English.

5.2.3   English stress rules

We have looked at the basic operation of the English stress rule in our discussion of metrical phonology. Most work in lexical phonology is in agreement on the following main points of English stress.

The rule that assigns the rightmost stress to a word is the quantity-

sensitive rule that we discussed in chapter 3 which assigns a foot – i.e. a stress – to the right-hand end of the word, skipping over extrametrical material and one light syllable, if there is one. However, a final syllable which is extrametrical but contains a long vowel will nonetheless be marked as stressed, i.e. will form a foot. This is possible primarily in the case of nouns, whose final syllable is extrametrical. Nonetheless, a final syllable with a tensed vowel will find that syllable stressed. In additional, certain adjectival suffixes may have this property of attracting stress, such as *-oid* and *-ory*.[37]

Reviewing, then, from chapter 3, the stress pattern of the sorts of words in (39) will be predictable. If the final syllable contains a long vowel, as in (39d), that final syllable is stressed. Otherwise, stress falls on either the penultimate or the antepenultimate, depending on the weight of the penultimate. If the penultimate is heavy, as in (39b) or (39c), it is stressed; otherwise, as in (39a), the antepenult is stressed.[38] A syllable, of course, can be heavy either by virtue of having a diphthong, or by virtue of being closed, as in the penultimate syllables of (39b) and (39c), respectively.

(39)	(a) America	(b) aroma	(c) veranda	(d) Naverone
	cinema	balalaika	agenda	magazine
	asparagus	hiatus	consensus	antidote
	metropolis	horizon	synopsis	
	javelin	corona	utensil	

These principles have been called the Main English Stress Rule; they are followed by a related rule or set of rules that gives rise to something close to a pattern of alternating stress earlier in the word. The precise nature of these rules is crucial for the formulation of the arguments concerning the cyclic nature of English stress assignment, as is the precise formulation of the several rules of destressing that are necessary, as we shall see.

Hayes (1982) argues that the principles of stress assignment involved here are as might be sketched in (40). He argues that the rule of stress retraction, which he calls Strong Retraction, is quantity-insensitive,

(40)   Stress retraction

ớ σ ớ σ ớ σ σ]

                    Main English
                    Stress Rule
                    (quantity-sensitive)
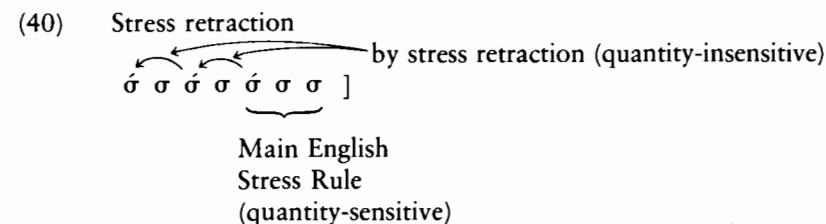
unlike the Main English Stress Rule. The arguments for this position are moderately complex, and reviewing all the arguments would require a book in itself; nonetheless, this assumption is crucial to one of the most often-cited arguments for cyclicity, which is essentially as follows. The words *compensation*, *condensation*, and *compurgation*[39] have essentially identical surface syllable structures. However, only the third of these is stressed in a non-cyclic fashion, and the lack of stress on the second syllable, despite its status as a closed (and thus heavy) syllable, shows that the pattern of stress retraction leftward across the word will skip over a heavy syllable as it constructs metrical feet, as in (40); this is illustrated in (41). Thus retraction is not quantity-sensitive.

(41)   x   x
          xx   xx   xx   x
          com   pur   ga   tion

But the stress pattern in *condensation*, it is suggested, has distinct stress on the second syllable, unlike both *compurgation* and *compensation*. This is the result of the stress (Row 1 grid mark) that appears on the second syllable of the base form, *condense*, arguably; no other factor could account for this distinct stress pattern.

There are two important destressing rules to bear in mind. Both may be viewed profitably as stress-clash-reducing principles. The first Hayes calls 'prestress destressing'; mnemonically, we might refer to it as the 'banana/bandanna' rule, because it is the rule responsible for the deletion of stress on the initial syllable of *banana*, which is originally assigned a stress pattern just like that of *bandanna*, as we see in (42). Prestress destressing (44), however, removes the stress from a (degenerate) monosyllabic foot which immediately precedes another foot, and which consists of a single, light syllable, thus removing the stress from the first syllable of *banana*, as in (43).[40] This rule applies only when the syllable that immediately follows is the main stress of the word as a whole, giving rise to differences in the stress on the first syllable in pairs such as *department/departmental*, or *relax/relaxation*. In both cases the second syllable is stressed, but only in the first of each does it bear the main stress of the word, and only there is the first syllable unstressed.

Following the *SPE* analysis, this rule is often conflated with the process that shortens and destresses long vowels in such forms as *explanation*

(42)   x   x
              x   x   x   x
            x   x   x   xx   x   x
          (a) banana   (b) bandanna

(43)   x
                x
              x x x
              banana

(44)   Prestress destressing
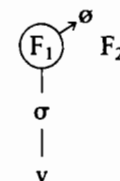
Condition: F_2 is main stress of word

(from *explain*) and *invitation* (from *invite*). This process of shortening does take place even when the stressed vowel does not bear the main word stress; for example, any verb ending in *-ate* will find that *-ate* shortened, and destressed, before the suffix *-ory*, as in *deprecate/ deprecatory*, *congratulate/congratulatory*.

The second rule of destressing applies rightward, deleting a stress foot on an open, light syllable immediately following a stressed syllable. The effect of this rule is particularly noticeable in the apparent alternations it gives rise to in a suffix like *-ory*, which has a full stressed vowel in words like *deprecatory*, *promissory* or *equivocatory*, where the preceding vowel is unstressed, but loses both length and stress in words like *advisory* and *illusory*, where the immediately preceding syllable is stressed.[41]

5.2.4 Strata in English

We have now completed a quick survey of the basic principles of English stress assignment, and we can begin to approach some of the thornier problems that arise in dealing with the question of the extent to which English derivation can be treated as cyclic phonology.

(1) A great deal is known about English morphology, and yet at the same time there is considerable controversy about how to interpret what is known. In this section I shall explain what the main dimensions are along which investigators have found differences in the affixation processes in English, and shall indicate several of the ways that have been suggested for integrating these observed differences into theoretically coherent programs, of which lexical phonology is one.

Let us start with a simple example or two. There are, it appears, two kinds of ways that suffixes in English can be found attached to their *base*. In the first, the base is an existing word, with a meaning, a stress pattern,

and a syllabification that it would have standing alone; and when the suffix is attached, it leaves that base alone, phonologically. This is most clearly seen in cases where the juncture between the base and the suffix forms a sequence that cannot be found within monomorphemic words. I have already mentioned a few examples of this sort. While we cannot have geminates word-internally, we find these in words like *coolly*, *cleanness, tailless*. We cannot have non-high vowels followed by another vowel word-internally, but we find this in words like *Indianaism* or *Bermudaize*. The same point holds for prefixes as well, yet we find geminates with the prefix *un-*: *unnatural*, for example. Following traditional terminology, we may refer to this as *affixation with open juncture*. Not all affixes ever participate in this kind of juncture, as we shall see. It is a fact – and not a logical necessity – that, in all cases of affixation with open juncture, the base is an independently existing word, with a stress pattern that remains unchanged when affixed, and a meaning that is composed to form the meaning of the whole word.

