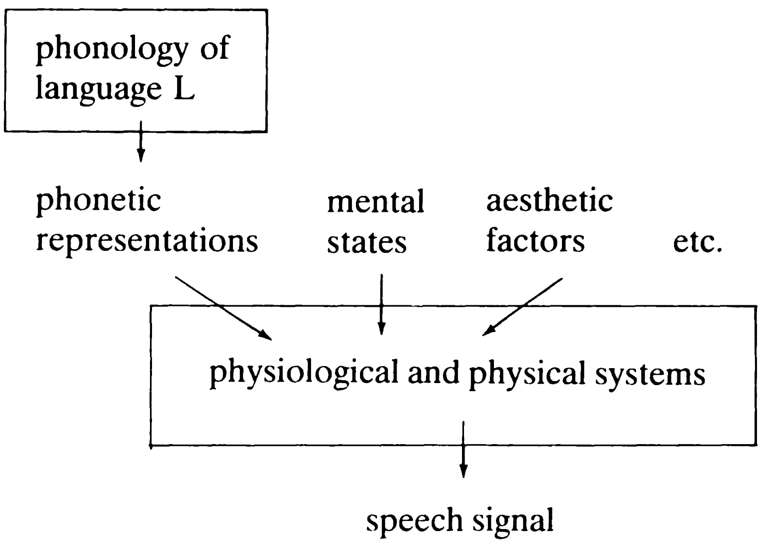
Our assumptions also lead us to the conclusion that implementational operations dissolve structure, and cannot therefore be structure-preserving (section 6.7). Finally, I shall show that the theory of phonetic implementation captures all the basic insights that lay behind the practice of classical phonemicists, without any of the undesirable consequences of the theory of classical phonemics (section 6.8).

6.2. SPEECH AS IMPLEMENTATION OF PHONETIC REPRESENTATION

It is well known that phonetic representations, which constitute the output of the phonological component of the grammar, are not direct records of speech. Rather, they are abstractions or idealizations of the linguistically relevant aspects of speech. No phonetic representation contains, for example, information about the voice quality of a speaker, or the effects of pharyngeal constriction due to anger or excitement. All phonologists and phoneticians agree that these details are different in nature from, say, the specification of aspiration of voiceless stops in English.

Phonetic representations may be thought of as instructions that the language module gives to the vocal organs in order to implement an utterance (Postal 1968). The details of voice quality (e.g. male vs. female, adult vs. child) are consequences of the physical properties of the articulatory system of each speaker, and are not included in the instructions that the language module issues to the vocal apparatus. Pharyngeal tension due to anger is not part of the instructions from the language module either, because it derives from the interface between the 'module' of mental states and the physiological system. One may, therefore, conceptualize the relation between phonetic representations and the speech signal as follows:

(10)



The module of the physiological or articulatory system implements the instructions from the phonological module in the production of speech, accepting, in the implementation, instructions from nonlinguistic sources as well (e.g. mental or physical states).

If we define the subject matter of the theory of phonetics as the physical implementation of speech, we see that phonetics is concerned with the following types of questions:

A. What is the nature of the linguistic *input* to phonetics, namely, phonetic representations?
B. What is the nature of the *output* of phonetics, namely, the speech signal?
C. What is the nature of the mapping between the linguistic input and the speech signal?
D. What is the nature of the mapping between the nonlinguistic inputs (mental states, articulatory setting . . .) and the speech signal?

Most traditional phonetics has been addressed to question A—C, though recent years have also seen an interest in questions of type D (e.g. Laver (1980), Nolan (1983)). We may think of B as the problem of PHYSICAL PHONETICS, and A and C as the problems of LINGUISTIC PHONETICS. Question A, for example, covers the distinctive features used in the characterization of phonetic representations, and question C covers the way distinctive features are articulatorily implemented (e.g. Ladefoged (1980), Halle (1983)). For the sake of subsequent discussion, I shall refer to the mapping covered by question C as the UNIVERSAL RULES OF PHYSIOLOGICAL IMPLEMENTATION of phonetic representations.

Observe that it is questions A and C which make the cooperation of the phonologist and the phonetician imperative. The phonologist must give a characterisation of the output of phonology, and the phonetician must give a characterisation of the input to phonetics, and the two characterisations should match (question A). Similarly, the phonologist and the phonetician should together decide what kinds of mappings belong to the domain of phonology, and what kinds of mappings belong to the domain of the universal physiological implementation of phonetic representations. These two questions are interdependent: the theory of physiological implementation depends on our conception of phonetic representations, and the conception of phonetic representations depends on our knowledge of how much information in the speech signal can be left to the rules of physiological implementation.

The issue of how much should be included in the instructions from the language module and how much can be left to other systems is not a trivial one. Take, for example, the specification of pace or tempo of speech (number of syllables per second). Whether one speaks at a rate of three syllables or six syllables per second (unlike the range of pitch at which one speaks) is not determined by the physical properties of the articulatory system. The rate of speech is a parameter for which each individual must choose a value in implementing speech. Once having made the choice, speakers generally do not change their rate of speech. The values chosen for articulatory parameters of this kind may be called the "articulatory setting" for each speaker (Honikman 1964), and is similar to the general instructions for spacing, margins, indentation for paragraphing etc. as part of the formatting instructions in a journal style sheet ("indexical features" in Abercrombie (1967)). Another analogy for articulatory setting would be the style of handwriting that characterises individual writers.

There are differences across speech communities in their articulatory setting. The speech one hears in Texas, for example, is slower in pace than the speech in Boston. In order to account for such differences, it is necessary to assume that a learner's choice of articulatory setting (which also includes the range of jaw movements, the degree of tension in making articulatory contacts in general, etc.) is influenced by the choice made by the speakers that the learner is exposed to, but it does not force us to assume that these choices are part of the knowledge of the language as such.[3] In fact, bilingual speakers continue using the same pace, the same 'clenching of teeth', etc., when they switch languages. These properties are part of the socio-cultural knowledge or mannerisms, rather than linguistic knowledge. Similar remarks may be made about the lowering of the soft palate, which is nonlinguistic in the production of the "nasal twang" which permeates throughout the speech, but linguistic in the production of nasal consonants.

More complex are the parameters which are set in the articulatory system in order to implement specific instructions from the language module. Consider, for example, the production of voiced plosives. The instructions from the language module would be [+voice] (= keep the vocal cords vibrating), [−nasal] (= raise the soft palate), [−continuant] (= make a central contact in the oral tract), and [−lateral] (= don't lower the sides of the tongue). In order to implement the instruction to keep the vocal cords vibrating, it is necessary for the supraglottal air pressure to be sufficiently lower than the subglottal air pressure. This pressure difference is maintained by speakers of languages like English by lowering the larynx, and thereby increasing the volume of the supraglottal cavity (Ladefoged 1971). Now, it has been discovered that the speakers of many Indian languages maintain an incomplete closure of the soft palate in the production of voiced plosives, the resultant leakage of air maintaining the pressure difference (Rothenberg 1968, Nihalani 1975). Another strategy is to increase the volume of the supraglottal cavity by expanding the pharyngeal walls. Thus, speakers with different language backgrounds choose different combinations of parameters for the implementation of voicing in stops, namely, lowering the larynx, lowering the soft palate, and expanding the pharynx.

A similar situation appears in the articulatory implementation of implosives. Ladefoged (1980), for example, points out that implosives in Hausa are said with creaky voicing, while those in Kalabari are said with full voicing. While the implosives investigated by Ladefoged are produced without an ingressive airflow at the point of release (Ladefoged 1971), those in Sindhi show an ingressive airflow (Nihalani 1985b).

From the kinds of data available, it is not yet clear whether the differences in the choice of physiological parameters of this kind are part of the idiosyncratic choice of individual speakers who happen to have been selected as subjects for the investigation, or whether they are properties of the community of speakers who use the same language. Assuming, for the present, that the situation corresponds to what Ladefoged (1980) claims, namely, that there is a correlation between the language and the parameters, we still have to decide whether these properties, like rate of speech, should be dealt with as part of the articulatory setting, or whether it is necessary to include them in the phonetic representations as part of the instructions to the articulatory system. It might take several years of research on the physiological implementation of phonetic representations before we can arrive at satisfactory answers.

