Phonetics and Phonology*

Abigail C. Cohn

In this chapter, I investigate a number of issues about phonology, phonetics, and their relationship. First, I discuss the nature of both phonology and phonetics, by considering the nature of rules, representation, and underspecification in each domain. I then focus on the nature of the relationship between phonetics and phonology and the question of distinguishing between the two. It is concluded that phonology and phonetics are indeed formally distinct in that what they manipulate is different: in the phonology, abstract qualitative representations are manipulated and in the phonetics such representations are realized quantitatively in both time and space. It is argued that the mapping from phonology to phonetics can be insightfully represented by a target-interpolation model. Such a model is taken as a starting point to account for phonetic implementation of the feature Nasal, the goal of the subsequent chapters of the study.

Phonetics and phonology are usually defined as distinct areas of study. Phonetics is the study of the physical properties of sounds used in human speech: their production, their acoustics, and their perception. Phonology is the study of how speech sounds pattern together. Yet there is also an implicit derivational relationship between the two: a phonological representation indicates the abstract, linguistic characteristics of sounds; the phonetic representation is the physical output or realization of the phonological representation, what the speaker actually produces or the hearer perceives. This relationship can be investigated from the point of view of production or perception.

There is evidence from psycholinguistics that abstract phonological representations exist for speakers and hearers and many researchers conclude that this is part of what speakers/hearers know about their language. Following this view, the nature of the relationship between abstract linguistic representations and the actual physical output is important for our understanding of language as a human cognitive process, since it offers us insight into one aspect of the linguistic behavior of both speakers and hearers in using human language. Thus not only the abstract patterns of sounds, but also the mapping processes between distinctive contrasts and physical events are central issues in our understanding of phonology, as one facet of linguistic knowledge.

* This work is the introduction of a book in progress, entitled From Phonology to Phonetics. It takes chapter 1 of my dissertation, Cohn (1990), as a starting point. Some of these issues are also taken up in Cohn (1993b) §2. Thanks to Ioana Chitoran, Allard Jongmin, Pat Keating, and Drag a Zec for comments. Thanks also to Bruce Hayes, Pat Keating, Peter Ladefoged, Donka Minkova, and Carlos Otero for general discussion of the issues presented here.
It is often assumed that mapping from an abstract representation to the physical output is a fairly mechanical process; the specifics of this process are usually only hinted at. However, recent work on this topic shows that the mapping from the phonology to the phonetics is in no sense trivial. (See notably Pierrehumbert 1980, Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, and Keating 1984; 1985a; and 1988a; see also Browman and Goldstein 1992 and work cited therein for a somewhat different view.) In this study, I investigate the mapping from an abstract phonological representation, viewed in terms of discrete and timeless segments, to a quantitative, phonetic representation, realized in time and space, through a detailed study of the feature Nasal, including its phonological behavior and its phonetic implementation. In this first chapter, I discuss the nature of phonology and phonetics and their relationship.

The widely held view of phonology and phonetics as two distinct areas of investigation is in fitting with a modular view of the linguistic grammar, a tenet of modern linguistic theory (see Chomsky 1988). Phonology and phonetics, while both concerned with sound structure, have distinct goals and are governed by different principles. Within a generative perspective, phonology provides a set of representations and rules for any given language. The set of representations allows all of the meaningful contrasts in a particular language to be made; while the rules flesh out these representations to provide a surface representation interpretable by the phonetics. The phonology serves the dual role of representing the linguistically relevant aspects of sound structure and providing the input to the phonetic realization.

Within the model of generative phonology, the phonology must encode the non-predictable aspects of pronunciation of all of the lexical items within any human language. It is generally agreed that only non-predictable information is represented in the mental lexicon and that predictable information is introduced by rule. A range of evidence strongly supports the view that phonological representations involve discrete units with a segmental component. (See Halle and Stevens 1990 for discussion of this point.) A form such as will /wɪl/ is assumed to consist of three discrete sounds or phonemes: /w/, /ɪ/, /l/. This view is based on distributional evidence, i.e. the fact that will /wɪl/ is distinct from bill /bɪl/, from well /wel/, and from win /wɪn/, and psycholinguistic evidence, e.g. slips of the tongue.

Within most current phonological theories, phonological representations consist of bundles of distinctive features. Although a priori, such features could be arbitrary, strong evidence exists for the view that features are characterized in phonetic terms. This set also serves to define systematic properties of linguistic sound systems, such as phonotactic
patterns and systematic alternations, which are affected by both articulatory and perceptual considerations. It is an important result that the same set of features serves to characterize these different sound properties and that this set of features is best defined in phonetic (articulatory and/or acoustic) terms. Although some debate exists, the set of features generally assumed for phonological representations is roughly the set proposed by Chomsky and Halle (1968, hereafter SPE), with some slight modifications. (See Keating 1988c for an overview of post-SPE features.)

A lexical item such as *will* /wɪl/ is represented abstractly as three distinct feature bundles, as schematized in Figure 1A. (The specific features used in this example are not central to the point being made.) The phonetic realization of such a form is not so clearly segmentable into three units, as shown by the spectrogram presented in Figure 1B.

A.  
\[
\begin{array}{c}
\text{-cons} \\
\text{+son} \\
\text{+back} \\
\text{+round} \\
\text{...} \\
\end{array} & \begin{array}{c}
\text{-cons} \\
\text{+son} \\
\text{-back} \\
\text{+high} \\
\text{-tense} \\
\text{...} \\
\end{array} & \begin{array}{c}
\text{+cons} \\
\text{+son} \\
\text{-lat} \\
\text{-tense} \\
\text{...} \\
\end{array}
\]

B.  
\[
\begin{array}{c}
\text{4000} \\
\text{3000} \\
\text{2000} \\
\text{1000} \\
\text{Hz} \\
\end{array}
\]

Figure 1. A. A possible phonological representation of American English *will* /wɪl/; B. A wideband spectrogram of the same form, spoken by a female speaker of American English. F1 and F2 have been highlighted with a black line.

1 While Chomsky and Halle (1968) focus on the articulatory aspects of their feature system, they explicitly acknowledge the importance of acoustic and perceptual correlates: "We shall speak of the acoustical and perceptual correlates of a feature only occasionally, not because we regard these aspects as either less interesting or less important, but rather because such discussions would make this section... much too long." (p. 299) It is important to bear in mind the relevance of both articulation and perception for phonology.
This spectrogram is an acoustic representation, one phonetic representation of the form *will*. In the spectrogram, we observe that the formants (the characteristic overtones that acoustically identify the quality of vowels and sonorants) do not show clear breaking points between the /w/ and /u/ or the /u/ and /v/, rather they are continuously varying. It is difficult to identify a precise point at which one segment ends and another begins. Hockett's (1955) analogy of an assembly line of mashed Easter eggs graphically captures this point.