Many cases of affixation exist, of course, where the base to which the affix is attached either does not exist as a free standing word (e.g. *patern-al, plagiar-ize, inocul-ate*), or has undergone some segmental modification (e.g. *buddh[]-ism*, where a schwa has vanished, or *jeopard[]-ize*, where a *y* of some sort has vanished, or *syllab-ic*, where an *l* has vanished, or *libid[in]-ous*, where an *o* has vanished and an *in* has taken its place). Despite certain complications which we shall discuss below, it can be said that, in all of this second kind of affixation, the base plus the suffix together satisfy the phonotactics of monomorphemic English words, and that the global stress pattern of such words also satisfies the rules of English monomorphemic words. For want of a better descriptive term, we will call this *close juncture affixation*.

This distinction between types of affixation has been described and categorized in different ways. The most familiar view attributes the difference to a difference of the *suffix*, and labels the two types of affixation we have described as affixations of, respectively, *stress-neutral* and *stress-affecting* suffixes; or *stratum 2* and *stratum 1* suffixes; or *#* suffixes and *+ suffixes*. From the examples we have looked at so far, it is not obvious, of course, that the difference in the kind of juncture is the responsibility of the particular suffix that was chosen – it may, after all, have been the base that determined the kind of juncture that occurred. Since it has been taken for so long to be the suffix that is responsible for the type of junction, however, we should consider more carefully the pros and cons of this position. One argument that has been given for ascribing the junction type to the suffix is that it has been observed that the same base can appear sometimes in close juncture with one suffix, and in open juncture with another:

Close juncture	Open juncture
curiosity	curiously
deceptive	deceiving
obligatory	obliging

Thus, goes the first argument, if the base were the object responsible for the kind of juncture we find in a suffixed form, no base would allow itself to be found in suffixed forms with different kinds of juncture. Therefore, it must be the suffix that determines the junction type.

The second argument for handing the junctural responsibility over to the suffixes derives from trying to identify minimal pairs among morphemes, where one always comes with close juncture, the other always with open juncture. If juncture type goes hand in hand with the particular meaning of the suffix, then surely it would follow that the suffix is formally responsible for the junction type. For example, there is a suffix *-ism* that marks characteristic speech forms, and which comes with open juncture, and a suffix *-ism* that indicates a philosophy, in a broad sense, with close juncture. This gives us a number of close and actual minimal pairs; cf. (45).

(45)	Close juncture	Open juncture
	cathóli[s]ism	cátholi[k]ism
	Buddhism	?Buddha-ism
	communism	??commune-ism
		Indiana-ism
		Indian-ism (from *Indian*, not from *Indiana*)

There is certainly something to this generalization, but most of what is right about it has to do with new formations that we might make up, not with existing words. If we make up a word with the speech-mannerism sense, we will, in all likelihood, employ a suffixation process with open juncture (*Turkism, Yankeeism*); but the other thing we might well do is attach *-ism* to a base ending in *-ic*, whereupon the *ic* will become *i[s]*, phonologically: *gallicism, italicism, slavicism, anglicism, sinicism* (or *sinocism?*), and so forth. Which of these already exist and which have just been invented I am not quite sure, and that is the point. It is certainly not clear, to say the least, that this second group is an instance of suffixation with open (#) juncture. And there do exist a good number of suffixed *-ism* words that have close juncture but the mannerisms meaning, rather than the philosophy meaning, of which one is perhaps

mannerism, depending on whether the base, *manner*, has the appropriate meaning (e.g. *archaism*, *hebraism*, perhaps *rhotacism*).

Thus, our generalization about the speech-forms sense of *-ism* always taking open juncture is not terribly strong; and, going the other way, *cannibalism* (although it may, irrelevantly, have the speech-form sense!) generally has the philosophy sense, but it has the stress pattern of the word *cannibal* with *-ism* attached as an open juncture – otherwise we would expect to get stress on the second syllable, as in *calamity*.

The point of this extended discussion is to indicate that the *prima facie* evidence in favor of ascribing *to the suffix* the determining role in deciding whether we get open or close juncture in a particular base–suffix combination may be a bit hasty. After all, the first part of the argument as sketched so far is that there are bases that occur with both open and close juncture in different words. But it is not hard to find – and we have already seen – cases where what looks for all the world like the same suffix is sometimes attaching in close juncture, and sometimes in open juncture. We have seen such cases with the speech-manner *-ism*; the same point can be seen with the open juncture *-ize* in *Bermuda-ize* and the close juncture in *notar[]ize*, *synchron[]ize*, or *Catholicize*. In the first two forms, a word-final /y/ is elided, a sign of close juncture, and in the third, the stress pattern has been modified drastically in the base. And perhaps it need not be said, but we should also be aware that the mere fact that certain suffixes (*-ity*, for example, as in *sanity*, *divinity*, or *electricity*) will not attach to certain bases (**tallity*, *?fabulosity*) is no more a fact about the *suffix* than it is about the *base*; that is, just as one can generalize across the various stems that *-ity* will not attach to, one can generalize over the various suffixes that *tall* will not attach to (**tallity*, **tallous*, **tallic*, **tallible*, and so forth).

(2) Historical interlude In order to discuss current approaches, it is necessary to sketch a certain number of generalizations that have been noted in the literature, and which serve as the basis for most of the models currently under discussion.

Siegel (1974) followed Chomsky and Halle (1968) in assigning to each suffix a boundary, thereby claiming that the juncture that was found was grammatically the result of the choice of suffix. Some suffixes came with a + boundary, leading to close juncture, in general (though she did not use that term), and others came with a # boundary; she referred to these as class I and class II suffixes, respectively, and gave the examples shown in (46).

Siegel made the following observations. (i) Class II suffixes attach, with only a handful of exceptions,[42] to independently existing words. Class I suffixes attach either to words or to stems, i.e. bound morphemes. (ii)

(46)

Class I	Class II
-ation	-ness
-able	-less
-y (nominalizing)	-ly
-al	-al (nominalizing)
-ity	-y (adjective forming)
-ate	-ful
-ify	-ment
-ic	

Class I suffixes will occasion a shift in stress from the pattern found on the independently existing base word, whereas class II suffixes do not. (iii) Class I suffixes can follow each other (e.g. *histor-ic-al*, *illumin-at-ion*, *in-determin-ate/in-determin-ac-y*), and class II suffixes can follow both class I suffixes (*fratern-al-ly*, *transform-ate-ion-less*) and class II suffixes (*weight-less-ness*). Class I suffixes can never, however, follow a class II suffix (**weight-less-ity*, **fatal-ism-al*) (though I will disagree with this point below). (iv) Certain class II suffixes place conditions on the stress pattern that independently exists on their base word; class I suffixes never do. (Siegel discusses the *-al* nominalizing suffix,[43] the *-(e teria* suffix, and the *-ful* suffix).