I shall now turn to available evidence on the nature of phonetic representations, and try to establish the boundary line between phonetics and phonology. After spelling out the kinds of information that the instructions from the language module must contain, and how these instructions can be stated in phonological theory, I shall go on to examine the consequences of these proposals for the theory of Lexical Phonology as sketched in (5).

6.3. THE NATURE OF PHONETIC REPRESENTATIONS

6.3.1. *Phonetic Features on a Scale*

The following assumption about phonetic representations appears to be uncontroversial:

(11) Phonetic representations are interpretable on the basis of universally applicable conventions (i.e. the instructions they contain can be implemented by the language independent physiological module).

In addition to (11), we also find assumption (12) being shared by most practitioners of phonology:

(12) Phonetic representations contain strings of phonetic segments.

SPE proposes a more specific assumption about the kinds of segments phonetic representations contain:

(13) Each segment in a phonetic representation is composed of a set of feature specifications along a scale.

While phonological representations contain segments which are specified in terms of binary features, phonetic representations make use of scalar values:

> The phonetic representation consists of a sequence of 'phonetic segments,' each of which is nothing other than a set of 'phonetic feature specifications.' A phonetic feature specification consists of a 'phonetic scale' (called a 'phonetic feature') and an integer indicating the position of the phonetic segment in question along this scale. The phonetic scales form a predetermined universal set, namely, the '(phonetic) distinctive features.' Thus a particular segment might be marked as 'noncontinuant' (i.e., 'minus' with respect to the phonetic feature 'continuant'), 'highly aspirated,' 'nonvoiced,' etc. In short, a phonetic representation is a 'phonetic matrix' in which the columns correspond to segments and the rows to features and in which each entry states the *extent to which a given segment possesses the corresponding feature* [italics mine; footnote omitted] (SPE, p. 164)

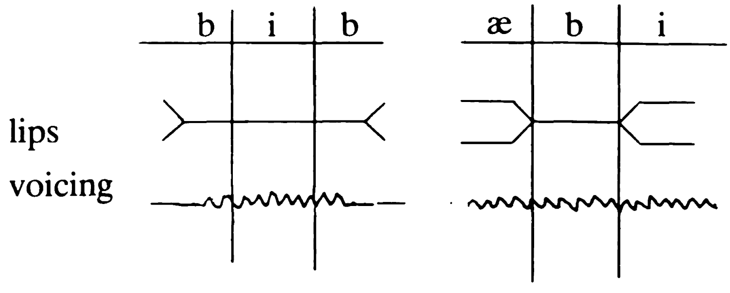
The need for gradient feature specifications (= feature specifications along a continuum, as opposed to binary feature specifications) is illustrated by the behaviour of voiceless aspirated stops in English. The degree of aspiration in English is dependent on the degree of stress: the greater the degree of stress, the greater the degree of aspiration. (This is not the case in languages like Malayalam and Hindi.) Thus, in *participation*, the **p** in the fourth syllable is more aspirated than the **t** in the second syllable, because the fourth syllable carries greater stress than the second syllable. Adapting the SPE alpha notation as '*n* notation' where *n* is a variable ranging over the integers, we may formally state the rule of aspiration as follows:

(14)
$$\begin{bmatrix}-\text{cont}\\ -\text{son}\\ -\text{voice}\end{bmatrix} \rightarrow [n \text{ aspirated}] \,/\, [_{\substack{\text{syll}\\ [n \text{ stress}]}} \;\_\_\_$$

(read: a voiceless stop is aspirated at the beginning of a stressed syllable, the degree of aspiration corresponding to the degree of stress)[4]

Another example of the need to use scalar features is found in the facts regarding the degree of voice in voiced obstruents. It is well known that in *bib*, when said in isolation, the **b** at the beginning and end are less voiced than the **b** in *abbey*:

(15)



The general principle governing voicing of obstruents in English may be stated as: obstruents are fully voiced only when preceded and followed by voiced segments.

Though physiologically motivated, this phenomenon of devoicing cannot be relegated to the module of universal physiological implementation, since there are languages which do not exhibit this kind of devoicing. As noted for Hindi and Telugu by Prasad (1950) and Rothenberg (1968), and for Sindhi by Nihalani (1975), the voicing of word-initial stops begins fairly early in Indian languages. My own informal examination of voiced stops in Malayalam shows that the voice onset time in this language is similar to that of Hindi, Telugu, and Sindhi, not to that of English.[5] Given that the word-initial voiced stops in Hindi, Telugu, Sindhi and Malayalam are not devoiced, it is necessary to specify the devoicing of voiced obstruents in the grammar of English, or its absence in the grammars of Telugu, Sindhi and Malayalam. Such a rule will have to refer, not to the binary values of [voice], but to the scalar values of voice.

Introducing scalar values of features towards the end of phonological operations takes us one step closer to the continuum of speech. I shall now take a further step, and claim that segments do not exist at the level of phonetic representation.

6.3.2. *How Abstract are Phonetic Representations?*

For the purposes of exposition, let us make the following assumption about phonetic representations as the most abstract position:

(16) A phonetic representation contains strings of segments which are specified in terms of binary features.

A phonetic representation of type (16) differs from the speech signal in three ways: it abstracts away (i) the degree of articulatory gestures, (ii) the transition from one gesture to another, and (iii) the overlap between articulatory gestures. The continuum of the DEGREE of articulatory gestures is represented in terms of binary categories, e.g., 'less' nasal and 'more' nasal become minus nasal and plus nasal respectively. The continuum of the TRANSITION from one articulatory gesture to another is represented as a discrete abrupt break from one gesture to another, e.g. the gradual transition of the tongue position from a high position to a mid position at the beginning of *yes* is represented as a sequence of two discrete entities, namely, [+high] followed by [−high, −low]. Diagrammatically, what is involved here is a representation of the physical reality of (17a) as (17b):

(17) a. speech signal b. phonetic representation

+high −high +high −high

The assumption implicit in (17) is that (17b) will be mapped onto (17a) by the universal rules of physiological implementation.

The OVERLAP, or lack of synchronization between articulatory gestures along various dimensions, is ironed out to make the gestures coincide. For example, even though the soft palate is lowered a few milliseconds prior to the contact between the alveolar ridge and the tip of the tongue in the word *in*, the word is represented as

(18)

	i	n
nasal	−	+
continuant	+	−

the assumption being (as in the case of (17), that the lack of alignment between the two articulatory gestures will be taken care of in the module of physiological implementation, converting (18) to (19):

(19)

	i	n
nasal	−	+
continuant	+	−

If we do not make these three types of idealizations (i.e. filtering out the effects of the degree, transition, and coordination of articulatory gestures as in (16)), we arrive at the most concrete hypothesis of phonetic representations (the one that is closest to the speech signal), stated in (20):

(20) A phonetic representation must specify the degree, transition, and overlap of articulatory gestures.

The SPE conception of phonetic representations is less abstract than (16) in that it allows for the incorporation of the degree of articulatory gestures, but more abstract than (20) in that it factors out the transition, and, more important for our purposes, the overlap. The question that I would like to raise is: should the feature specifications along different parameters of articulation in a phonetic representation be in alignment with each other, or should we allow the articulatory gestures to overlap? This question may be rephrased in terms of (12) as: should phonetic representations contain segments?

6.3.3. *The Status of Segments in Phonetic Representations*

The SPE conception of phonetic representations is one that incorporates (12). In contrast, I would like to explore the consequences of assuming that phonetic representations are more like the one in (19), in which the feature specifications for different articulatory gestures do not coincide with each other, but are aligned to a 'timer' independently of each other. This proposal is illustrated in the following steps in the phonological derivation for *Ben* and *bet* ((21)—(24)):

(21)

	Ben			*bet*		
	C	V	C	C	V	C
continuant	−	+	−	−	+	−
nasal	−	−	+	−	−	−
voice	+	+	+	+	+	−
high		−			−	
low		−			−	

(22) scalar features

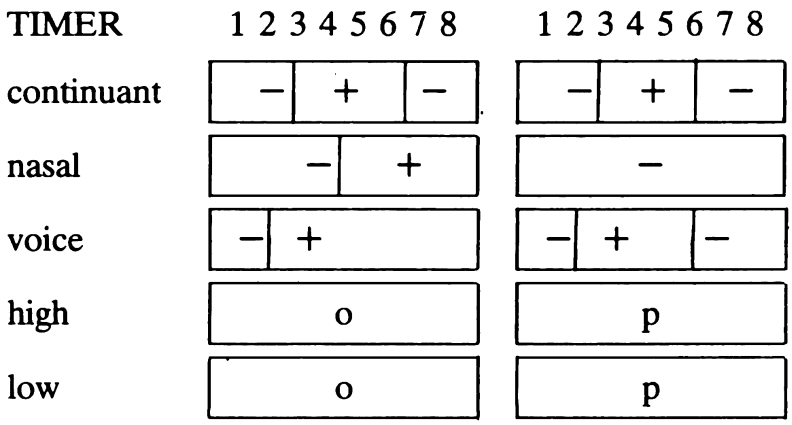
	Ben			*bet*		
	C	V	C	C	V	C
continuant	−	+	−	−	+	−
nasal	−	−	+	−	−	−
voice	+	+	+	+	+	−
high		o			p	
low		o			p	

(o and p indicate scalar values.)