"Imagine a row of Easter eggs carried along a moving belt; the eggs are of various sizes, and variously colored, but not boiled. At a certain point, the belt carries the row of eggs between the two rollers of a wringer, which quite effectively smash them and rub them more or less into each other." (p. 210)

He then goes on to discuss an inspector who he likens to the hearer, whose job is not to put the eggs back together, but to identify what they were.

In a sense, this is the crux of the problem in relating phonological representations to phonetic ones. How can we characterize the relationship between the abstract discrete representation in Figure 1A, and the continuous one in Figure 1B? Note that we can ask this same question, going either from an abstract linguistic representation to a physical one or from a physical to an abstract linguistic one. I frame the question here in terms of the mapping from an abstract representation to a physical one.

In the remainder of this chapter, I turn first to a discussion of the nature of phonological rules and representations (§1), and phonetic ones (§2) before delving into the issue of distinguishing between phonetics and phonology (§3). I conclude with an outline of the remaining chapters (§4).

1. **Phonological background and assumptions**

Although strong evidence exists that segments are a psychologically relevant level of representation, it is also clear that segments consist of smaller units or features. Distinctive feature theory as developed by the Prague School of Linguistics is a central tenet of modern phonology (see notably Trubetzkoj 1939 and Jakobson, Fant, and Halle 1963 and also Anderson 1985 for discussion of these historical developments). In more recent work, it has also been argued that particular features may function independently, as separate tiers (Goldsmith 1976, Clements 1976) and that there is some kind of logical hierarchical structure among features internal to the segment, termed *feature geometry* (Clements 1985, Sagey 1986, McCarthy 1988). A well-formed underlying phonological representation involves feature specifications assumed to bear a universally determined structural relationship to one another.
Rules of the phonology consist of elemental processes of feature association: spreading (assimilation), schematized in (1a); delinking (deletion, or loss), shown in (1b), and deletion followed by feature fill-in (dissimilation), as in (1c).

\[(1) \quad \begin{array}{cccc}
   & \cdot & \cdot & \cdot \\
   \_ & \_ & \_ & \_ \\
   +F & +F & +F & +F \\
\end{array} \]

Within this approach, explicit predictions are made about the simplicity of rules. Those involving only these elemental processes are the simplest, and are thus predicted to be very common cross-linguistically. Further, only individual features or those dominated by a single abstract node in the feature geometry are predicted to pattern together in these processes. The feature geometry is one part of a well-formed phonological structure, which includes both a segmental representation and a prosodic one (that is, syllabic structure, metrical structure, and so forth).

1.1 Phonological rules

The recent phonological literature includes much discussion of patterns of phonological rule behavior and how they interact. As summarized in (2), the following clustering of properties has been observed.

\[(2) \quad \begin{array}{ll}
   \text{sensitive to morphological structure} & \text{not sensitive to morphological structure} \\
   \text{cyclic} & \text{non-cyclic} \\
   \text{derived environments} & \text{across-the-board} \\
   \text{exceptions} & \text{no exceptions} \\
   \text{structure preserving} & \text{introduce new sounds} \\
\end{array} \]

A large class of phonological rules has been identified that are sensitive to morphological structure. Most commonly such rules might apply if a form is morphologically complex, but not if it is monomorphic. Such rules typically apply only in derived environments (unless they are structure building), apply cyclically, allow exceptions, and usually manipulate only those sounds which are part of the underlying inventory of the language. On the other hand, rules which are not sensitive to morphology typically apply across-the-board, are non-cyclic and apply in an exceptionless fashion. This clustering of properties is accounted for within the framework of Lexical Phonology (Kiparsky 1982, 1983 and

\[2 \text{ Some recent work leads to a rethinking of the status of phonological rules in general (see notably Goldsmith 1993, McCarthy and Prince 1993, and Prince and Smolensky 1993); in particular the derivational metaphors that have dominated the generative literature are questioned. However for ease of exposition, I maintain the canonical generative phonology view of rules in the discussion here.} \]
Mohanan 1982), by positing two distinct sorts of rules. The rules that interact with the morphology are said to be *lexical* rules; they are interleaved with the morphology and have the chance to reapply after each morphological operation. These rules are inherently cyclic.\(^3\) Rules which apply in an across-the-board fashion, those which are not affected by morphological structure, are called *postlexical* and are argued to apply at the end of the grammar, after all of the lexical rules. This division gives a view of the grammar such as the one schematized in Figure 2.

\[
\text{lexical phonology} \quad \xrightarrow{\text{phonological rules}} \quad \text{morphological rules} \quad \xrightarrow{\text{postlexical rules}} \quad \text{postlexical phonology}
\]

**Figure 2.** View of the grammar following Lexical Phonology.

The lexical rules correspond with what has been called the *deep phonology*. The postlexical rules are nearer the surface and include the class of allophonic rules. Mohanan (1982) argues that the last level of the lexical phonology has a special status. In particular, speakers appear to judge sameness and distinctiveness at this level, not at the underlying level or a more surface level. Speakers' awareness of a rule's effect may be taken as support for something being part of the lexical phonology; e.g. speakers of English are typically sensitive to alternations due to Trisyllabic Laxing (*opaque ~ opacity*) a lexical rule. In contrast, naive native speakers are usually insensitive to allophonic variation in their language; e.g. speakers of English typically judge the /p/ of *pit* [pʰɪt] and the /p/ of *spit* [spɪt] to be the same (Mohanan 1982, Kenstowicz 1994). These distinctions will be important below in §3 in the discussion of rule types.

### 1.2 Phonological representations

It is worth considering in some detail the nature of phonological representation, in both time and space. Within the phonology, representations are discrete and categorical. As

---

\(^3\) Some debate exists about the nature of cyclicality and the status of level ordering, another independent tenet of Lexical Phonology, see Kaisse and Shaw (1985) for an overview and Hargus and Kaisse (1993) for a range of recent perspectives on Lexical Phonology and also Halle and Vergnaud (1987) for a somewhat different point of view.
mentioned above, it is widely assumed that segments consist of a hierarchically structured set of distinctive features, organized into a tree. While most researchers agree on the basic outline of this hierarchical relationship, debate still exists over the specific set of features and their precise relationships. Additionally there are two somewhat different views of the nature of, or motivation for, this hierarchical structure. The view put forth by Clements (1985) and McCarthy (1988), among others, is that this structure is a logical structure, motivated by phonological evidence. The view espoused by Halle (1983) and Sagey (1986), among others, is that this phonological structure is articulatorily motivated, with the structure mirroring the vocal tract. The position taken here is that the feature geometry must necessarily be a logical phonological structure and the degree to which this structure mirrors the vocal tract is an empirical question. It is important to bear in mind that such abstract phonological representations are necessarily both articulatorily and perceptually motivated. I take the geometry proposed by McCarthy (1988), schematized in Figure 3, as a starting point. One of the many issues, as yet not fully resolved is the location of the feature Nasal in the geometry, a point to which we return in Chapter 5.