Siegel united these observations in the hypothesis that the morphological process of suffixation should be divided into two layers, with the stress rules applying after all class I suffixation but before all class II suffixation, as indicated schematically in (47). This proposal was quite influential, and led in an almost direct path to the construction of the model of lexical phonology. To arrive at the conception of Kiparsky (1982a,b), the notion was needed in addition that the stress rules operate cyclically, after each of the class I suffixes is added,[44] and that class I suffixes are simply the first-stratum suffixes. The class II suffixes could

(47)

```
                        Roots
                          ↓
        ┌─────────────────────────────────────────┐
        │  first set of affixation processes – class I  │
        └─────────────────────────────────────────┘
                          ↓
               ┌──────────────────────┐
               │   cyclic stress rules   │
               └──────────────────────┘
                          ↓
        ┌─────────────────────────────────────────┐
        │ second set of affixation processes – class II │
        └─────────────────────────────────────────┘
                          ↓
               ┌──────────────────────┐
               │    word-level rules    │
               └──────────────────────┘
```
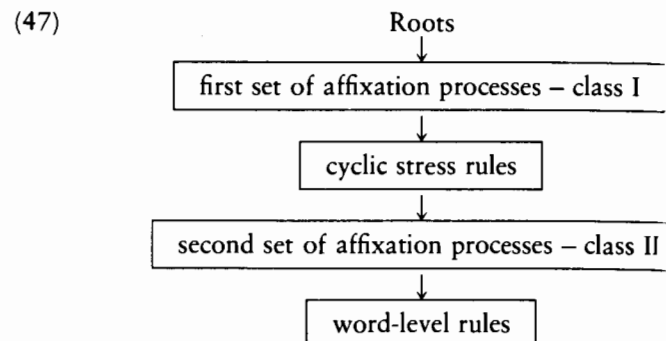
now be added on a later – perhaps the second – stratum. Any stratum 2 process would leave unaffected the stress pattern of the derived word.[45]

Inherent in the model proposed by Kiparksy (1982a,b) is the idea that close juncture is what is produced by forms going through the stratum 1 rules; a form that has managed to get through the roster of stratum 1 rules will satisfy the word-internal phonotactics of English. Suffixes affixed in the second stratum are the sources of open juncture, for they are not subject to the stratum 1 rules that would be responsible for such affairs as the deletion of the schwa in *Buddhism*. Thus juncture is translated into strata, so to speak; word-level phonotactics (which is to say, close juncture) becomes a matter of stratum 1 rule effects.

Determining which stratum a given suffix is in, though, can be harder than it may seem at first. This is apparent just from comparing the several lengthy studies that have looked into this hypothesis (Siegel 1974; Allen 1978; Aronoff 1976; Selkirk 1982b); their classification of this basic point for various suffixes is by no means uniform. Selkirk is, for her part, explicit about the point that some suffixes are in both classes, but she suggests that 'the vast majority' of English affixes are clearly in one class or the other.

We should be careful in evaluating the result of being faced with suffixes that individually belong to both classes. On the one hand, if it is true that individual suffixes do, in fact, come with different kinds of juncture, then we need a model that can recognize that, and treat it adequately, and not require that each suffix be assigned to one kind of juncture or affixation. On the other hand, assigning a suffix to both classes without independent justification can, under certain circumstances, be just a sign that the model is in trouble, and is making wrong predictions. Consider an example of the second sort.

The class I/class II model in (47), and its descendants, make the strict prediction that class I suffixes never appear to the right of class II suffixes. Now, the suffix *-al* is a clear case of a class I suffix, since it easily attaches to non-words (*patern-al*), and changes the stress pattern of its base (*parent-al*). The suffix *-ment*, however, has to be a class II suffix; as Siegel (1977) points out, only that assumption would give the right stress pattern for *abolishment*; for even if (in more recent terminology) the final syllable were extrametrical, we would expect the stress to fall on the penultimate. Only if we let *abolish* derive its stress on the first stratum can the form be derived correctly. The problem is, though, that the sequence *ment-al* is quite common, as we see in words like *governmental, ornamental, elemental, supplemental,* and so forth. Selkirk suggests that this shows that there are two suffixes *-ment*, one from class I, which appears in *govern+ment* and *orna+ment*, and another from class II which appears in *employ#ment*. Only the former can appear with the

class I suffix *-al* following it. This, of course, is merely recognizing the problem, not solving it.[46] We shall offer a solution to it below.

(3)   In section (1) we reviewed the various characteristics of English word formation that motivate distinguishing between word formation with open juncture and that with close juncture. In section (2) we considered the various approaches that have been proposed to incorporate certain further hypotheses regarding these kinds of junctures, all of them following upon Siegel's perspective (what Selkirk calls the *Affix Ordering Generalization*: close juncture is always closer to the root than any open juncture), including most importantly the 'stratification' of morphology in lexical phonology. In this section, we will review the various arguments that have been given subsequently which cast serious doubt on the Affix Ordering Generalization and its descendants.[47] It is important to bear in mind, however, that rejection of this hypothesis is independent of the question of cyclicity, which we return to in the next section, where we will provide an analysis of a class of data in English that is simultaneously an argument *for* cyclicity and *against* the Affix Ordering Generalization, in that an open juncture suffix precedes a close juncture suffix, both of which apply cyclically.

The arguments against the Affix Ordering Generalization, and thus the division of affixes into strata, include (i) those based on prefix/suffix incompatibilities; (ii) those based on inverted orders among the suffixes, with stratum 2 suffixes preceding stratum 1 suffixes; and (iii) problems regarding the placement of compounds within a stratal account. I shall present an example of the second kind in detail in the next section, and leave the third kind to interested readers.[48] The arguments based on prefix/suffix incompatibilities are of considerable interest, and I will indicate their general form here.

Referring to a number of striking observations in Williams (1981), Strauss (1982b) observes that the Affix Ordering Generalization has problems when it encounters words like *ungrammaticality, reorganization,* and even forms like *South American,* under certain assumptions about compounding. In each case, the suffix is a stratum 1 suffix, but there is good reason to believe that the base to which it is attached has undergone stratum 2 prefixation (*un-grammatical, re-organize*), or even compounding, (*South America*) *before* the affix is attached. This is sketched graphically in (48).

In cases (48a,b), the prefixes have the familiar properties of layer 2 affixes (they attach to existing words, they are thoroughly productive, they allow segment sequences at their juncture that are not found morpheme internally), but the argument that the correct hierarchical structure is as given in (48) is quite strong. In case (48a), *un-* attaches

(48)   (a)

$$N_{word}$$

$$A_{word} \qquad N_{affix}$$

$$A_{affix} \qquad A_{word}$$

un   grammatical   ity

(b)

$$N_{word}$$

$$V_{word} \qquad N_{affix}$$

$$V_{affix} \qquad V_{word}$$

re   organize   ation

(c)

$$N_{word}$$

$$N_{word} \qquad N_{affix}$$

$$N_{word} \qquad N_{word}$$
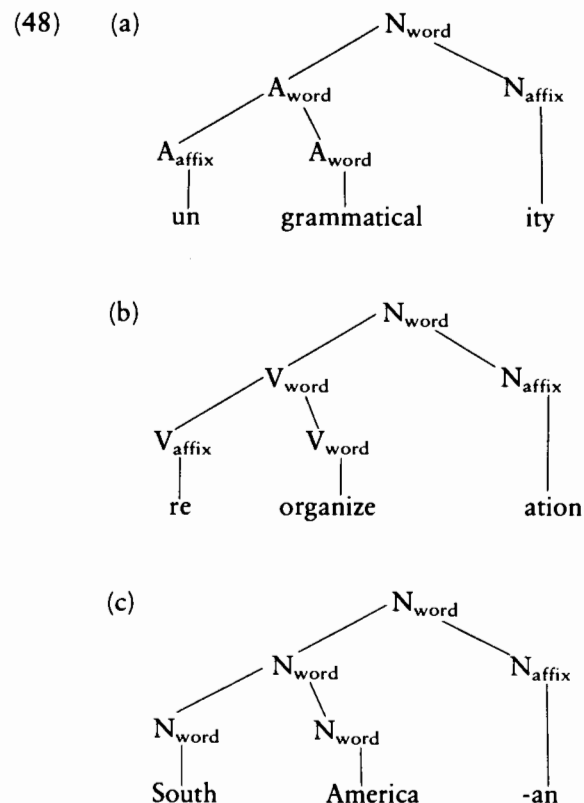
South   America   -an

productively to adjectives, not to nouns; *ungrammatical* exists, but *un-* does not productively attach to nouns. *Re-* similarly attaches productively to verbs, not to nouns, and of course in case (48c), the adjectival form *South American* is based on the compound *South America*.[49]

Strauss characterizes these 'paradoxes' as involving prefixes and suffixes, and it certainly does seem to be correct that it is difficult or impossible to assign to morphemes that flank a stem a relative *order* that is consistent with what we know about the rest of the language.[50] In the next section I will suggest further reasons for not incorporating level ordering into our account of morphology.