(23) timer alignment

	C	V	C	C	V	C
TIMER	1 2	3 4 5 6	7 8	1 2	3 4 5	6 7 8
continuant	+	+	−	−	+	−
nasal	−	−	+	−	−	+
voice	+	+	+	+	+	−
high		o			p	
low		o			p	

(24) feature alignment with timer



The phonetic representation I would like to advocate is the one given in (24) (similar to (19)), which, to borrow terminology from autosegmental phonology, is an orchestrated score of phonetic features aligned to the time axis.[6] The representation in (21) corresponds to the conception in (16). The next step in the derivation, (22), specifying the scalar values of features, corresponds to the SPE position. Further steps in the derivation, (23) and (24), take us closer to (but do not reach) the position in (20). The mapping from (22) to (23) aligns the segments or features linked to the Xs at the skeletal tier to a timing axis, and eliminates the Xs. The mapping from (23) to (24) frees the individual feature specifications from each other and aligns them (independently of each other) with the time axis.

There are several advantages in thinking of phonetic representations in this fashion. First, we have a formal apparatus for stating the timing rules that govern segment duration in natural languages, as in the mapping from (22) to (23). There have been a number of studies in the recent literature which indicate that the durational phenomena reported in House & Fairbanks (1953),[7] responsible for the timing difference between the vowels in *Ben* and *bet*, are not entirely universal. Thus, Port et al. (1980) and Port & Mitleb (1983) show that Japanese has temporal compensatory effects which keep the overall duration of words constant for a given number of moras, while Arabic has none. In Japanese, the duration of a vowel is affected by both the preceding and the following consonants: the longer the consonant on either side, the shorter the vowel. Thus [a] has a duration of 79 msec before [t], 89 before [d], and [92] before [r], which correlates with the durations of [t], [d] and [r], namely, 55 msec, 36 msec and 25 msec. The duration of [a] after [t] is 90 msec; after [d] it is 96 msec, and after [r] it is 100 msec. In contrast, vowel duration remains constant in Arabic irrespective of the duration of the preceding or following consonant. English appears to lie halfway between Japanese and Arabic: a vowel is shortened if the following consonant's duration is long, but the vowel is unaffected by the duration of the preceding consonant.

One of the advantages of the formalism in (23), (24) is that it lends itself to the statement of segment durations of this kind.

More important, perhaps, is the promise that (24) holds for the statement of phonetic phenomena due to the misalignment in the coordination of articulatory gestures. It is well known, for example, that speakers of English insert a short stop between a nasal or a lateral and a following fricative:

(25)

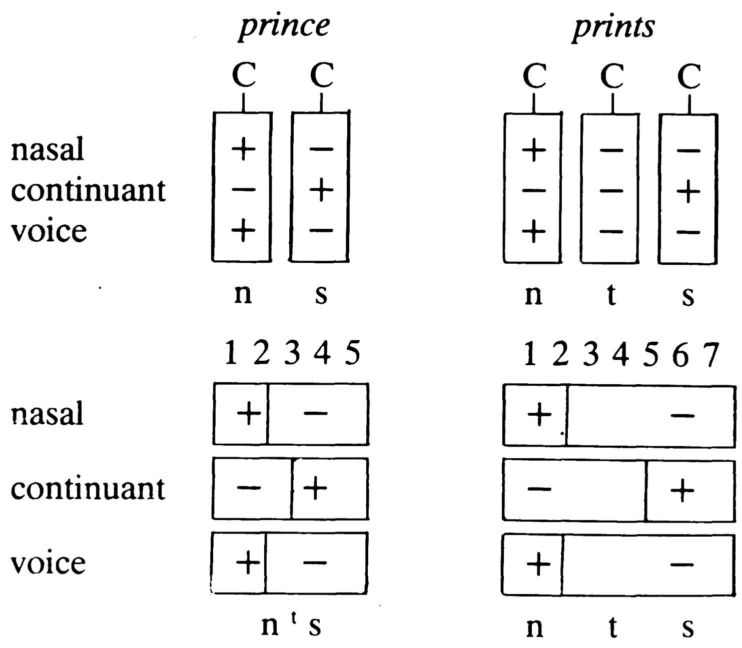
	A	B
	a. prince [. . $n^{t}s$]	b. prints [. . nts]
	c. wins [. . $n^{d}z$]	d. winds [. . ndz]
	e. pulse [. . $l^{t}s$]	f. cults [. . lts]
	g. wells [$l^{d}z$]	h. welds [. . ldz]

Most speakers of English distinguish between A and B in terms of the duration of the stop. The stop between [n] and [s] in *prince*, for example, is about 25% shorter than the corresponding stop in *prints* (Fourakis 1980).

The articulatory basis of the stop insertion in (25A) is fairly obvious. What is happening in (25a,c) is a misalignment or lack of coordination between two articulatory gestures, namely, the raising of the soft palate and the release of the contact between the alveolar ridge and the tip of the tongue. If the two gestures are simultaneous, the result is a perfect [ns] or [nz]. If, on the other hand, the soft palate is raised a few milliseconds prior to the release of the tongue tip, the result is a period of complete stoppage of airflow, which would be [^{d}] if the vocal cords are still vibrating, and [^{t}] if they are not.

In terms of the conception of phonetic representations in (24), the contrast between [$n^{t}s$] and [nts] may be given as follows:

(26)



The stop insertion in (25e,g) involves a similar misalignment of the raising of the sides of the tongue and the release of the tongue tip.

Given that what is happening here is so tightly bound up with the timing of the articulatory gestures in the implementation of speech, one may enquire whether the phenomenon that we are dealing with belongs to the linguistic system, or whether it should be left to the physiological system. The way to find out is to see if there are languages which do not exhibit this phenomenon. Now, Fourakis (1980) reports that the stop insertion between sonorants and fricatives does not happen in South African English. The fact that the delay in the release of the oral closure in relation to the closure of the nasal cavity appears in American English, but not in South African English, shows that it is a *physiologically motivated linguistic phenomenon*, not a *purely* physiological phenomenon, and therefore must be represented in the outputs of grammars of languages. What is happening here is a 'grammaticalisation', or an absorption of the entities of the extralinguistic systems into the linguistic systems, parallel to the grammaticalization of concepts like 'agent' and 'experiencer' in syntax. Agency and experience are entities that belong to the conceptual system of human beings, but they have been grammaticalised and absorbed into the syntactic module.

A clearer instance of the grammaticalisation of the misalignment of articulatory gestures is found in the insertion of stops between a stop and a nasal in Malayalam. Thus, words like **swapnam** 'dream' and **yugmam** 'pair' are pronounced with a short [^{t}] and [^{b}] as [swap^{t}nam] and [yug^{b}mam]. The phonetic notation that represents this misalignment, not observed in languages like English, is as follows:

(27) *swap^{t}nam*

	1	2	3	4	5	6	7	8	9
voice	−						+		
nasal	−						+		
coronal	−				+				
continuant	−								
		p			t		n		

The stretch of time at 5/6 is associated with the features [−voice, +coronal, −continuant, −nasal], which is what is heard as a brief [^{t}]. What happens in the case of **yugmam** is similar, except that the voicing in **g** simply continues to **m**.

What makes Stop Insertion in Malayalam interesting is its interaction with a rule that changes the voiceless dental stop **t̪** to a voiced alveolar lateral when followed by an obstruent in colloquial speech.