Figure 3. The feature geometry, following McCarthy (1988).

The space dimension is represented as the logical structure of the feature geometry with feature values for each feature. A binary feature may be positively specified (+), negatively specified (-), or unspecified (Ø). Thus, possible representations for the feature Nasal include only [+nasal], [-nasal], and [Ønasal] (assuming it is a binary feature), as exemplified in the English word can /kæn/, shown in (3).
(3) \[ \begin{array}{ccc}
  k & æ & n \\
 [-\text{nusal}] & [\text{Ønasal}] & [+\text{nasal}] \\
\end{array} \]

Segments also include an abstract duration or timing unit, represented independently from the feature matrix itself. It is the timing or abstract duration of the segment that allows the segment to be related to higher level prosodic structure, which organizes segments into syllables, feet, and so forth. The possible timing contrasts available within the phonology are very limited. The inherent durations of different segment types do not play a role in the phonology. Only a very gross characterization of timing appears to be relevant. The relevant categories within a segment appear to be two units, one unit, and part of a unit, unspecified for specific duration or proportion. (Steriade 1990 makes a similar point.) More complex structures might arise through phonological derivation, e.g. Inouye's (1989) analysis of flaps in English, as well as her reanalysis of medio-nasals in Kaingang (as discussed by Anderson 1976), both of which are argued to involve structures with three branches. But there is little evidence for structures with more than two branches in underlying representation.

These categories are captured within current phonological theory as timing units (either C's and V's (e.g. Clements and Keyser 1983) or x's (e.g. Levin 1985), or as moras or weight units (e.g. Hyman 1985, McCarthy and Prince 1986). The mapping between segment quality, that is the features, and timing units represents phonologically relevant relationships, as exemplified in (4) (using x's).

(4) \( \begin{array}{ccc}
  a. & x & b. & x & x & c. & x \\
  \backslash & \backslash & & / & & / & \backslash \\
  \alpha & \alpha & & \alpha & \beta \\
\end{array} \)

The representation in (4a) is that of most segments, one set of features for one timing unit. (4b) represents one set of features taking up two timing units, such as a geminate consonant or a long vowel. (4c) represents a contour segment, such as an affricate, a prenasalized stop, or a short diphthong. These can be viewed as two feature matrices attached to one timing unit, or a feature matrix with one branching feature (see Sagey 1986). It has been observed that, although geminates are consistently longer in duration than single segments, the relative timing of single segments and long segments varies from language to language. This does not invalidate the notion of timing units, but rather shows that they are abstract entities. On an abstract level only, geminates are twice as long. But the actual physical timing is a matter of phonetic realization which is, to some degree, language specific.
The representation in (4c) implies that two feature specifications share one timing unit. Each specification gets part of the whole. It does not say what the specific timing relationships are; this again would be language specific. It has been shown by Chan and Ren (1987) that prenasalized stops vary in their relative timing, yet I assume that this would not affect the phonological representations. A prediction that follows from this is that no language will distinguish between single unit contour segments of different proportions (e.g. \([m^b]\) vs. \([m^b]\), where superscript denotes shorter duration). To my knowledge, no language makes a phonological contrast of this sort. In recent work, Steriade (1993, 1994) has proposed a rather different type of structure to account for contour segments, arguing that only [-continuant] consonants may show these contour effects, as they consist of both a closure and release phase. Thus a prenasalized stop consists of a closure phase which is nasal and a release which is oral. This provides a much more constrained theory of contour segments. The implications of such representations will be explored in Chapter 4.

1.3 Degree of phonological feature specification

An area of interest in the recent phonological literature has been the degree of specification in phonological representations. In other words, is it the case that for every segment each feature is specified for a plus or minus value? If so, the representation is a fully specified one. If not, if no specification is possible, there is said to be underspecification. In SPE, even though lexical representations were underspecified, with values provided through markedness and predictability of certain feature values from other feature specifications, it was explicitly stated that all features must be fully specified in the phonology. Thus all values were filled in at the level of underlying representation of the phonology. This led to a constrained view of feature representation, but a rather unsinsightful approach to certain phonological processes, such as harmony rules, which were represented as feature changing rules. Recently this constraint on full specification has been rejected (Kiparsky 1982, Pulleyblank 1983, and Archangeli 1984); it is now generally assumed that some degree of underspecification is used by the phonology. Yet there is much debate in the literature as to the degree of underspecification allowed, the way in which underspecification should be constrained or restricted, and how rules interact with representations which are not fully specified (cf. Kiparsky 1985, Archangeli & Pulleyblank 1986, Steriade 1987, Archangeli 1988, Christdas 1988, Clements 1988, 1993, Mester and Itô 1989, and Steriade 1995 for discussion of these issues).

Two types of approaches to underspecification have played a central role in these discussions: (a) the approach taken by Kiparsky (1985) and Archangeli and Pulleyblank
(1986), often termed Radical Underspecification, in which no reference is made underlyingly to either redundant or unmarked values; following this view, the unmarked value of a feature cannot be referred to underlyingly or in the phonology, until filled in by rule; (b) the type of approach taken by Steriade (1987) and Clements (1988), referred to here as Contrastive Underspecification, where in case of contrast, both feature values are specified in the underlying representation, but in the case of redundant specification, no value is present. See Archangeli (1988) and Mester and Itô (1989) for insightful discussion of the development of underspecification theory and comparison of these two types of theories, and Clements 1993 and Steriade 1995 for a recent perspective on some of these issues.

Relevant to the present discussion is the widely agreed upon assumption that redundant feature values are underlyingly unspecified; I assume that there is some degree of underspecification and that certain phonological patterns may be due, not to individual feature specification, but to more general principles. For the sake of explicitness, I follow a Contrastive Underspecification approach. While in some of the cases investigated here Contrastive Underspecification offers a more transparent interpretation of the observed patterns than Radical Underspecification does, nothing crucial hinges on this choice and the facts in the present study are compatible with both types of approaches.

As an example, consider the specification of the feature Nasal in both French and English, languages differing in their phonological use of the feature, as presented in (5).