5.2.5   The cycle in English

A number of arguments have been given for a cyclic account of English stress placement. As I indicated above, many of these rest on elusive data, but I will offer the following argument in favor of a word-based cyclic account of English stress. As noted above, the notion of benign cyclicity

can be reduced to the less striking statement that words are formed from other words, and that word-level rules can therefore have applied to subparts of words as well as to the whole word itself. But a cyclic account leads to some rather surprising results, as well.

We have seen that the metrical grid assigned to a given word forms a chart, the bottom row of which associates to the skeletal tier in ways that we have explored in preceding chapters. Halle and Vergnaud (1987) have proposed that the notion of cyclicity in stress assignment be understood in terms of autosegmental charts as well. They suggest that each cycle produces a distinct metrical grid, associated to the skeleton, and that there are limited ways in which information can pass from the grid corresponding to an earlier cycle to the grid of the later cycle.[51] They suggest that in English, a word-level stress (that is, the main stress of a word, a grid mark on the third row) is copied over to the next cycle, though other (secondary) stresses are not copied.[52] This predicts a derivation as in (49) for the pairs of words *condense, condensation*, in which the second syllable should be stressed in both forms.

(49)   (a)

```
        x
    x   x
   xx  xx
  condense
```

(b)
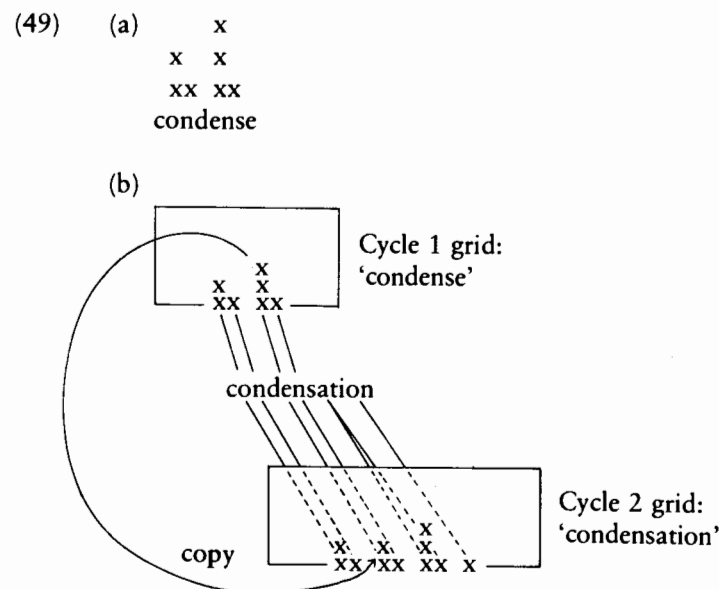
The strongest argument for a word-based cyclic approach to English stress would be one that (i) involves a restriction on morphology that accounts for combinations of morphemes of a sort that have never been heard, but which sound totally unacceptable, and (ii) follows naturally from, and only from, an account with more than one grid chart simultaneously associated with a word. The argument I will provide, in

fact, is based on the behavior of words with *three* simultaneous grid charts, making the case even stronger.

Consider first the set of data given in (50), first discussed in Strauss (1983), who points out that in (50a), there are a number of words formed with the suffix -*ist* that do not allow for acceptable suffixed forms with -*ic*, while in (50b), the forms are fine. In parts (c) and (d), we find examples of trisyllabic -*ist* words which allow -*ic* suffixation, and in (e), trisyllabic forms that do not allow -*ic* suffixation.

(50) (a) sexist   *sexistic   (b) sadist   sadistic
         faddist   *faddistic   fascist   fascistic
         rightist *rightistic   linguist linguistic
         leftist   *leftistic   deist   deistic
         racist   *racistic   theist   theistic
         rapist   *rapistic   jurist   juristic
         putschist *putschistic   sophist   sophistic

     (c) fatalist   fatalistic   (d) masochist masochistic
         regalist   regalistic   plagiarist plagiaristic
         humorist   humoristic   populist   populistic
         socialist socialistic   atheist   atheistic
         humanist   humanistic   aphorist   aphoristic
         realist   realistic   hedonist   hedonistic
         communist communistic   anarchist anarchistic

         (e) cartoonist *cartoonistic
             escapist   *escapistic
             falangist *falangistic
             alarmist   *alarmistic
             defeatist *defeatistic
             adventist *adventistic
             conformist *conformistic
             extremist *extremistic
             reservist *reservistic
             careerist *careeristic

Strauss proposes that the generalization is that '-*ic* may attach to an X+*ist* base provided the final syllable of X is not primary stressed if X is a lexical item.' Thus an example in (50a) (e.g. *rightistic) is bad because *right* is a lexical item with final stress; *rightist* is formed from it; and therefore we cannot add on -*ic*. Likewise, we cannot form, *cartoonistic from *cartoonist*, because it has been built from an existing word, *cartoon*, with final stress. *Sadist*, however, is not synchronically built from an existing word (surely not from *de Sade*, synchronically) and thus

Strauss's restriction does not apply, allowing *sadistic*, just as *fatalistic* is allowed, because although *fatalist* is built from an existing word, *fatal*, that word does not have final stress.

This does seem, admittedly, like magic. The generalization is too good to be true – almost literally, for, although the correct predictions are being made, it seems impossible that one's own judgments could be driven by Strauss's generalization. How, after all, could we have learned to put such a affix- and language-particular restriction on the -*ic* suffixation rule? Surely we must look deeper.

Let us step back a moment and listen to the impossible word *alarmistic. What makes the word impossible is not too hard to determine: it is that we cannot figure out how to stress it, because it feels like we want to give the second syllable a word stress, and also the third syllable, and we simply cannot figure how to do both at the same time, that is, with a single articulation – so we give up, and say it is unacceptable.

The problem, then, is clearly one of stress-clash, but how, and why? Let us review a few things about stress-clash in English. First of all, some stress-clashes are fine. There are many bisyllabic words where both are stressed (e.g., nýlòn, Révlòn, Ràngóon) and polysyllabic words with adjacent stresses (Ìllíni, Tìcònderóga, etc.). Monomorphemic words put no major constraints on the pattern of adjacent stresses.

Furthermore, there are some kinds of suffixation that look like they will lead to a stress-clash which is then resolved one way or the other with no great difficulty – not every potential clash leads to the phonological breakdown felt in *alarmistic*. A word like *admonish*, for example, with penultimate stress, has a stress pattern just like that of *alarmist*; but because its stress is in an open syllable, it shortens and destresses when the -*ion* suffix is added, as in *admonition*, which has (or can have) an unstressed second syllable. A word like *abnormal*, again with stress on the second syllable, does not have to lose that stress when the suffix -*ity* is added, since it is in a closed syllable, but the stress on the third syllable is clearly stronger (as can be the stress on the first syllable) than that of the second syllable in *abnormality*. Both of these cases involve close juncture, and neither seems to be troubled by stress-clash – in one case because the stress is deleted, in the other because it is somehow sufficiently subordinated.