Thus, relatively careful pronunciations like [ut̲bhawam] ‘beginning’, [saakṣaat̲kaar̄am] ‘realization’, and [ut̲sawam] ‘festival’ are replaced by [ulbhawam], [saakṣaalkaar̄am] and [ulsawam] in more casual speech. One also finds careful/casual pairs like [aat̲maawə]/[aal^{p}maawə] ‘soul’, and [pat̲mam]/[pal^{p}mam] ‘lotus’, but never *[aalmaawə] and *[palmam], clearly indicating that the rule that inserts the stop between t̲ and **m** feeds the rule that changes t̲ to **l**. I give the derivation for [pal^{p}mam] in (28):

(28)

<table>
<tr><td></td><td colspan="2">1 2 3 4 5 6 7</td></tr>
<tr><td>voice</td><td>−</td><td>+</td></tr>
<tr><td>nasal</td><td>−</td><td>+</td></tr>
<tr><td>coronal</td><td>+</td><td>−</td></tr>
<tr><td>lateral</td><td colspan="2">−</td></tr>
<tr><td>continuant</td><td colspan="2">−</td></tr>
<tr><td></td><td>t̲</td><td>m</td></tr>
</table>

<table>
<tr><td></td><td colspan="3">1 2 3 4 5 6 7</td><td></td></tr>
<tr><td>voice</td><td colspan="2">−</td><td>+</td><td></td></tr>
<tr><td>nasal</td><td colspan="2">−</td><td>+</td><td>by stop insertion</td></tr>
<tr><td>coronal</td><td>+</td><td colspan="2">−</td><td></td></tr>
<tr><td>lateral</td><td colspan="3">−</td><td></td></tr>
<tr><td>continuant</td><td colspan="3">−</td><td></td></tr>
<tr><td></td><td>t̲</td><td>p</td><td>m</td><td></td></tr>
</table>

<table>
<tr><td></td><td colspan="3">1 2 3 4 5 6 7</td><td></td></tr>
<tr><td>voice</td><td>+</td><td>−</td><td>+</td><td></td></tr>
<tr><td>nasal</td><td colspan="2"></td><td>+</td><td>by t̲ → l/___stop</td></tr>
<tr><td>coronal</td><td>+</td><td colspan="2">−</td><td></td></tr>
<tr><td>lateral</td><td>+</td><td colspan="2">−</td><td></td></tr>
<tr><td>continuant</td><td colspan="3">−</td><td></td></tr>
<tr><td></td><td>l</td><td>p</td><td>m</td><td></td></tr>
</table>

The third advantage of allowing phonetic representations to have feature overlaps aligned against time specification is that it provides a way of encoding the phonetic correlates of abstract entities like stress. It is well-known that [stress] is a phonological entity (not a phonetic entity) which makes a syllable more prominent than the unstressed syllables. It has various phonological *consequences* (e.g. there are phonological rules which are conditioned by stress). Different languages choose different ways of making a syllable more prominent than others, but three basic parameters of the phonetic implementation of stress have been identified in the literature (Lehiste 1970): a syllable may be made more prominent by making it longer (DURATION as a phonetic correlate of stress), by making it louder (AMPLITUDE as a phonetic correlate of stress) or by making its pitch different from that of the other syllables (FUNDAMENTAL FREQUENCY as a phonetic correlate of stress). Languages may choose one or more of these parameters in encoding stress. English, for example, makes use of all the three parameters, unlike Malayalam, which does not employ amplitude to encode stress. The stressed syllable in English is said at a higher pitch than the unstressed syllable, unless phrase-level intonation obscures this correlation. In contrast, the primary stressed syllable in a word in Malayalam is said at a lower pitch (section 4.2.).

This situation makes it obvious that stress is not an entity that figures in phonetic representations, which are to be interpreted by the physiological system without having recourse to language-specific information.

> ... there is no evidence from experimental phonetics to suggest that these contours [i.e. stress contours KPM] are actually present as physical properties of utterance in anything like the detail with which they are perceived. Accordingly, there seems to be no reason to suppose that a well-trained phonetician could detect such contours with any reliability or precision in a language that he does not know ... (SPE, p. 25)

If stress is not part of the phonetic representations, how can the correlates of stress, from which phonological stress can be reconstructed, be represented in the phonetic notation? The conception of phonetic representation in (24) already contains a provision for representing the subtleties of the gradient relationship between stress and duration (the greater the stress, the greater the length). What we need, in addition, are two more tiers in our orchestral representation, namely, those of relative loudness and pitch. In principle, they can both be incorporated into our notation, though the details of separating the filtering-out of the purely physiological aspects of pitch (e.g. the effect of stops on pitch) from the linguistically determined pitch curves requires a great deal of further investigation. It is clear, however, that the proposals in Liberman & Pierrehumbert (1983), which contains the most explicit theory published to date of the phonetic implementation of the phonological representations of pitch patterns, can be incorporated into the notion in (24) without serious difficulty.

In sum, we have discussed three reasons for adopting the nonsegmental conception of phonetic representations as in (24). This conception allows for the statement of (a) the gradient phenomena of segment duration, (b) the misalignment of articulatory gestures, and (c) phonetic correlates of abstract phonological entities like stress.

6.4. LANGUAGE-SPECIFIC IMPLEMENTATIONAL PHENOMENA

The reader might have noticed that the kinds of language-specific processes which were adduced in partial support for the nonsegmental or parametric approach to phonetic representations are radically different in nature from more abstract phonological processes of the sort discussed in chapters 2—5. The examples that we have looked at specify the details of the degree, duration, and coordination of the articulatory gestures needed to implement the information encoded in segmental strings. These mappings ultimately lead to the dissipation of the segments in the overlapping of feature specifications aligned along the parameter of time. Phenomena of this kind have been variously called rules of phonetic implementation, phonetic rules, low level rules, etc., in the literature (Anderson 1975, Prince 1980, Liberman 1983, Pulleyblank 1983, Liberman & Pierrehumbert 1983, etc.). These rules, such as the ones responsible for [ˈ] insertion in *prince* and the lengthening of the vowel in *bid* (as opposed to *bit*), are closely linked to the physiological systems of speech production, and yet they are part of the linguistic systems in that they carry language-specific information without which the physiological systems would not be able to interpret adequately the output of the phonological module. In order to separate the linguistico-physiological processes from the purely physiological implementation of speech, I refer to the latter as PHYSIOLOGICAL IMPLEMENTATION, reserving the term PHONETIC IMPLEMENTATION for linguistically significant phenomena.

Liberman (1983) notes the following as properties of phonetic implementation:

A. Phonological rules are operations on discrete entities. In particular, phonological rules operating on distinctive features are binary in the sense that they change the class membership of segments. In contrast, rules of phonetic implementation may be gradient in nature, i.e. they may be operations on a continuous scale.

B. The number of phonological entities is bounded; the number of phonetic entities is in principle unbounded.

C. The consequences of phonetic rules often involve matters of timing or duration, and coordination.

D. Rules of phonetic implementation cannot have lexically conditioned exceptions.

A and C cover what we have called the degree, timing and coordination of articulatory gestures. Liberman goes on to suggest that the entire class of postlexical operations is implementational. In the light of the preceding discussion (e.g. the gemination in (6)—(9)), however, this attractive hypothesis appears to me to be too strong. I shall therefore continue to take a more conservative position, allowing a nonimplementational postlexical module:

E. The postsyntactic module is the module of phonetic implementation.

As we did in the case of rules applying in the lexical module and the postlexical module, we assume that the phonological rule system contains a single set of rules, and that they are specified for application in the implementational module, in the syntactic module, in any of the strata in the lexical module, or in multiple modules. In other words, we will not distinguish between phonological rules and implementational rules, but between rule applications in the lexico-syntactic modules and applications in the implementational module, allowing the possibility that the same rule may apply in implementational and nonimplementational modules. The properties listed as A—D above should therefore be seen as properties of the *mapping* from the syntactico-phonological representation to the phonetic representation, rather than as properties of a class of implementational *rules*. A gradient operation, for example, must take place in the implementational module, but a binary (= discrete) operation may take place in the implementational module or in any of the nonimplementational modules.

In what follows, I shall discuss in detail the types of subsegmental phenomena which are characteristic of phonetic implementation, and examine the consequences of assumption E above for the theory of Lexical Phonology.