(5) Phonological specification for Nasal

<table>
<thead>
<tr>
<th>Type</th>
<th>French</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Stops</td>
<td>[-nasal]</td>
<td>[-nasal]</td>
</tr>
<tr>
<td>Nasal Consonants</td>
<td>[+nasal]</td>
<td>[+nasal]</td>
</tr>
<tr>
<td>Continuant Consonants</td>
<td>[Ønasal]</td>
<td>[Ønasal]</td>
</tr>
<tr>
<td>Vowels</td>
<td>oral</td>
<td>[Ønasal]</td>
</tr>
<tr>
<td></td>
<td>nasal</td>
<td>[+nasal]</td>
</tr>
</tbody>
</table>

Both French and English make a contrast between nasal and oral stops, e.g. French net /net/ 'neat' ~ dette /det/ 'debt', English net /net/ ~ debt /det/. I therefore assume that, in both languages, the nasal consonants are specified as [+nasal] and the oral stops as [-nasal]. In both languages, the [+continuant] consonants show no contrast and I assume remain unspecified for the feature Nasal. French, however, has a contrast among the vowels as well, e.g. beau /bo/ 'beautiful' ~ bon /bo/ 'good'; I conclude that the nasal
vowels are specified as [+nasal] and the oral vowels as [-nasal].\textsuperscript{4} English shows no such contrast among the vowels; thus, in the absence of evidence to the contrary, I assume that vowels in English remain unspecified for Nasal throughout the phonology.

An alternative view of feature specification is that some or all features are privative, that is, they are present or absent (see Lombardi 1991 and Steriade 1995). In particular, some researchers (e.g. Itô and Mester 1989, Steriade 1993, 1995, and Trigo 1993) have argued that the feature Nasal is privative. This claim has important implications for phonological representation, as well as for phonetic implementation; the phonological and phonetic consequences of this view are discussed in Chapter 3. For the moment, I maintain the view that Nasal is a binary feature.

In sum, within the phonology, phonological features are hierarchically arranged; feature specifications are represented as only plus, minus, and unspecified; and timing is represented in abstract units: x, two x's, or part of x. At the end of a phonological derivation, the output is still a discrete, static one, with logical structure, relative ordering relationships between segments (or feature specifications), but no actual (concrete) durations, or quantities along the physical dimensions corresponding to features. Rather it is the job of the phonetics to assign real durations and actual quantities.

2. Phonetic rules and representations

In this section, I consider the nature of phonetic representation (§2.1), phonetic implementation (§2.2), and the role of underspecification in the phonetics (§2.3).

2.1 Phonetic representations

In SPE, it is assumed that there is a phonetic level of feature representation, the phonetic transcription, at which point binary values are translated into a small number of discrete categories. The phonetic transcription is described as follows:

The phonetic transcription can therefore be taken to be a two-dimensional matrix in which the columns stand for consecutive units and the rows stand for different features. At this level of representation each feature is to be thought of as a scale. A particular entry in the matrix, then, indicates the position of the unit in question on the given scale. The total set of features is identical with the set of phonetic properties that can in principle be controlled in speech; they represent the phonetic capabilities of man and, we would assume, are therefore the same for all languages. (pp. 294-5.)

\textsuperscript{4} It has been widely debated in the literature whether the nasal vowels of French are underlying or derived (cf. Schane 1968, Dell 1970, Tranel 1981, Prunet 1986, among others). What is important here is that, in either case, a contrast exists at the level of representation relevant to the present discussion; since, if the nasal vowels are derived, it is in the deep phonology.
This, then, is a linguistically relevant, discrete, phonetic representation. Everything beyond the phonetic transcription is assumed to be universal. Anderson (1974) offers a concrete instantiation of such a phonetic representation. As an example, he discusses three possible phonetic values for the feature Nasal in the Breton dialect of Plougrescant – [0nasal], [.3nasal] and [.7nasal], on a scale of zero to one – to describe oral vowels in an oral context, oral vowels next to a nasal, and distinctively nasal vowels, respectively. Such an approach, in effect, codifies the phonetic values that exist at the level of the phonetic representation (which Anderson notes are only significant in their relative relationship to each other).

This SPE-type view gives a formal status to the phonetic representation – a still discrete, but detailed, representation. It is not obvious that such a level of representation is necessarily warranted. (See Pierrehumbert 1990 for discussion of this point.) It might also be the case that the status of the phonetic representation is different for different features. For any particular feature, we can ask whether there are a certain number of points along the scale that are linguistically relevant. This is an empirical question. Keating (1984) argues that, for the feature Voice, there is evidence of an intermediate formal level of phonetic representation in terms of categories distinct from the phonological categories. She offers a detailed and formal proposal for such a level of representation for Voice. She argues that the phonological feature Voice is represented at an intermediate stage as three phonetic categories: [voiced], [voiceless unaspirated], [voiceless aspirated]. Different languages will map the two binary values [+voice] and [-voice] differently into the phonetic mapping. This accounts for the kind of similar phonological processes observed across languages with respect to Voice, even though what is voiced in one language may count as voiceless in another. (However, in more recent work, Keating 1990b no longer ascribes the same formal status to an intermediate level of representation.) But even if this is a correct view for Voice, it does not follow that such a level of representation necessarily exists for other features.

Much of the discussion of the relationship between phonology and phonetics has focused on the nature of the phonetic representation (Chomsky and Halle 1968, Anderson 1974, Ladefoged 1977, van Reenen 1982, Mohanan 1986). This emphasis on representation rather than process results, I believe, from the formal status attributed to the phonetic transcription as defined by Chomsky and Halle. But such a level of phonetic representation may or may not be relevant to implementation. Following the view of implementation proposed by Pierrehumbert (1980) and Pierrehumbert and Beckman
(1988), the phonetic representation is an extension of the phonological representation. The phonologically discrete units get increasingly blurred through the implementation of different parameters. What is taken as the phonetic representation is a result of the phonetic implementation; it may or may not be the case that such representations should be formally codified. Thus for example, the [.3nasal] level observed on vowels next to a nasal consonant in Breton might arise from the transition between an oral segment and a nasal segment. The [.3nasal], although descriptively accurate, may not be directly relevant to our understanding of the mapping. It might be the case that such a value is the result not of a phonetic target, but of a transition between two more extreme values. Values such as [.3nasal] may arise as a result of change over time. If this is the case, we can still describe the phonetic event in these terms, but do not necessarily want to ascribe a formal status to this description. In the analysis in this study, although phonetic representations are presented, I focus primarily on the process of implementation, in the belief that the phonetic representation of the feature Nasal follows from its implementation.

2.2 Phonetic implementation

The formal properties of phonological rules have been a topic of ongoing research in generative phonology; the formal properties of phonetic implementation rules are much less studied and much less clearly understood. The nature of phonetic implementation is a complicated issue for a number of reasons.

First, phonetics, representing the physical properties of speech, includes a number of possible representations. Denes & Pinson's (1993) characterization of the speech chain, presented in Figure 4, illustrates this multifaceted nature of phonetics.

![Speech chain diagram](image)

**Figure 4.** The speech chain (adapted from Denes & Pinson 1993, p. 5).

At a minimum, we can talk about the articulatory event, the acoustic event, and the auditory event, as being part of the phonetics. We might further distinguish between the higher
level articulatory gesture and lower level motor events and between auditory and perceptual events. The implicit question that shapes researchers' assumptions about the nature of phonetics is which facets of phonetic activity are under speaker or hearer control. When a speaker produces a labial gesture, as in the sound /b/, are the individual muscles commanded to fire, or are the individual muscle firings a result of a more global command to make a labial gesture? Research shows that higher level aspects of such events are under linguistic control, as are some of the lower level ones. I focus here on higher level events; while in a sense this is an arbitrary decision, it seems reasonable to assume that higher level events are more likely to be relevant to linguistic structure.