There *are* productive morphological processes that clearly are directly more sensitive to stress-clash, however, such as -*ize* suffixation. We have already observed that the suffix -*ize* can be found with either close juncture or open juncture: the cases included *catholicize/notarize* and *Bermudaize*, respectively. The task of determining which kind of juncture we have here is not that easy, however. On any of the accounts of verbal

stress assignment discussed in this chapter and chapter 4, the correct stress pattern can be obtained whether the suffix -ize is added on stratum 1, with cyclic stress assignment, or on stratum 2, even if we assume no application of the Main English Stress Rule (if we allow a rule to stress word-final long vowels in the second-stratum phonology).[53] To see why, consider the case of a word like standardize or cannibalize. If stress is assigned to standard and cannibal on the first cycle, and then if -ize is attached on a second stratum 1 cycle, stress will be assigned to the final syllable (-ize) because it is heavy, and nothing further will happen. Stress retraction, discussed above (see (40)), does not apply when there are no 'unfooted' syllables, i.e. no syllables that have not yet been gathered into feet.[54]

In short, a monosyllabic verbal suffix (like -ize) is going to turn out not to affect the stress pattern of the base to which it is attached on stratum 1, just because of the way the English rules are set up, and because stress retraction obeys the Strict Cycle Condition, and thus does not *change* the metrical structure that it finds, but instead only applies to fill structure in. On the other hand, if -ize is a stratum 2 suffix, then all the more so it will not affect the stress pattern of the base to which it is attached, just by the way we have set things up: the Main English Stress Rule does not apply on stratum 2.

So if we want to determine in a particular case whether we have open or close juncture with a particular -ize verb, and if we recognize that the decision may have to be made on a case-by-case basis, we can only conclude that stress will not help us in this task.

I suggest that, in cases of indeterminacy, the default assumption to make is that we have open juncture — that is, from the point of view of stratal ordering, a stratum 2 juncture. The judgment is a delicate one, but it is based in part on the following peculiarity.

The suffix -ize is highly productive, but it has some phonological limitations. Consider the forms and judgments in (51). The fact that we can *autumnize* an heating system, for example, but not *fallize* it, suggests two things. First, it affirms our decision to treat this as open juncture, since the /n/ of autumnal does not appear in this form; that is, we have a true word inside [[autumn]ize]. Second, it suggests that there is a phonological, not a semantic or pragmatic, reason why we do not have *fallize*. The generalization is a simple one, to be sure; -ize does not attach to a word with final stress, which includes all monosyllables. But should this be stated as a fact about -ize, as we have put it, or is it part of a larger generalization? In fact, it seems that this is part of a larger generalization that prohibits stress clash over an open suffix juncture.

We suggest that there is a prohibition in English against adjacent stresses across open juncture when the material on the right is suffixal (52). We label

(51)   (a) winterize   (b) alphabetize   (c) *Montrealize
            summerize   radicalize   *New Yorkize
            autumnize   departmentalize   Bostonize
            *fallize   *?cartoonize   ?Chicagoize
            *springize   journalize
                             ?*magazineize
                             ?*reviewize
                             *bookize
                             publicize
                             legalize

the *morphological* bracketing with its corresponding *phonological* juncture type, open (#) or close (+). (Note that the prohibited clash involves successive syllables, not moras.)

(52)   Prohibited: *   *   foot
                       *   *   syllable
                     σ]# σ]

Such clashes are not prohibited, however, as long as they do not occur across an open juncture (which is written with a word-labelled bracket in filter (52)). As we have observed, stress-clashes word-internally do not offend the phonological sensibilities of the English language, nor are such clashes across a close juncture impossible, as *abnormality* illustrates.

With filter (52) in mind, we can return to the account of Strauss's generalization involving *alarmistic*. The productive, word-based formation with -ist is again one with open juncture, just as is the case with -ize. That is the crucial point, for when we add the next suffix, -ic, stress will fall on -ist, yielding a stress-clash across open juncture. This is illustrated in (53).

As we observed, this account violates the Affix Ordering Generalization; class I suffixes follow class 2 suffixes, and in general I believe that is correct, and not theoretically ruled out. As indicated above, when the literature has faced this observation, the most frequent response to the perceived problem has been to say that a given morpheme can play the role of either a class I or a class II suffix. But as noted above, in reference to Selkirk's treatment of the problem of why ment+al will not attach to free-standing verbs, that is a recognition of the problem, not a solution.

We can now return to the problem now of why we find the pattern in (54), and can immediately see the reason why the forms in the third column of (54b) are bad. They are bad for the same reason that *alarmistic* is bad: they require a stress-clash across an open juncture.

In summary, then, we have here an argument against the Affix Ordering Generalization, and thus against a stratal approach to English

(53)

by *banana*-rule
copy

ø← x
x x
a l a r m + i s t # i c

by *banana*-rule
ø← x
copy

copy
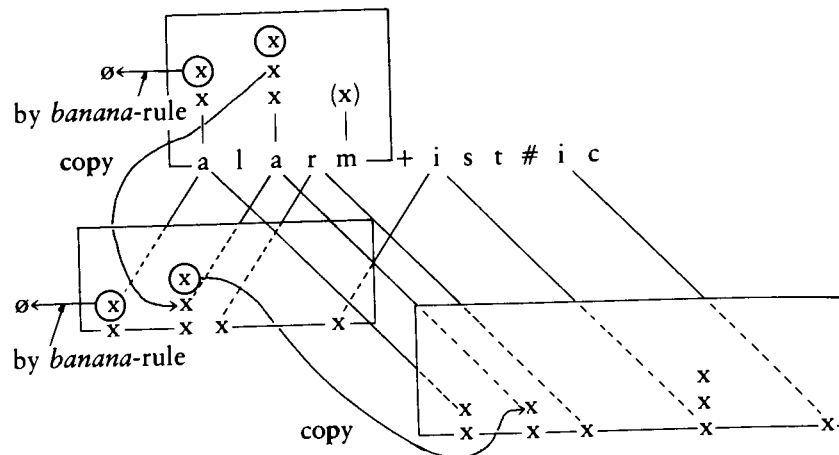
(54)

(a)	develop	development	developmental
(b)	employ	employment	*employmental
	recruit	recruitment	*recruitmental
	govern†	government	*governmental
(c)		ornament	ornamental
		government	governmental
		supplement	supplemental

x	Row 2
x x	Row 1
x x x	Row 0

(d) [[employ] ment]
         #   Second cycle

x	Row 2
x x x	Row 1
x x x (x)	Row 0

[[employ] ment] al]
   #   +   Third cycle

†In the linguistic sense; e.g., a verb 'governs' an object.

morphology and to lexical phonology.[55] On the other hand, it is a positive argument for a cyclic and word-based approach to English stress assignment. What is crucial is the kind of juncture that is found between a base and a suffix; the default case is open juncture, since English adheres to a word-based morphology.[56] Such a juncture does not cause restructuring of the metrical grid. However, open juncture does just that, while preserving earlier word main stress, as Halle and Vergnaud have elucidated.

In this short study of a few morphological processes in English, we have focused on open juncture operations. In the case of words built with close juncture, corresponding to stratum 1, or class I, formations, the case must be built with considerable care, and we will not look much further into the details in this book. A certain amount of the discussion (e.g. Kiparsky 1982a: 35–45) might lead one to believe that accepting a cyclic account of these words necessarily meant accepting accounts of words like *hypocrisy*, in which the first syllable is underlying long, and later shortened by rule, or *presidency*, which is formed by the suffixation of a *y* on stratum 1.