6.5. TYPES OF SUBSEGMENTAL PHENOMENA

6.5.1. *Timing of Articulatory Gestures*

The specification of the duration of vowels when preceded or followed by different kinds of consonants is an example of the phenomena that govern the timing of articulatory gestures (Liberman's C). Yet another example is that of the reduction of vowel duration in unstressed syllables. The contrast between vowel *reduction* and vowel *deletion* also illustrates the gradient nature of implementational phenomena (Liberman's A). Vowel reduction is an instruction to shorten the vowel, which, in the extreme case, results in the total disappearance of the vowel. Thus, the first vowel in words like *potato* and *divinity* are pronounced with considerable shortening (as opposed to, say, the final vowel in *Canada*, which is not followed by a consonant). In casual speech, speakers of English tend to have various degrees of vowel length in unstressed vowels between consonants, including zero length in [pteyDow] and [dviniDi]. While the reduction of vowel length to the point of the disappearance of the vowel is a gradient operation, vowel deletion (or vowel insertion) is a binary (= discrete) operation: the change is from full vowel to zero vowel (or the reverse), with no intermediate stages. Thus, the vowel which surfaces as [ə] of *idea* is deleted, not reduced, in *ideology*, with no trace of the vowel ever appearing on the surface.

At this point the reader might legitimately enquire how much of vowel reduction need be specified in phonetic representations. Would it not be sufficient, for example, to specify the reduced vowel as, say, [+reduced], leaving the details of shortening to physiological implementation? Such an instruction to reduce the duration of the vowel is not sufficient, because the degree of reduction depends on the language in question. First, notice that in both English and Malayalam, a vowel is reduced to extinction only between two consonants, not when it is word-initial or word-final. A brief examination of the segmental conditions under which the extinction takes place in Malayalam shows that these conditions are language-specific. Consider the following examples:

(29)	slow speech	casual speech	gloss
a.	kaṭala	kaṭla	'chick peas'
b.	urumpə	urmpə	'ant'
c.	aṭappə	aṭppə	'lid'
d.	kamalam	kamlam	'Kamalam'
e.	waasana	waasna	'fragrance'
f.	waluppam	walppam	'bigness'
g.	ampalam	amplam	'temple'
h.	poṭiccu	poṭccu	'powdered'
i.	maar̄aṇam	maar̄ṇam	'sorcery'
j.	calanam	calnam	'movement'
k.	malaṭe	malṭe	'hill's'
l.	waakaṭe	waakṭe	'a type of tree's'
m.	aananam	aannam	'face'

As (29) illustrates, the vowel disappears in a wide range of segmental environments. One important condition that prevents the disappearance of a vowel (without preventing shortening as such) is that the preceding and following consonants should not be identical if they are obstruents:

(30)	slow	casual	gloss
a.	kaṭaṭe	*kaṭṭe	'shop's' (cf. (29k—m))
b.	eekakakṣi	*eekkakṣi	'single party'
c.	upapaaṯam	*uppaaṯam	'misfortune'

In contrast, reduction of vowels between identical stops is allowed in Warlpiri:

(31) a.	yardipipurikangu yardippurikangu	'stole the hip joint'
b.	karlijijurutardilykipunam karlijjurutardilykipunam	'He broke the boomerang.'
c.	ngajulparnawartiturni ngajulparnawartturni	'I collected wild potatoes.'

(data from Ken Hale, personal communication)

It is not sufficient, therefore, to specify that the vowel that undergoes extinction is [+reduced]. In Warlpiri, vowels are free to disappear when preceded and followed by identical stops; in Malayalam, they are allowed to shorten considerably but not disappear completely under the same conditions. These restrictions, therefore, must be built into the grammars of the two languages.

6.5.2. *Coordination of Articulatory Gestures*

We have already seen some examples of language-specific processes which are consequences of the coordination between different articulatory gestures (the [^{t}] insertion in *prince* in English, the [t] insertion in *swapnam* in Malayalam). These examples show that a number of regularities of speech, which have so far been dismissed as the domain of physiological implementation and left for phonetics to deal with, are in fact language-specific, and must, therefore, be handled within an adequate theory of phonology. By enriching the notion of phonetic representation as outlined in section 6.3., and incorporating the module of phonetic implementation to yield such a representation, we now have a mechanism that will allow these phenomena to be brought within the domain of phonology.

As pointed out earlier, implementational phenomena are firmly rooted in physiological systems. Many of them might be seen as grammaticalizations of physiological phenomena. A telling case of the way such grammaticalization works is seen in the absorption of vowel quality by the preceding consonants in Malayalam.

It is well-known that the consonants in human languages anticipate the tongue and lip positions of the following vowels. Thus, the **t** in *tea* is said with spread lips, while that in *two* is said with rounded lips, because they anticipate the lip positions of **iy** and **uw** respectively. Similarly, while making the bilabial closure for **p** in *pea*, the front of the tongue is already raised in anticipation of the following **iy**, in contrast to the raising of the back of the tongue in *pooh*. Now, the retroflex **ṭ** in Malayalam in words like **aṭuppə** 'oven' anticipates the back raising of **u**, and is therefore said with a different tongue posture from that of the **ṭ** in **aṭappə** 'lid', which is said with the tongue body in a neutral position. As we noted earlier, the medial (unstressed) vowels in this environment are subject to vowel reduction, and as an extreme form of reduction, the vowel becomes extinct. What is remarkable is that in spite of the complete disappearance of the vowel, speakers of Malayalam produce **aṭappə** and **aṭuppə** differently, and listeners can tell them apart without any effort. This applies to the reduced versions of **aṭaccu** 'closed' and **aṭiccu** 'beat' as well. The phonetic distinction between the words in each pair lies in the tongue body articulation associated with the retroflexes in the two words. In spite of the absence of overt vowel articulation, the retroflex in **aṭuppə** is still said with the back of the tongue raised, which distinguishes it from **aṭappə**, and that in **aṭiccu** is said with the front of the tongue raised.

(32)	aṭappə	aṭuppə	
	aḌappə	aḌuppə	intervocalic lenition
	———	aḌ̂uppə	raising of the back
	aḌppə	aḌ̂ppə	vowel reduction

It is obvious that the anticipatory raising of the tongue should be recorded in the phonetic transcriptions of Malayalam. Yet, coarticulatory phenomena of this kind are typically those which have been banished from the domain of phonology in most theories. One of the advantages of the formalism in (24) is that it allows us to deal with derivations like (32).

6.5.3. *Degree of Articulatory Gestures*

The binary value + in [+high] specifies that the body of the tongue should be raised. In order to implement this gesture correctly, the articulatory organs need to know *how high* the tongue should be raised. This specification is what is taken care of by the use of scalar values of distinctive features in SPE, designed to provide information about the degree of articulatory gesture. Within our conception of phonology, the question "How high?" is answered only in the implementational module.

An example of degree specification of an articulatory gesture is that of the reduction of vowel quality (as opposed to vowel duration) in reduced vowels. Vowel quality reduction may be taken as an instruction to get *closer* to the central position or a position in which the tongue is maximally relaxed during the production of speech. Sidestepping the question of the exact phonetic details involved, we note the following: just as the extreme form of phonetic length reduction produces results identical to that of phonological deletion, the extreme form of phonetic quality reduction produces results which are identical to that of phonological segment neutralization. Thus, many speakers of English pronounce the reduced (boldface) vowels in the following words with different vowel qualities in careful speech, but not in colloquial/fast speech:

(33) a.	sol**i**d	vs.	meth**o**d
b.	goodn**e**ss	vs.	min**u**s
c.	**e**xact	vs.	**a**bout
d.	Ros**y**'s	vs.	Ros**a**'s, ros**e**s[8]

In Malayalam, vowel reduction centralizes **i**, and unrounds and centralizes **u**. The distinction between the two, however, is not neutralized, because the reduced form of **u** (which I shall symbolize as [ə̂˃], which is a raised and retracted [ə]) is further back than the reduced form of **i** (which I shall symbolize as [ə̂˂], which is a raised and fronted [ə]). As a result, the distinction between the **i** in **kurippə** 'note' and the **u** in **kuruppə** 'a surname' is maintained in normal speech as [kurə̂˂ppə] and [kurə̂˃ppə]. The vowels [ə̂˂] and [ə̂˃] are phonetically distinct from the reduced version of **a** (which I shall represent as [ə̌], a lowered [ə]) in words like **karappə** 'blackness'). Clearly, distinctions of this kind are not expressible in terms of binary oppositions of features.