Second, even if we could fully resolve the first issue, we would still be faced with the issue of how should we best describe or model phonetic information. Consider the schematic example of an articulatory gesture, a movement of the velum, presented in Figure 5A. Here the velum begins in a high position at the beginning of the vowel (i), starts to lower part-way through the vowel, reaching a low position by the beginning of the nasal consonant (ii) which is maintained throughout the duration of the nasal consonant (during the oral closure of the consonant), then it starts to rise again during the following vowel (iii).

![Diagram](image)

**Figure 5.** A schematic velic gesture in /ana/, A. movement of the velum, B. static model of this movement, C. dynamic model of this movement.

We might argue that this phonetic event is best characterized by the high and low positions of the velum, with transitions connecting the targets. This view suggests that the
fundamental information is provided by the maxima and minima, while the transitions are, in effect, a by-product. I characterize this as a static view of the phonetic representation, since it is the endpoints, not the movement, which are taken to be primary. Such an approach is basically a target-interpolation model: maxima and minima are targets, connected through interpolation. Target-interpolation models of phonetic implementation have been discussed in the literature in two different contexts. First there are the rule-based systems of speech synthesis (e.g. Holmes, Mattingly, and Shearme 1964; Mermelstein 1973; Hertz 1982; Allen, Hunnicutt, and Klatt 1987, among others; see Klatt 1987 for a review). More recently target-interpolation models have been applied to the problem of the mapping of phonology to phonetics (Pierrehumbert 1980, Pierrehumbert and Beckman 1988, Keating 1985a). (These two domains of investigation are not necessarily distinct as shown by the linguistically based speech synthesis work of Hertz 1991 within her Delta system.) Within a target-interpolation model, we might model the event in Figure 5A as shown in Figure 5B.

Alternatively we might take the movement to be the primary event and model the velic gesture as a gesture or an oscillation, as schematized in Figure 5C. I characterize this sort of view as a dynamic view of the phonetic representation: the movement is primary and the endpoints are a byproduct of the timing and amplitude of the oscillation. This is basically the view espoused by Browman and Goldstein in their recent work (e.g. Browman and Goldstein 1992).\(^5\)

A third issue that influences the notion of phonetic representation is the theoretical position we take vis-à-vis the relationship between phonetics and phonology. As argued below in §3.1, evidence supports the conclusion that the phonetics is not just the automatic, universal realization of phonological patterns. Rather the phonetics is, at least in part, language specific and therefore under linguistic control. Phonetic patterns are quite systematic and therefore assumed to be rule governed, the basic tenet of the growing field of generative phonetics. Furthermore, our notion of the phonology and the representation of the output of the phonology will influence what we take to be a simpler or more straightforward model of the phonetics.

Models of phonetic implementation can be characterized in terms of the form of the phonological output, discrete or continuous, categorical or gradient, static or dynamic; the nature of the phonetics, static or dynamic; and the role of timing within the phonetics, as intrinsic or extrinsic to phonetic events. Most models involve an explicit point of

---

\(^5\) For Browman and Goldstein the phonology is also characterized in gestural terms.
interpretation or translation, though some provide a more seamless view of the connection between phonology and phonetics. (See Pierrehumbert 1990 for discussion of different types of models.)

As discussed above, the view taken here is that phonological representations are discrete, categorical, and static. One possible view of phonetic implementation would be to add concrete durations to autosegmental associations. A model consistent with this view of phonology is the Look Ahead model of Henke (see Keating 1985a for a discussion of this model and its history). As Keating discusses, in this model a feature value changes as soon as it is able to. One can think of this as on-off switches, with the principle being to change to the next value as soon as possible, predicting extensive anticipatory, to no carry over coarticulation. As has been argued elsewhere in the literature, such a model is not adequate to account for the facts of nasalization (see Kent, Carney, and Severeid 1974 and Benguerel, Hirose, Sawashima, and Ushijima 1977, among others). However, more generally, such an approach alone would not be sufficient, as autosegmental representations are inherently qualitative; rather a quantitative evaluation of such abstract representations is required.

A richer phonetic interpretation of phonological information is needed. The job of phonetic rules is nicely characterized by Pierrehumbert and Beckman (1988, pp. 4-5):

The phonetic rules are like phonological rules in that they seek to describe complex regularities in sound structure through the interaction of a few general principles. They differ from phonological rules in the representations that they manipulate. They take as input phonological representations, but their output consists of quantitative functions, representing facts about articulations or sounds.

The class of target-interpolation models allows quantitative evaluation of such representations with few assumptions. Target-interpolation models are typically assumed to assign static targets to phonological feature values. Recent work suggests that targets may have inherent duration (Pierrehumbert and Beckman 1988, Huffman 1989) or more than one target per phonological segment may be required (Huffman 1989, Laniran 1992). Additionally, nothing excludes the possibility that discrete phonological representations could be mapped to targets which are basically gestural in nature (a point to which we return in Chapter 6); such a view is argued for by Zsiga (1993).

An outstanding example of phonetic implementation within a target-interpolation model is Pierrehumbert's (1980) dissertation, in which she models intonation in English going from an abstract (sparse) phonological representation to observed physical output. This approach has been developed in further work (notably Beckman and Pierrehumbert 1986
and Pierrehumbert and Beckman 1988). Such an analysis relies heavily on phonetic implementation rules. Pierrehumbert develops the traditional distinction in implementation between evaluation and interpolation rules. The former rules evaluate feature values, translating them into phonetic targets located in both time and space. The latter rules connect up targets. This approach offers a possible model for examining segmental processes, as proposed by Keating (1985a).

Within a target-interpolation model, feature values leaving the phonology are translated into phonetic targets, such that a plus value translates to relatively more of the physical value that implements a particular feature than a minus value; these targets are then hooked up through interpolation, as illustrated in Figure 6:

```
Phonological Output:  +F  -F  +F
Phonetic Targets:      high  •       •
                      low  •
Phonetic Interpolation: high  ———
                         low
```

**Figure 6.** The target-interpolation model of phonetic implementation.

In this example, we see a sequence of segments with values for some feature F. The vertical scale is the physical dimension. The target evaluation rules assign phonetic targets based on phonological specification, with the +F specifications receiving a high target, and -F a low target, for the particular phonetic parameter (or parameters) that corresponds to the phonological feature. I assume for the moment that targets consist of single points located in the middle of the duration of a segment, with linear interpolation between the targets.