The fact is that the phonological effects of close juncture suffixes are quite varied, and are difficult to phonologize. Nonetheless, lexical phonology's core idea does make a contribution to our understanding here, for there is a strong generalization lurking behind these analyses: the vowel length and stress modifications that are occasioned on the final syllable of a word by the suffixation of class I suffixes are changes that simplify the word, in the sense of making it look more like a monomorphemic word. Lexical phonology's premise – that the lexical phonological rules are one and the same with the set of rules defining natural underlying forms – states this direction of change correctly. The changes induced with suffixation are strong tendencies; a vowel in English tends to be short when followed by an unstressed vowel in non-final position, and we call the tendency in this direction 'trisyllabic shortening' (rule (33) above). When stems shift in this direction under affixation, we see this rule applying in a structure-changing fashion, though it may have exceptions, both scattered (such as *obesity*) and systematic (such as *exploitative*, since the diphthong *oy* will never shorten: it has no shortened form).[57] This is an important result which any version of a lexical phonology must maintain in its theoretical core; and, along with lexical phonology's contribution to our understanding of the importance of underspecification in the deeper part of the phonology, it forms the basis of a major contribution to phonological theory.

16 Sources: (i) Wright (1971), McCarthy (1979) and references there; (ii) McCarthy (1980); (iii) Kenstowicz (1980). See also Prince (1980), Welden (1980).

17 See McCarthy (1979: 451–2), on Classical and Cairene, and McCarthy (1980: 96, fn 5), on Damascene. The former source suggests that final CVVs are stressed in Cairene and unstressed in Classical Arabic. Kenstowicz (1980) is clear on the point that final CVV syllables are stressed in Cairene.

18 Examples from Hayes (1982).

19 It is important here to focus here on the presence of stress, not the presence of main (word) stress. We return to the treatment of word stress in English in section 5.2.3, where we will see that the location of main stress in English is predictable once the (secondary) stresses have been placed, roughly speaking – the guiding principle being that the rightmost stressed syllable that is *not* word-final receives the word stress.

20 I modify the form of his suggestions to suit the version of metrical theory we are employing here.

21 This accounts leaves open why long vowels in final syllables are unambiguously treated as heavy, surprisingly. However, as we noted in Chapter 3, the second half of the long vowel (or diphthong) in English is normally in complementary distribution with the non-coronal obstruent in coda position; hence when both appear, the offglide of the vowel must be in an appendix position.

22 In this respect, then, English differs from Arabic, where the appendix does associate with a grid position.

23 For example, on phrase-level stress, readers may look at the lengthy exposition in Hogg and McCully (1987) and the references given there, including Hayes (1983), (1984), and Selkirk (1984). With regard to the relation between the foot and segmental processes, one may see, for example, the papers by Leer in Krauss (1985), and the extended discussion in Nespor and Vogel (1986).

Chapter 5   Lexical Phonology

1 An early discussion of this is in Strauss (1982a), a revision of Strauss (1979). More influential have been the papers by Kiparsky (1982a,b), which develop a set of approaches to a broad class of traditional phonological problems, and Mohanan (1982, 1986). Pesetsky (1979) and Siegel (1974) were also extremely influential in the development of this approach.

2 Cyclicity of this sort is a hallmark of classical generative work, and was developed further in Aronoff (1976), under the influence of Brame (1972a, 1974).

3 Readers skeptical about one or both of these rules are invited to wait for a discussion of these points below. The formulation given of velar softening in (1) is more for expository convenience than out of conviction with regard to the phological environment.

4 Readers may fairly be warned that this is a somewhat unorthodox reading of

the literature on lexical phonology, but one that I shall attempt to justify in the course of this chapter. A more orthodox lexical phonological view of the matter would hold either that there are no lexical rules of the first sort mentioned above, or, more likely, that such rules are theoretically insignificant or irrelevant. The issues involved here are (i) the degree to which the rules are automatic (i.e. are not governed by particular morphemes), and (ii) the degree to which the rules are transparent (in Kiparsky's sense) and interpretable as aiming at satisfaction of a phonotactic. We return to this at the very end of the chapter; see n. 57.

5 Certainly one could ask whether such post-lexical rules are necessary in all cases. That is, one should look for language-specific evidence that features filled in by such default rules really are motivated in the phonology of the language. The alternative is to leave the matter in the hands of a phonetic theory that does not utilize phonological representations, but we will not explore this question here.

6 Wonderly (1951); see also Kenstowicz and Kisseberth (1979: 35ff.).

7 One view (see Kiparsky 1982a) takes feature filters to be special cases of the rules such as in (4a) which specify the unmarked value of a lexically contrastive feature, but as I have indicated, we will consider an alternative view in section 5.1.5 below.

8 In fact, the main difference between traditional structuralism and lexical phonology as regards what information must go in the lexical phonology as distinctive, and what must go in the post-lexical phonology, derives from the difference between the 'hearer's perspective', the fundamental point of view of structuralist phonology ('phonemics takes the point of view of the hearer,' says Wells 1947), and the speaker's (or, perhaps better, the knower's) point of view in generative phonology. See the remarks on the notion of 'segment' from the structural and generative points of view at the beginning of ch. 1.

9 This notion is also discussed in the recent literature in Anderson (1969, 1975) (see further references there), and Koutsoudas, Sanders, and Noll (1974).

10 In Goldsmith (1984a), I argue that the notion of precedence in application must be kept distinct from that of linear rule ordering, on the basis of a complex set of data in Tonga; a similar situation is found in Sukuma (Goldsmith 1985a).

11 Pulleyblank (1986a: 221ff., 235–6, fn. 20, and elsewhere) notes similar doubts about this interpretation of structure-preservation, and calls the reader's attention to similar points made in Mohanan and Mohanan (1984).

12 I have changed the relation of the tiers slightly from Kiparsky's formulation for clarity of exposition, but in a way that does not affect the validity of the argument one way or the other.

13 These heterorganic clusters *would* be evaluated as slightly more marked than the homorganic sequences, since the former would have their nasal point of articulation left unspecified underlyingly. But as we shall see when we discuss underspecification theory in more detail below, this difference would be the equivalent of the difference between, for example, *nightingale*, with its long first vowel, and *Canada*, with its short first vowel – hardly the

difference between impermissible and permissible sequences.

14 Kiparsky suggests just that, in prose (1985: 100): '[w]hen not assimilated, unspecified nasals ... are assigned the unmarked feature values... If we block these default values from being assigned lexically, the Elsewhere Condition will guarantee that they will not be assigned in the postlexical phonology until after Nasal assimilation has applied there.' If we take the filter in question to act like other lexical phonological rules, and enter into an Elsewhere relationship with the rule of nasal assimilation, then the filter will not block the assimilation rule, as needed; but in that case the filter will not *block* anything, because a lexical phonological rule is always overriden by the Elsewhere Condition, and a specific lexical entry 'overrides', in that sense, a lexical rule. If such feature filters do not participate in the Elsewhere relationship to lexical phonological rules like nasal assimilation, then the filter will block assimilation within the lexicon, which would undo the whole analysis. Of the two choices, the latter appears to be Kiparsky's theoretical position; cf. Kiparsky (1985: 98) where he says that feature filters should block the application of assimilatory rules, in accordance with the change-inhibiting interpretation. We may consider a third logical possibility, and suggest that feature filters enter into an Elsewhere relationship with lexical phonological *rules* but not with lexical entries – saying, in effect, that lexical entries may *not* violate the feature filters, but lexical phonological rules can create structures that do. This position requires abandoning one of the most important aspects of the interpretation of lexical phonology under discussion here, the notion that lexical entries and lexical phonological rules are the same kind of theoretical object, in the relevent respect. The correct solution to the present case, I will suggest shortly, is that no negative feature filter is required for most of the work, and the one that is required derives from the notion of autosegment licensing.