The specification of degree is needed, not only in reduced vowels, but also in full vowels. The phonetic realization of the vowel quality of full vowels differs from language to language, and dialect to dialect. The vowel nucleus in *mean*, for example, has a higher tongue position in most dialects of English than the vowel nucleus **miin** 'fish' in Malayalam. The vowel in *met* varies a great deal depending on the dialect in question, ranging from a high mid to a low mid in different dialects of English. (See Wells (1982), who calls them "realizational" differences.) Differences of this kind are not statable in terms of binary oppositions. They are, nevertheless, part of the phonological system of a language or dialect.

It may be pointed out that the specification of the exact vowel quality cannot be part of the articulatory setting referred to in section 6.2., since it is not the case that a vowel has exactly the same phonetic quality in every environment. Thus, the vowel **o** in **koṭṭi** 'drummed' in Malayalam is slightly higher than the **o** in **koṭṭa** 'basket': vowels are raised when followed in the next syllable by a high vowel. After specifying the details of the vowel quality of **o** (and other vowels), therefore, the phonological rules of the language must be allowed to make finer adjustments on this quality.

We also note that the degree of gestures does not remain constant for the given length of a segment. Thus, we need language-specific instructions to spell out the changes in vowel quality. Some of these are determined by the environment, as in the slight diphthongisation of short vowels before voiced sounds in most dialects of American English (e.g. the vowel in *bed* has a [ᵊ]-like offglide, in contrast to the absence of the offglide in *bet*). Others are context-free realisational features, such as the diphthongisation of the short low front vowel **æ** in New York English as [eə] and [æə] (Wells 1982, p. 503). In Boston English, **æ** begins with an onglide that approaches [I]. Though **æ** is a *phonetic* diphthong in these dialects, it is not a phonological diphthong: the rules responsible for deriving [ey] from /æ æ/ are phonological rules, those responsible for deriving [æ ə] from **æ** are phonetic implementation rules.[9]

6.5.4. *Enhancement as Phonetic Implementation*

In addition to the three types of processes which are unique to the implementational module, namely, the specification of the timing, coordination and degree of articulatory gestures, there seems to be a case for a process which introduces "enhancement features" (Stevens & Keyser 1985) in this module. Stevens & Keyser use the term 'enhancement' to refer to redundant phonetic properties which are not distinctive by themselves, but merely serve to intensify a phonological contrast encoded by other features. In a language like Malayalam that has the vowels **i**, **e**, **a**, **o** and **u** at the lexical level, the feature [round] is an enhancement feature, since the features [high], [low] and [back] are sufficient to encode the lexical contrasts. All that rounding is doing is enhancing the backness of the nonlow vowels. Similarly, the feature of aspiration, not needed to encode phonological contrasts in English, enhances the voicelessness of plosives in some environments. I incorporate this notion into Lexical Phonology as follows:

(34) An enhancement feature is one that is not required to make phonological distinctions at the level of lexical representation.

Note that the level that is crucial for this definition is the lexical level, not the underlying level. The feature [tense], for example, has been argued to be redundant in underlying representations in English (Halle & Mohanan 1985), but since it is needed to encode lexical contrasts, [tense] is not an enhancement feature but a distinctive feature for English vowels.

Having characterised the notion of enhancement, we place an additional restriction on the phonological systems of natural languages by making the following assumption:

(35) Enhancement features are made available only at the implementational module.

Given (35), it would follow that the rule that aspirates voiceless stops in English can apply only in the module of phonetic implementation. It would also follow from (35) that articulatory gestures which are not used to make lexical distinctions in any natural language are universally constrained to be introduced at the implementational module. An example of this kind of enhancement feature which is nondistinctive in all human languages is that of lip protrusion (the increasing of the length of the oral cavity by extending the lips in the horizontal dimension) as opposed to lip rounding (the narrowing of the lip orifice in the vertical dimension). In English, lip protrusion accompanies lip rounding as an enhancement, but not all languages implement lip rounding with this accompaniment (Ladefoged 1971). The instrumental studies made by Nihalani (1985a), for example, show that the back nonlow vowels in Sindhi are said with lip rounding, but no lip protrusion. Thus, the specification of lip protrusion is part of the grammar of natural languages, but given (35), it can de dealt with only in the implementational module.

To summarise, then: in addition to the three types of processes which Liberman identifies as being restricted to the implementational module, namely, the timing, coordination and degree of articulatory gestures, we identify one more, namely, the enhancement of articulatory gestures.

6.6. UNDERLYING AND LEXICAL ALPHABETS

Let us assume that the mappings in the lexical and syntactic modules are operations on a finite set of segments characterisable in terms of binary distinctive features drawn from a universal alphabet. In contrast, mappings in the implementational module may dissolve phonological segments. Following Mohanan & Mohanan (1984), I shall use the term UNDERLYING ALPHABET to refer to the inventory of segments used at the underlying level, and LEXICAL ALPHABET to refer to the inventory of segments at the lexical level.

Given the notion of underlying and lexical alphabets, there exist four possibilities of relating the two:

(36) a. The underlying and lexical alphabets are identical.

b. The underlying alphabet is a subset of the lexical alphabet.

c. The lexical alphabet is a subset of the underlying alphabet.

d. The underlying and lexical alphabets are identical in the unmarked case.

Condition (36b) prohibits absolute neutralization rules (e.g. the yers in Russian, the Maltese ʕ discussed in Brame (1972), etc.). If we accept the need to postulate underlying segments which do not surface, we must reject this condition. A comparison of the underlying and lexical nasals in Malayalam shows that (36c) cannot be maintained either:

(37)	underlying:	m n ṇ
	lexical	m n ṇ n̲ n̄ ŋ′ ŋ

(see section 3.2.1.)

In English, the underlying alphabet does not require the feature [tense], but the lexical alphabet does, in order to distinguish between the lexical segments [I] and [i] (section 2.3.2.).

Kiparsky (1983) has proposed that lexical operations are "structure preserving", while postlexical operations need not preserve structure. By structure preservation, Kiparsky also includes a ban on the introduction of segments which are not part of the underlying inventory (36c): the lexical alphabet must obey the segment structure constraints of the underlying alphabet. The Malayalam and English facts cannot be explained if the constraint of structure preservation is interpreted in this fashion (I refer the reader to Mohanan & Mohanan (1984) for details of (37)).

Similar options must be considered for the relation between the segments at the level of lexical representation and syntactico-phonological representation. I assume, in the absence of counterevidence, that the alphabet used for syntactico-phonological representations is the lexical alphabet. If so, it follows that the mapping from the lexical to the syntactico-phonological level cannot introduce novel segments.

As stated earlier, underlying and lexical alphabets constitute a small number of entities characterisable in terms of binary features drawn from a universal inventory. The condition that operations in the lexical and syntactic modules are operations on the union of the underlying and lexical alphabets yields directly all the properties that we listed earlier as properties of the implementational module. Given that the alphabets are characterised in terms of binary entities, it follows that the lexical and syntactic modules do not allow gradient operations; given that the alphabets constitute a small set, it follows that only the implementational module allows operations on unbounded entities. Operations in the lexical and syntactic modules may be SEGMENTAL (altering the class membership of segments, deleting segments, or inserting segments), or STRUCTURAL (building or changing metrical or autosegmental structures on segments). Given the condition that operations in these modules are bounded by the lexical alphabet, it follows that they cannot allow SUBSEGMENTAL operations (those that specify the timing, coordination, degree, and enhancement of the articulatory gestures used in producing the segments).

Our condition on alphabets predicts that operations which are unstatable in terms of underlying and lexical alphabets (i) cannot have access to morphological or syntactic information, and (ii) cannot precede those operations which are governed by morphological or syntactic information. Thus, the specification of lip protrusion (as opposed to the specification of lip rounding: see section 6.5.4.) is not available in the universal inventory of binary features, and hence cannot be part of the alphabet of a natural language. The specification of lip protrusion must therefore take place in the implementational module in all natural languages. Similarly, the details of vowel duration which operate on a continuous parameter (see note 7) are not statable in terms of a set of binary features, and must take place in the implementational module. The same condition applies to the rule of the aspiration of voiceless stops in English (14), which takes the gradient of stress in its structural description, and yields the gradient of aspiration as its output. Given our assumptions, these phenomena cannot be governed by morphological or syntactic structure or features (including exceptionality features), nor can they precede operations which are governed by these features. To the best of my knowledge, these predictions are consistent with the known facts about natural languages.