This model raises issues about the nature of targets, their assignment, and interpolation. We return to the nature of targets in Chapter 3. Many patterns of interpolation appear to follow from the linear connection of targets. This is the case for some of the intonational patterns described by Pierrehumbert or the case of jaw lowering (see Keating 1987). But consideration of a broader range of data makes it clear that a restrictive theory of interpolation which allows only for straight lines is not empirically adequate. Some of the patterns presented by Pierrehumbert, as well as the case of tongue backing in Arabic discussed by Keating (1987) show greater complexity. Pierrehumbert (1980) proposes a model of interpolation, including three types of curves to account for the patterns that she
presents. Choi (1992) also finds both linear and non-linear patterns of interpolation in his phonetic study of frontness vs. backness in Marshallese. See Choi (1992) for a comprehensive discussion of the issues relating to the nature of interpolation. In the course of the present study, I assume linear interpolation, although this is something of an oversimplification.

This view of phonological and phonetic structure and the nature of the mapping between them can be contrasted with the kind of approach taken by Browman and Goldstein (1992 and work cited therein), within the model of Articulatory Phonology. Browman and Goldstein argue that both phonology and phonetics can be represented in terms of gestures, articulatory events with inherent timing, indicating local constrictions in the vocal tract. Following this view, the relationship between phonology and phonetics is essentially seamless, as the same structures are manipulated by both the phonological and phonetic components. The categorical nature of lexical representations is accounted for by imposing constraints on gestural overlap and timing. This model could be characterized as one which is inherently quantitative, with the more qualitative aspects of the phonology accounted for by the imposition of constraints. (Such constraints have yet to be explicitly laid out. Until such constraints are made more explicit, the categorical representations of the phonology appear ad hoc.) This view takes articulation to be primary at both the phonetic and phonological levels.

Both approaches have been used quite successfully to model phonetic events and in each case a number of developments and amplifications have been made. While Articulatory Phonology is attractive in a number of respects, I follow Steriade (1990), Clements (1992), and others, in arguing that an autosegmental, feature geometric approach offers a more predictive, constrained view of phonology. I take such a view as a starting point for the mapping from phonology to phonetics accounted for by a target-interpolation model. My primary goal here is to investigate how explicit phonological representations can be interpreted by the phonetics. While I will argue that phonetic patterns of nasalization are insightfully captured within a target-interpolation model, this does not rule out the possibility that features are not all uniform in this respect. Features may differ as to whether they are best modeled in static or dynamic terms. It should be noted that the phonetic representations that I develop are still quite abstract in a number of respects. In particular, I assume the existence of an extrinsic timing model, not developed here.
2.3 Phonetic underspecification

An issue basic to the discussion of phonetic implementation is whether there is necessarily full specification of features at the output of the phonology. Full specification leaving the phonology is of course assumed in any theory which does not allow phonological underspecification. But even under certain views of phonological underspecification, it is often assumed that there is full specification at the end of the phonological derivation, due to default fill-in rules (e.g. Archangeli and Pulleyblank 1986). In contrast, in order to account for patterns of pitch, as measured from F0 traces, Pierrehumbert (1980) assumes a sparse phonological output. Not every syllable which is potentially the bearer of a pitch accent receives such an accent. In effect then, the phonological output is not fully specified along the dimension that is implemented as intonation. Other examples include the realization of the vowel-like quality of intervocalic [h], which Keating (1988b) argues is phonologically unspecified for vowel quality, the vowels of Marshallese, argued by Choi (1992) to be unspecified for backness, and the nasalization of continuant consonants in Sundanese, as discussed below in Chapter 5 and by Cohn (1993a). In order to account for these observations, it is necessary to abandon the assumption that there is full specification leaving the phonology.

A target-interpolation model taken together with the possibility that there is not full feature specification leaving the phonology leads to the following schematic view of implementation: a plus value of a feature should translate to a relatively high level; a minus value to a low level, and an unspecified value would be expected to be determined by phonetic context. The rules of the phonetic implementation evaluate the targets and connect them through interpolation. Three examples are given in Figure 7.

\begin{verbatim}
Phonological Output:    +F  -F  -F    +F  ØF  -F    +F  ØF  +F
Phonetic Targets:       high *  high *  high *  high *  high *  high *
                        low  *  low    low  *  low
Phonetic Interpolation: high *  high *  high *  high *  high *  high *
                        low  low
\end{verbatim}

**Figure 7.** Schematic examples of three patterns of phonetic interpolation allowing for phonetic underspecification.

In Figure 7A the form is fully specified for the feature F leaving the phonology. Targets are assigned, along some scale for the particular feature. These targets are then hooked up
through interpolation, showing a fairly rapid transition between the neighboring high and low targets. In Figure 7B, the first segment is [+F], the second segment is unspecified leaving the phonology and the final segment is [−F]. Only the first and last segments receive phonetic targets. There is interpolation between the targets and the middle segment receives a transitional amount of the scale for the feature F from the phonetic context. In Figure 7C, both the first and last segments are specified as [+F] and the middle segment is again unspecified. Targets are assigned and in this case, there is interpolation straight through the middle segment. The unspecified segment receives a high value of the relevant phonetic parameter due to the phonetic context, giving the (erroneous) impression that it had a relatively high target for Feature F, when in fact it has no target of its own.

Let us look more closely at the vowel-like qualities of intervocalic [h], as discussed by Keating (1988b), as an example. It has commonly been assumed that [h] gets its vowel-like qualities from a phonological assimilation rule, where [h] assimilates to the vowel features of the following vowel (e.g. Ladefoged 1993). Were this the case, we would expect that [h] would have similar formant structure (the characteristic acoustic pattern of a vowel or vowel-like segment, indirectly related to the phonological vowel quality features) throughout most of its duration as that of the following vowel. Keating studied formant patterns of [h] and concluded that the formant structure was not due to a phonological rule, but due rather to phonetic context. Consider relevant spectrographic examples, [iha] and [aha] presented in Figure 8A & B, respectively.

A. [iha]

B. [aha]

**Figure 8.** Wideband spectrograms of a female speaker of American English saying A. [iha] and B. [aha].
As shown in Figure 8A, [i] is characterized by having a low F1 (centered here at about 300 Hz) and a high F2 (centered at about 2700 Hz); [a] is characterized by a high F1 (centered at about 1000 Hz) and a low F2 (centered at about 1500 Hz). The formant structure is constant throughout the duration of the vowel. Were it the case that [h] gets its formant structure from a phonological rule of assimilation from the following vowel, we would expect the [h] to have formant values like the following [a] for most of its duration. This is not the case; rather, the F1 and F2 values are transitional throughout the full duration of the [h]. This can be explained in a straightforward fashion, if it is assumed that [h] is unspecified for vowel quality leaving the phonology; it therefore gets no phonetic targets of its own and its vowel-like qualities are from the phonetic context, due to interpolation. Note that this conclusion is partially dependent on the segmentation of the [h] and the adjacent vowels. It is difficult to determine precise segmentation of this sequence, since the [h] is voiced throughout. The realization of F1 and F2 for [h], then, looks very much like the pattern of phonetic implementation schematized above in Figure 7B.