15 Kiparsky suggests that the principles of underspecification theory should surely not be taken to mean ruling out the possibility of specifically referring to the unmarked value. We believe that this *is* an important characteristic of an underspecification model of phonology, and that such references to unmarked values should in fact be ruled out. However, segmental positions that are unspecified for a particular feature (i.e. with respect to a particular autosegmental chart) may undergo assimilatory rules that other segments do not undergo in order to satisfy well-formedness conditions (phonotactics), such as that all consonants must be associated with a point of articulation at the phonetic level.

16 Put another way, the nasal segment that is still unmarked for point of articulation by the end of the phonology surfaces as an alveolar.

17 There may be further restrictions on point of articulation in the coda; this analysis is consistent with or without further conditions.

18 See the discussion of the Obligatory Contour Principle in ch. 6.

19 This conception of rule application is somewhat orthogonal to the subject of this book, which is the nature of representations rather than of rules or phonological levels. However, the conception of rule application in lexical phonology as Kiparsky has proposed it incorporates the traditional *SPE*

conception of rule application, which is precisely what is at issue here. For more discussion, see pp. 322ff.

20 This follows Kiparsky's analysis as well.

21 See Schuh (1972), and Newman and Salim (1981), though our analysis differs from theirs considerably. Within the present context, the synchronic and historical motivation for Klingenheben's Law is clear: the coda does not license a point of articulation. Hausa differs from most languages with this restriction in that it turns neither to epenthesis nor to cluster simplification for a resolution, but to glide formation.

22 No structural description need be specified if we follow the proposal mentioned just above (and discussed further in section 6.5) that word-level rules apply if and only if their application reduces the set of violations of phonotactic conditions.

23 This point is made forcefully in Halle and Vergnaud (1982). The fact that the specific rule of Sukuma mentioned also applies post-lexically may make the example less than an ideal illustrative example. See Pulleyblank (1986a: 221–3) for a parallel discussion.

24 Other recent work has applied a version of structure preservation that does not inhibit changes, but subjects the output of rules to higher-level procedures to insure structure preservation. Hayes (1986b), for example, develops a 'convention' for the feature [sonorant] whereby a segment that becomes [+nasal] by rule also becomes [+sonorant], and a non-continuant that loses its [+nasal] specification also becomes [−sonorant].

25 At least one suggestion has been made, in principle, that would indirectly put constraints on such a move; but the constraints are so indirect, and require such considerable analysis in order to put the constraints into effect, that in actual practice the braking effect of the constraint would be seen less rather than more frequently. The suggestion comes from Mohanan's work (1982), and involves the proposal that, if a rule is present in more than one stratum, then the sequence of strata that the rule is found in must form a continuous span in the phonology.

26 Thus once again, we see that lexical phonology is structured in such a way as to permit the maximal *phonologization* possible of processes affecting phonological segments, as opposed to any morphologization. To put the matter in a way suggested to me by B. Darden, the Strict Cycle Condition began as the Alternation Condition of Kiparsky (1968), which said something quite unobjectionable – that neutralizing sound changes effected strictly morpheme-internally are no longer synchronically active, having become part of the underlying form. Kiparsky (1982a) and later work on lexical phonology takes this not as a result of how languages are learned, but rather as a fact about the formalism of rule application, and interprets it as a formal device that would *explain*, in some sense, why a proposed rule should fail to apply morpheme-internally. To the extent that such an account is available to explain non-applications, to be sure, more freedom of movement to propose phonological rules applying at morpheme boundary is available to the linguist.

27 There is another process that lowers a suffixal *i* to *e*, and one might therefore

wonder whether this form *heer* derives from an intermediate *ha+er*. However, the rule lowering *i* applies after radicals with a mid vowel *e* or *o*, not after radicals with a low vowel *a*. Thus a radical of the form *CaC* takes the applied suffix *-ir-*, not *-er-*.

28   There is a useful discussion of the notions dealt with here in Steriade (1987), which unfortunately contains a large number of arguments that are vitiated by a change in the vowel features that are adopted for dealing with vowel systems; see ch. 6 below, and Goldsmith (1987a). To my knowledge, the first suggestion of underspecification theory as it has come to be known within the generative context is in Ringen (1988, [1975]), in the context of the treatment of vowel harmony. Clements (1988) also presents a clear discussion of the issues. As this book goes to press, a collection of papers on the subject has appeared (*Phonology*, vol. 5, n. 2, 1988), with a useful introduction by Archangeli (1988b).

29   They also are interpreted very naturally, as Pulleyblank (1983, 1986a) was the first to emphasize, within the framework of autosegmental phonology. We have used these notions quite a few times already, especially in attempting to license only specified (contrastive) features.

30   There is another important difference between these two types of restriction countenanced by underspecification theory. Feature specifications that are ruled out by the first ('eliminate redundant features') principle are ruled out throughout the lexical phonology; sonorants in English, that is to say, are not voiced at any point in the lexical phonology. Such constraints are thus part of the 'structure' that is 'preserved' during the lexical phonology. The rules that fill in the unmarked value of an opposition are lexical rules, and hence are not conceived of as *changing* a structure, or, put another way, are not thought of as creating a structure that is preserved throughout the lexical phonology.

31   And of course therein lies a difficult question: how are such generalizations collected by the language learner? Does he/she rely crucially on alternations noticed when a word takes on a suffix, and finds its vowel shortened, or is that just evidential icing on the cake? The question remains a knotty one.

32   See, for example, Dressler (1985) for a thorough review of the positions in the literature on the relations among the various rules of morphophonemics, in the broadest sense.

33   We would prefer a different formulation of this constraint, as careful readers will observe. Bearing in mind our discussion of the Catalan nasals above and the fact that syllabification is imposed on the word-level representation in Spanish, we would say that in Spanish no point of articulation specification may be present on a coda that is specified as nasal.

34   To wit, all words with main stress on the final syllable, such as *boutique* or *Peking*. Of course, if the only stressed syllable in a word is the final syllable, it will receive the main stress, but this is hardly an exception to anything (e.g. *balloon, remark*). Other questions arise with words such as *legislature*, which we cannot go into here.

35   An important case that is difficult for some speakers to judge involves the pair *compensation/condensation*, which Chomsky and Halle (1968) offer as a pair in which the second syllable is unstressed in the first word but stressed

in the second, as their theory predicts, based on their derivations from *compensate* and *condense*, respectively. Halle and Vergnaud (1987) mention three other examples, where a secondary stress is perceived on the second vowel, suggesting earlier derivation, including *electricity*, *infestation*. Judgments here are too delicate for this linguist's ear.

36   Including notably Liberman and Prince (1977) and Hayes (1982), among many others.

37   Hayes suggests that assigning stress to such final syllables by means of a process distinct from the main English stress rule is supported by the prediction that in such cases the *next* stress, scanning leftward, will be assigned in a quantity-sensitive way (so-called 'weak retraction'). On Hayes's account, no other stress assignment rule is quantity-sensitive, for the rhythmic assignment of stresses to alternating syllables leftward across a word will be done in a quantity-insensitive fashion.