6.7. PHONOLOGICAL STRUCTURE AND PHONETIC IMPLEMENTATION

If we adopted (36c), the conditions that defined the structure of segments (segment structure conditions in SPE) would have to be identical at the underlying and lexical representations. As pointed out above, we reject this assumption. Instead, what we must say is that conditions which define the structure of segments can be imposed only within the lexical and syntactic modules, since segments exist only in these modules. Operations in the implementational module dissolve phonological segments. Since we have adopted the idea that the internal structure of segments is hierarchical (Mohanan 1983, Clements 1985), the implementational module must dissolve the trees at the subsegmental level. The suprasegmental trees (syllable trees, stress trees) are built on segments, and if segments are dissolved, it must be the case that these structures are dissolved as well. Thus phonological structures at all levels exist only in the lexical and syntactic modules, and are dissolved in the implementational module.

McCarthy (1979) proposes the following condition on phonological rules: "A phonological rule may apply to a form only if its output can be properly syllabified. If the output cannot be syllabified, the rule is blocked from applying".

In the light of what we said about dissolving structures, McCarthy's condition is inapplicable in the implementational module. It would automatically follow that the operation that reduces the duration of vowels (as opposed to vowel deletion), leading to the disappearance of the first vowels in *divinity* and *potato*, is not subject to the condition that it preserve syllable structure. This conclusion is indeed correct, as the syllable structure of English does not allow the onsets **dv** and **pt**. Similarly, Malayalam does not allow syllabic nasals, but the implementational rule of vowel reduction results in what looks like a syllabic nasal in words like **r̄aamad̲aasan** [r̄aamd̲aasn] 'a name'.

Another example of the failure of phonetic implementation to observe structural conditions is provided by words like *barren* and *barrel*, listed with an italicized [ə] (indicating optionality) after [r] in Jones (1977). The forms without [ə] violate the structural constraint that **r** can occur only before vowels in the dialect of English given in the dictionary. The mystery is solved when we realize that italicization of [ə] in Jones is a way of marking the vowel which can undergo vowel reduction to extinction. Since vowel reduction is a process that takes place in the implementational module, it is free from the structural constraint on **r**.

A similar situation is found in the interaction between vowel reduction and **l** Velarisation (section 2.4.3.). When the post-lateral vowel in words like *California* and *Palestinian* is reduced to extinction, yielding consonant sequences like [lf] ([kælfoniə]) and [lst] ([pælstiniən]), the lateral is still a clear [l], not [ɫ]. This means that the rule of **l** Velarisation precedes vowel reduction. Given our assumption that syllable structures do not exist when vowel reduction takes place, the ordering could not have been the other way round (vowel reduction preceding **l** Velarisation), since **l** Velarisation crucially refers to syllable structure. We predict, therefore, that no dialect of English can have alternations like [kælifoniə] (normal speech)/ [kæɫfoniə] (casual speech).

Our last example comes from Malayalam, in which **r̄** can occur only before a V element: [awar] 'they', [awar̄uṭe] 'their', [awwarkkə] 'to them'. The results of vowel reduction, however, obscure this constraint, as shown by [kar̄impə]/[kar̄mpə] 'sugar cane', [maar̄aṇam]/[maar̄ṇam] 'sorcery', etc. Any sequence of **r̄C** which is not paired with a **r̄VC** in slow speech is illformed in the language, showing at once the need for the structural constraint, and the inapplicability of the constraint in the implementational module.

6.8. PHONETIC IMPLEMENTATION AND CLASSICAL PHONEMICS

The striking similarity between the theory of phonetic implementation and the theory of classical phonemics must be obvious to the reader by now. I believe that the actual practice of classical phonemicists was governed by significant insights into the nature of speech, many of which were unfortunately not pursued with the advent of SPE. As pointed out in chapter I, what was troublesome about classical phonemics was the way in which intuitions about the phonemic level were built into a formal theory: the principles of the theory were incorrect, and furthermore, they did not match either the intuitions or the practice of classical phonemicists. This led to an inconsistency between theory and practice that most of the practitioners were not aware of. The theory of implementational phenomena brings back to generative phonology the valid intuitions behind the practice of classical phonemics. The lexical alphabet, for example, captures the intuition behind classical phonemes, and the distinction between implementational phenomena and nonimplementational phenomena captures the intuition behind allophonic and phonemic changes. It is necessary, therefore, to draw attention to the technical and conceptual differences between classical phonemics and Lexical Phonology in capturing these intuitions.

6.8.1. *Conditions Relating the Phonemic and Phonetic Levels*

Chomsky (1964) identifies linearity, invariance, biuniqueness and local determinacy as the assumptions which characterize the classical phonemic theory. The major point of similarity between the classical phonemic level and the level that we have identified as the syntactico-phonological level is that the mapping from either of these levels to the phonetic level has no access to nonphonological information. Thus, both mappings obey the condition that Chomsky (1964) referred to as LOCAL DETERMINACY. However, the remaining conditions, namely, INVARIANCE, BIUNIQUENESS, and LINEARITY, do not apply to the mappings in the implementational module.

It must be pointed out that classical phonemics took for granted the segmental nature of phonetic representations (12), while the present version of Lexical Phonology does not recognise segments at this level. The discussion of linearity, biuniqueness, etc., is meaningless within a theory in which phonetic representations are not segmental. For the sake of exposition, however, I shall ignore this discrepancy.

The condition of linearity entails that the phonetic and phonemic representations of a form must contain identical numbers of segments. This condition does not apply to the mapping of phonetic implementation, since we allow the syntactico-phonological representation of *prince* to be **prins**, and its phonetic representation to be [prinᵗs].

Biuniqueness is a condition that forces a one-to-one relation between phonetic and phonemic mappings, and invariance is a prohibition against absolute neutralization. The way Lexical Phonology deals with the facts of homorganic nasal assimilation in English shows that the theory does not observe either of these conditions. First, we observe that there are two types of homorganic nasal assimilation in English. One of them applies in the lexicon, changing the **n** to ŋ in *long* [loŋ], *lo[ŋ]ger* and *co[ŋ]gress* but not in *co[n]gressional* (Halle & Mohanan 1985). The other rule applies post-lexically, changing the **n** to ŋ in fast/casual speech in *congressional* (leaving the **n** unaffected in slow/careful speech), as well as across words in *ten pounds* (n → m), *ten cats* (n → ŋ) and *ten things* (n → n̲). Within the same syllable, this rule applies obligatorily, not only in casual speech, but also in careful speech: *te[n̲]th*, not **te[n]th*. Given our assumption that fast speech is handled by phonetic implementation, we must recognize that this assimilation (i.e. in *congressional, ten things*, etc.) applies in the implementational module. We now have the following mappings:

(38) a.	ten things	b. ten pounds	
	n θ	n p	
	n̲ θ	m p	assimilation

The [mp] sequence derived from **np** is phonetically identical to the [mp] sequence derived from **mp**. The derivation in (38b) leads to absolute neutralization, violating invariance as well as biuniqueness.

6.8.2. *The Nature of the Mapping*

With the wisdom of hindsight, it is easy to see why classical phonemic theory needed principles like biuniqueness and invariance, and how the classical phonemicist intuitions didn't lead to a theory similar to the implementational module in Lexical Phonology. We can now see reasons for this both in the conception of phonology and in the conception of linguistics as a science.