Consider now the spectrogram in Figure 8B. Here again the [a]'s both preceding and following the [h] are characterized by high F1's and low F2's. The formant structure of [h] is also characterized by a high F1 and low F2. Had we not considered the evidence in Figure 8A, we might (erroneously) assume that [h] has its own targets, high F1 and low F2, but based on that evidence, it is much simpler to assume that here too [h] has no vowel feature specifications of its own, therefore no phonetic targets, and that it too gets its vowel-like quality from the phonetic context, through interpolation between the neighboring targets. This then is like the pattern of phonetic implementation schematized in Figure 7C. It is important to note that such cases can easily be misinterpreted as all three segments, including the intermediate segment, having the same value specified. Comparison with forms where a transition might be expected throughout are crucial for determining such cases of underspecification leaving the phonology.

Keating's account of intervocalic [h] offers strong support for the view that there is not necessarily full feature specification leaving the phonology. The facts are amenable to an analysis in terms of the target-interpolation model (but could clearly be accounted for with other models of phonetic implementation as well). The assumption that there is not full specification leaving the phonology has important consequences. (1) It is not necessarily the case that there are rules which assign phonological feature default values at the end of the phonology. (2) Since a well-formed phonological output might not be fully specified along all dimensions, it is an empirical issue to determine whether a particular
representation is fully specified and whether or not default phonological values have been assigned.

A particular view of phonetic underspecification is Keating's (1988a, 1990a) Window Model. Keating proposes an explicit model of interpolation and phonetic spreading which assumes that rather than having a single target in the vertical dimension (space) for each specified value, there are ranges within which interpolation must stay. Thus a segment with a very narrow window would appear to be one with a specific target and one with a wide window would appear to be underspecified. This approach offers an account for certain kinds of variability, as well as a means of constraining possible types of interpolation.

In the present study, I take a target-interpolation model as the starting point to account for the phonetic implementation of the feature Nasal. I argue that such a model needs to be modified and extended to account for the observed facts. The characterization of targets in terms of both time and space is extended from the concept of a single point located in the horizontal and vertical dimensions. I argue that targets have inherent duration and a specific account of such duration will be proposed. Further, I argue that windows, or ranges of values, play a role in implementation.

3. Phonology vs. phonetics

With this brief sketch of rules and representations, both within the phonetics and phonology, we can now turn to the question of the relationship between phonetics and phonology.

3.1 Background on the relationship between phonetics and phonology

In order to study the relationship between phonology and phonetics and the nature of phonetic implementation, we have to be able to distinguish between them. Within the framework of generative phonology, phonological representations and derivations have been assumed to be a central part of the grammar, but until recently phonetics was often assumed to fall outside of the domain of the linguistic grammar. A widely held view, following SPE, was that phonetic implementation was universal. This was discussed explicitly in terms of coarticulation. Coarticulation, the overlapping or blending of neighboring segments, was assumed to follow from general non-linguistic principles, such as the physiology of the vocal tract. Stated more generally, phonetic implementation or the physical realization of the abstract patterns represented by the phonology was assumed to be mechanical. As a consequence, a phonological output was assumed to have a unique physical realization. For example, in English, vowels are longer preceding a voiced stop
than a voiceless one: the vowel in *bad* /bæd/ is longer than in *bat* /bæt/. This was not taken as evidence of a phonological length difference in English, but, rather, was assumed to follow from mechanical implementation of the voicing contrast. It was also assumed that the same differences occurred cross-linguistically. Under this view, the distinction between phonetics and phonology appeared clear-cut. Phonology involved language specific rules, whereas phonetics was the universal, mechanical realization of the phonology. This view, where only the phonological rules were part of the linguistic grammar, is schematized in Figure 9.

```
phonological rules  = the linguistic grammar
                         \  
universal phonetic implementation rules
                         \  
physical output
```

**Figure 9.** Traditional view of the linguistic grammar (SPE).

Since the phonetics was *automatic*, there was no notion of there being a phonetic rule system as such. As the mapping between phonological rules and phonetic implementation was thought to be universal, little attention was paid to the phonetic implementation of phonological representations from a linguistic point of view. Such questions were the purview of phonetics alone and not thought to be of interest to the phonologist.

More recently, this view has changed. Extensive evidence has been provided of the language specific nature of phonetics. Many phonetic processes that were assumed to be mechanical, on closer inspection, turn out to show significant differences between languages. In such cases, a purely mechanical explanation is no longer tenable. One such example, widely discussed in the literature, is the observation about vowel length mentioned above (see Chen 1970, Fromkin 1976, Anderson 1981 and Keating 1985b, among others). Physiological explanations have been offered, but these do not account for the observed cross-language differences, in which some languages suppress these differences and others enhance it.

Another such example is velum movement. It has been argued that contextual nasalization of vowels in English is due to the *sluggishness* of the velum. But studies have shown that the velum can move much more quickly than the degree of coarticulation
observed in English would suggest (see Ohala 1975 for discussion). Furthermore, cross-linguistic work has shown significant language differences with respect to the timing of contextual nasalization (Clumeck 1976). At best, we can say that these effects are physiologically motivated tendencies (see Anderson 1981).

Phonetic differences can even occur between closely related dialects. Such is the case with the realization of the time course of nasal vowels in European French and Canadian French, as shown in Figure 10.

![Time course of nasal vowels](image)

**Figure 10.** Time course of nasal vowels of European French & Canadian French, after van Reenen 1982 p. 74, showing a graphic representation of averages of 12 European French and 31 Canadian French nasal vowels in nonnasal environments, based on measurements from tracings in Brichler-Labaeye (1970) and Charbonneau (1971). Time in centiseconds on the x-axis and the value for N% on the y-axis.

\[
N\% = \frac{\text{nose coupling}}{\text{mouth coupling} + \text{nose coupling}} \times 100,
\]

where N = Nose Coupling, the opening in mm\(^2\) of the nasal port in a cross-section perpendicular to the airstream at the point of greatest constriction between the velum and the pharyngeal wall.

In Figure 10, we see the time course of velum movement during a nasal vowel in European French compared with Canadian French (e.g. [œ] as in bon [bʊ] 'good' (m.)). This figure (adapted from van Reenen 1982) shows averages of nasal opening vs. time for nasal vowels in non-nasal environments, measured from x-ray tracings for European French and Canadian French from Brichler-Labaeye (1970) and Charbonneau (1971) respectively. We see that there is a different time course, with the onset of velum lowering (onset of nasalization) occurring much sooner in European French than Canadian French.
The difference is not a mechanical one; there is no reason to assume that the physiology of Canadian speakers of French is different from that of European speakers. Yet it cannot be accounted for in an obvious way by differences in the phonological system, since the value of nasal vowels in the two systems is the same. Finally, these language specific differences are systematic ones, and thus appear to be rule governed.