38   A good range of simple words like *vanilla* exist with stress on the penultimate where the principle suggested here predicts antepenultimate stress. These accounts have only limited things to say about such words. A large proportion of them have a stressed [æ] and end in schwa, as in *Alabama*, but as *vanilla* shows, not all of them have these properties. All of the examples in (57b, d, e) of ch. 4 above similarly illustrate aspects of non-predictable English stress patterns.

39   A medieval form of refereeing.

40   The same rule appears to apply in an extended way to certain heavy syllables when they form Latinate prefixes (e.g. *con-taminate*, where the prefix has a reduced vowel because the following syllable is stressed, contrasting with *con-sternation*, where the prefix has a full vowel).

41   This rule and trisyllabic shortening share a common feature that our present analysis does not capture, unfortunately. There are generalizations about *derived* stress feet in English that we have not considered; see Selkirk (1980) for some remarks that can be interpreted along these lines. She offers an account that includes feet with two unstressed syllables following a stress syllable (as in *Canada*, for example). While such feet have no systematic place as such in the analysis discussed in the text above, there do appear to be significant properties about such derived feet whose recognition would allow us to simplify the formulation of the destressing rules. For example, such derived feet do not normally contain more than three moras. Thus, none of the syllables in such ternary feet can be heavy, either by reason of containing a long vowel or by reason of containing a closed syllable. This generalization stands behind trisyllabic shortening, to be sure. (A small class of exceptions exists to this generalization, such as the much-discussed *obesity*, with a long stressed vowel, or the proper name *Sepulveda*, with main stress on the second syllable.)

42   She lists five: *gruesome, hapless, feckless, winsome,* and *fulsome*.

43   The nominalizing suffix *-al* attaches only to bases that independently have final stress, as in *arriv-al, renew-al* (cf. *arrive, renew*), but not **perjural* (cf. *perjure*, with initial stress).

44   A suggestion along these lines was made by Pesetsky (1979) for Russian, and by Strauss (1982b) for English, on the basis of some material we will reconsider below.

45   This division even suggests a natural difference in the phonological behavior of the two kinds of 'zero derivation' in English, the word-category change that can take place with no overt affixation, as in pairs like *contrast* (v) / *cóntrast* (n). In cases such as this one, where the two forms are either apparently of the same degree of 'basicness' in the language, or in cases where the noun appears to be derived from the verb, we find that the noun displays the stress pattern appropriate for nouns, and the verb displays the stress pattern appropriate for verbs. On the other hand, when a verb is clearly created from a noun, it retains the stress pattern of the base noun (e.g. *to pattern*, as in 'the adjectives pattern with the nouns phonologically'; as we have already seen, a verb ending with two consonants would normally receive final stress, which is not the case here).

46   Selkirk (1982b: 104) cites Siegel (1977) as having drawn the same conclusion regarding the behavior of *-ment*. As I read Siegel, she proposes that the *ment* of *government* is neither class I nor class II, but rather that there is a boundary symbol (#) inserted with all class II suffixes which can also irregularly appear in a few underlying forms. We may distinguish three kinds of morphophonological theories: the boundary-driven type, as in Chomsky and Halle, and Siegel; the rule-block driven type, as in lexical phonology; and the domain-labeling type, adumbrated in Selkirk and discussed further below.

47   An excellent discussion of this can be found in Aronoff and Sridhar (1983). I have also benefited from reading Badecker (1988).

48   See Aronoff and Sridhar (1983), Mohanan (1986), and references there.

49   Williams (1981) proposes an account according to which the correct structure is the one predicted by the Affix Ordering Generalization (i.e. [un[[grammatical]ity]]]), and he suggests that structures of the form [x[y[z]]] will be recognized by the grammar as being 'related' to other structures of the form [x[y]]. If correct, this proposal would have the amusing property of relating non-constituents in compounds as well, such as *union suit* and *union suit-tailor* (since the latter has the form [union [suit [tailor]]]. The former is a kind of garment; the latter a progressive worker. One can construct other examples of this sort, such as *maternity suit* and *maternity suit-case*.

50   Pesetsky (1985) discusses these problems from a different perspective, though my remarks in this section and the next indicate why I am not in sympathy with his particular solution to the 'paradox'.

51   An interesting precursor of this approach is seen in Chung (1983), who is concerned with choosing between transderivational constraints and cyclic accounts in the face of phonological processes in Chamorro that optionally take 'earlier cycles" stress specification into account.

52   Halle and Vergnaud actually achieve this through a mechanism of *conflation* of the second and third row, essentially the same as what we referred to as suppression of secondary stresses. This leaves a grid mark on the second row only on the syllable with the main (word) stress, which is then copied.

53   We are using stratal terminology here, to be sure, though in the final analysis we will not conclude that it is appropriate to account for the present data.

54   This formulation of the restriction is tree-oriented in its statement. A less constituent-oriented formulation would be that the right-to-left perfect grid (quantity-insensitive) application that constitutes stress retraction applies only to stretches of unstressed syllables, which is the way Perfect Grid always works; in addition, there is no forward clash ovrride, again the unmarked case. Kiparsky (1982a) offers one example of a case where stress retraction does not work this way: the case of *solidify*, where he suggests that the 'unfooted' *i* of the suffix *-ify* is enough to trigger stress retraction. It is equally reasonable to suppose that this is a case of close juncture, precisely as Kiparsky proposes for such 'irregular' forms as *democratize*, where stress retraction does indeed appear to have overriden the stress pattern of the base.

55   The same point is made on independent grounds in Fabb (1985).

56   See the typological remarks in Booij and Rubach (1987).

57   The skeptic who was trying to establish a case against treating trisyllabic shortening as a phonological process might proceed suffix by suffix, looking a bit more closely to see whether alternations (like the one cited by Kiparsky 1982a, *omen/ominous*) holds up across the range of words in the English lexicon. A glance through a backwards-alphabetized dictionary, such as *Walker's Rhyming Dictionary*, suggests that there is no large class of words formed with an *-ous* suffix attached to an independently existing word base. Some exist, to be sure, such as *humorous*, related to *humor*, or *scandalous*, or *perilous*. Others, like *ominous*, the example mentioned by Kiparsky, are rather distantly related to the word that looks like it might serve (or might once have served) as a base, since calling a sky ominous is not to call it an omen. Still others look like they are formed with an *-ous* suffix but have no plausible base; the pattern *felicitous/felicity* is not matched by *jealous/jelly*.

     A glance, then, at the eight or nine hundred *-ous* suffixes in English turns up only one clear case where the adjectival form has a short vowel, but the related base has a long vowel: *libido/libidinous*, but the short vowel [ɪ] found in the derived form is not the expected vowel (we would expect ɛ]) — nor, of course, is the *in* expected. A good number exist that violate the shortening prediction, such as *cretinous, mountainous, libellous, poisonous, sonorous* (for some speakers, such as myself), *scrupulous, numerous, gratuitous, odorous,* or *cumulous*. It may not be accidental that most of these examples involve long round vowels, which may simply not fall under the generalization expressed by trisyllabic shortening; but be that as it may, the *-ous* forms cannot be said to provide positive suport for the existence of trisyllabic shortening as a rule in English. What the examples do suggest, however, is that the direction of change that is induced in the stem by the juxtaposition of affixes is in the direction of what would be found in nonderived forms.

Chapter 6   Further Issues

1   Or rather, the intersection of these intersecting sets with the set of segments in the language at hand. On a related point, one interpretation of a theory of