Let us begin with reasons internal to phonology. Basic to most of the principles of phonemicisation developed in classical phonemics was the idea that the mapping between phonemic and phonetic representations was statable as a mapping between the *alphabets* of the two levels of representation, rather than between the *representations* themselves. Rather than talking about the relation between a phonemic string and the corresponding phonetic string, classical phonemics talked about the relation between a phoneme and its allophone. This led to the formulations of contrastive distribution, complementary distribution and free variation in terms of substitutability of segments rather than substitutability of strings. Two phonetic segments were in contrastive distribution if they were mutually substitutable, and the substitution created a change of meaning. They were in free variation if they were substitutable but the substitution did not produce a change of meaning, and in complementary distribution if they were not mutually substitutable (e.g. Pike (1947:62)). Given this approach, one could say that [i] and [e] were in contrastive distribution in English, [t] and [t˺] in free variation, and [t] and [th] in complementary distribution, but the theory could not express or make use of generalisations like postpausal [e] and [ʔe] being noncontrastive, and the [nᵗs] in *prince* filling the distributional gap of [ns]. As a result, instead of pairing [ʔeg] with [eg] and [prinᵗs] with [prins], classical phonemics paired [ʔeg] with [beg] and [prinᵗs] with [prints]. Given these minimal pairs, the theory forces a conclusion which is intuitively unacceptable, namely, that [^{t}] and [t], [ʔ] and [b] are distinct phonemes in English.

The conception of a mapping in terms of alphabets rather than in terms of representations led to the distinction between "phonemic alternations", which involved the alternation between two phonemes (two entities in the phonemic alphabet), and "allophonic alternations", which involved the alternation between two allophones of the same phoneme. Thus, the [n]/[ŋ] alternation in *convert/congress* was a phonemic alternation, while the [n]/[n̪] alternation in *ten/tenth* was an allophonic alternation (Gimson (1980:293—4)). Lexical Phonology captures the essential intuition behind this distinction by claiming that the former takes place in the lexical or syntactic module (lexical module in the example given above), while the latter takes place in the implementational module. Where Lexical Phonology differs from classical phonemics is in dealing with the [n]/[ŋ] alternation in *ten/ten cooks* in the implementational module. Within the classical phonemic theory, this alternation is phonemic, while the [n]/[n̪] alternation in *ten/ten things* is allophonic.

Another assumption that determined the shape of the classical phonemic theory was that the mapping between phonemic and phonetic representations could be stated as a mapping that directly *linked* the two representations with each other, rather than through a series of representations, with the output of one function serving as the input to another. That is, the rules that characterised the phonemic-phonetic mapping applied simultaneously to the phonemic representation, and there was only one rule scanning. As a result, a sequential derivation like /bediŋ/ and /betiŋ/ first changing to **be.diŋ** and **betiŋ**, and then changing to [be.Diŋ] and [beDiŋ] was impossible in the theory. The idea of sequential application of rules (with or without extrinsic rule ordering) was alien to this conception of phonology. Coupled with this conception of derivations was the idea that the distributional statements on allophonic variation were made on the phonetic inventory, not on the phonemic inventory. Thus, one could say that phonetic segment [*A*] occurs before phonetic segment [*B*], not that phonetic segment [*A*] occurs before phonemic segment /*C*/. This assumption, subsequently formulated as the 'True Generalisation' condition in Natural Generative Phonology, makes it impossible to state the generalisation that [e.] occurs before /d/, not /t/. The result is the undesirable conclusion that [e] and [e.] belong to distinct phonemes in American English.

Similar problems appear in the case of preconsonantal clear [l] in examples like *California* and *Palestinian*, discussed in section 6.7. In classical phonemics **l** Velarization is an allophonic rule, and in Natural Generative Phonology, it is a P-rule: neither can explain why the clear [l] is allowed to occur in these examples, since sequential rule application that creates surface opacity is disallowed in these theories. One may add to this list examples like *barren* [bærn̩] and *barrel* [bærl̩] in nonrhotic accents of English, in which vowel reduction leading to vowel extinction obscures the generalizations that [r] never occurs preconsonantally and [n] and [l] are never syllabic after [r].

A more interesting example of the failure of the theory that disallows sequential operations is seen in the interaction between nasal assimilation and nasal spread in Malayalam and English. Recall that we analysed the [n̲n̲] sequences in Malayalam as underlying /Nd̲/ sequences. In Malayalam, the processes that yield the surface forms apply in the lexicon (section 3.2.1). The same processes apply in English, converting /nð/ sequences into surface [n̲n̲] (n → n̲/___ð, ð → n̲/[+nasal]___), yielding forms like *and then* [ən̲n̲en], *when that* [wen̲n̲æt], *on the* [ɔn̲n̲ə] in casual speech for most speakers of English, forming minimal pairs like *and though* [ən̲n̲ow] and *and no* [ənnow]. Unlike what happens in Malayalam, nasal assimilation, and the subsequent nasalization of the consonant that triggers it through nasal spread, take place in the implementational module in English. This distinction corresponds to the classical phonemicist's intuition that [n] and [n̲] belong to different phonemes in Malayalam, but to the same phoneme in English. This intuition (which we accept) does not follow from the classical phonemic theory (which we reject): given the minimal pairs [ən̲n̲ow] and [ənnow], the principles of phonemicisation would assign [n̲] and [n] to different phonemes in English as well. Needless to say, this conclusion is unacceptable to all phonemicists.

NOTES

1 A prosodic phrase, also called "tone group" or "intonation phrase" (Halliday 1973, Kingdon 1958), is a phonological string carrying a single "tune", "pattern of intonation" or what Liberman (1975) calls "intonation word". It carries one and only one phrasal nuclear stress (also called "tonic" or "nucleus of intonation").

2 In the model of Lexical Phonology proposed in Mohanan (1982), the syntactic module was not distinguished from the postsyntactic module. Because of the difficulties that the model presented in accounting for the way pauses affected the *a/an* alternation, I had to assume that the process took place in the lexicon, with conditions on the phonological environment that followed each form. This contrivance is not needed in the model given in (5).

3 SPE quotes Marouzeau (1943, p. 38), who defines "articulatory base" as "the system characteristic of articulatory movements of a given language that confer upon it its general phonetic aspect".

4 This notation may or may not turn out to be appropriate for the statement of rules of this kind. I am using the SPE formalism to bring out the gradient nature of the rule more clearly, without any implication that this is the formalism that I want to advocate.

5 In a sense, it is to be expected that Hindi, Telugu, Sindhi and Malayalam differ from English in this fashion. In most Indian languages, there is a lexical contrast between aspirated and unaspirated stops, in addition to the contrast between voiced and voiceless stops. As a result, [pi . .] and [phi . .] would be perceived as different words, both distinct from [bi . .]. It is quite possible that the devoiced [b̥i . .] is avoided in these languages since

it might be perceived as the realization of lexical **pi . .** In English, on the other hand, there is no danger of preceiving [b̥i .] as lexical **pi . .** , since the latter is realized as [phi . .], not [pi . . .], and hence the structure of the language permits the luxury of delayed voice onset for articulatory ease.

This functional explanation predicts that delayed voice onset would be disfavoured in English in environments in which aspiration does not appear. Thus, one would expect the delay in the **b** of *baroque* to be much less than the **b** in *bed*, since it is in an unstressed position in which aspiration does not occur. I do not know of any data that bears upon this prediction.

[6] The time axis is not to be interpreted as a specification of physical time (in terms of milliseconds) but as a specification of relative time. Thus, we interpret the continuant stretch for *Ben* to be longer than that for *bet* (23), whatever the actual physical duration (which is also dependent on the speaker's rate of speech).

See also Abercrombie (1965, pp. 120—124) who advocates the "dynamic approach" or the "parametric approach" to phonetic representation, which is essentially a nonsegmental view of phonetic representation, similar to (24). The parametric approach contrasts with the traditional 'static approach', which views phonetic representations as in (21)—(23), incorporating (12). It must be mentioned that Abercrombie advocates the right approach for the wrong reason, as what he cites as evidence for the parametric approach is the raw data from spectrograms and kymograms. Instrumental studies tell us that the speech signal is nonsegmental, but not necessarily that the phonetic representations should also be nonsegmental.

[7] Vowels are shortest before [p] and [k]. Assuming an average duration of 160 msec., House and Fairbanks compute vowel duration as follows: add 20 msec. if the following consonant is coronal, 35 msec more if the consonant is continuant, and 80 more if it is voiced. If the following consonant is a nasal, subtract 10 msec. Thus, the length of a vowel before [z] is $160 + 20 + 35 + 80 = 295$ msec.

[8] Some of these differences are represented in Jones (1977) in terms of the [ə]/[I] distinction, e.g. *goodness* [gudnIs], *about* [əbaut], *Rosa's* [rowzəz], *roses* [rowzIz].

[9] The processes that change **ii** (a monophthong [i:] in British English) to a diphthong [iy] in American English and [əy] in Australian English are also rules of phonetic implementation.