Based on such evidence, it has been convincingly argued that at least some phonetic rules are indeed part of the linguistic grammar and are not to be relegated to the universal phonetic component, leading to a view such as the one schematized in Figure 11, where both the phonological and the language specific phonetic rules are part of the grammar.

![Diagram showing the relationship between phonological rules, language specific phonetic rules, universal phonetic implementation rules, and physical output.](image)

**Figure 11.** Current view of the linguistic grammar.

This more complex view of phonetics leads us to seek a principled characterization of the difference between phonological and phonetic rules which follows from the mechanisms involved. Language-specific differences in phonetic realization identify aspects of the phonetics which are not universal and therefore not mechanical. Such differences may serve as a device for identifying linguistically relevant phonetics. Yet how can these cases be systematically distinguished from the phonology?

One possible view of the relationship is that language specific phonetic rules are a subset of phonological ones and should be accounted for with phonological rule mechanisms and representations. Another view is that phonological rules and language specific phonetic rules differ crucially in that what they manipulate is different: phonological rules manipulate discrete categorical representations, whereas phonetic ones manipulate quantitative ones. This fundamental difference necessitates that different formal mechanisms are involved. It is this latter position that will be taken here. We should
therefore seek a principled characterization of the difference between phonological and phonetic rules which follows from these different properties.

3.2 Distinguishing between phonetics and phonology

The conclusion that there is indeed a class of language specific phonetic rules leads us to seek a principled way of characterizing such rules and distinguishing them from phonological rules. As discussed in the preceding subsection, language specific differences in phonetic realization identify aspects of the phonetics that are not universal and therefore not mechanical. Such differences may serve as a device for identifying linguistically relevant phonetics. On the other hand, if it is claimed that such rules are distinct from the phonology, we need a systematic way of making this distinction.

Keating (1990a, p. 452) argues that phonological rules are expected to affect most of a segment in a significant way, whereas phonetic rules can have more gradient effects. "Phonetic rules can thus, for example, assign a segment only a slight amount of some property, or assign an amount that changes over time during the segment". Keating (1987, 1988a) proposes the following clustering of properties:

(6) Phonological rules
    - categorical
    - discrete & timeless segments
    - static effects
    - full segment affected

Phonetic rules
    - gradient/quantitative
    - continuous in time and space
    - segment may vary in quality continuously
    - part of a segment affected

By static, we mean output that does not change noticeably during the duration of time assumed to be associated with a particular segment or sequence of segments; such effects might be characterized as a plateau (along the space dimension). Gradient effects might be of different types; of particular interest to us here are cases of change in space over time, resulting in a cline-like effect. That static, categorical outputs are the result of phonology and gradient outputs from transitions between segments are the result of phonetics is not surprising. More interesting is the fact that phonetic rules may affect a full segment or more in a continuously varying way. This result follows directly from certain assumptions about the nature of phonetic implementation, most importantly the observation that there is not necessarily full specification of features leaving the phonology. The distinction is not always an obvious one. Phonetic output may at times appear to be categorical. Recall our discussion of the vowel-like qualities of [h], where we saw cline-like effects throughout the [h] when the preceding and following vowel were different, but static effects when the
adjacent vowels were the same. We also cannot exclude the possibility that a phonological
distinction may appear gradient; as when, for example, a target has only a brief point-like
duration. Such a clustering of properties can be, at best, taken as guidelines.

Of particular interest are cases which appear to involve gradience in the phonology or
where other aspects of what has been described as phonological rule application look
phonetic. Keating (1988a, 1990a) discusses a number of cases of assimilation involving
gradient rule application that she argues are phonetic and not, as previously assumed,
phonological. As discussed above, she suggests that the vowel-like qualities of [h] are a
result of phonetic implementation and not phonological assimilation. Additional cases
where Keating argues a phonetic rule (applying gradiently) has been misanalyzed as being
phonological (applying categorically) include vowel allophony in Russian and the
spreading of emphasis from the emphatic consonants in some dialects of Arabic.
Pierrehumbert and Beckman (1988) argue that some aspects of pitch accent in Japanese
result from phonetic implementation, not phonological tone spreading. It is important to
note that such observations depend crucially on instrumental phonetic data as the difference
between gradient and categorical rule application cannot necessarily be determined through
impressionistic observation. Careful listening alone is not sufficient to determine the status
of a rule, since we as hearers impose categories (see Repp 1984 and work cited therein),
even on gradient phenomena.

There are several interesting cases discussed by Kiparsky (1985), in which gradience is
involved. He discusses cases where the same rule seems to apply at different points in the
grammar, once categorically, once gradiently. He argues that this is a result of lexical and
postlexical application of the same rule. I believe that these cases may be amenable to
analysis in terms of phonetic interpolation through an unspecified span. A case discussed
by Kiparsky, of particular interest here, is the case of Nasal Harmony in Guaraní. It has
been widely observed that there are long distance spreading effects of nasalization in
Guaraní. Noteworthy is the fact that some of these effects appear to be categorical,
whereas some are gradient. Kiparsky proposes that these effects are the result of the lexical
and postlexical application of the same rule. It is nasalization of the class of unstressed
continuants which results in gradient effects. This is a class of segments which
characteristically does not contrast for the feature Nasal cross-linguistically. There does not
appear to be evidence of the nasal specification of these segments playing a role in the
phonology. It seems plausible that this class of segments remains unspecified for the
feature Nasal and that these segments receive quantitative levels of nasalization through
phonetic interpolation. An analysis along these lines is proposed for the case of Sundanese nasalization, as discussed in Chapter 5.

Identification of gradient behavior is taken as evidence of phonetic, not phonological, rule application precisely because the mechanisms of phonology and phonetics are distinct: Phonological rules manipulate discrete, timeless segments, whereas phonetic rules manipulate variables which are continuous in time and space. Gradient or cline-like behavior is a result of phonetic implementation, while categorical output or plateaus result from phonological rule application. (See Pierrehumbert 1990 for an interesting discussion of the mechanisms involved in phonology and phonetics respectively.)

Note, I distinguish here between gradient and variable. I take gradient to refer to change in space over time, while variable refers to applicability of a process independent of phonological context. A process may be variable in two senses, first it may or may not apply (this is often accounted for by saying that a rule is optional), second it may apply to varying degrees, e.g. deletion can be seen as the endpoint along a continuum of shortening. Some researchers assume that gradience and variability go hand in hand. Although an interesting hypothesis (not yet explicitly tested to my knowledge), it does not follow from either definition. Optional rule application seems equally consistent with phonological or phonetic rule application, while variability in the second sense suggests phonetic implementation. I return to the issue of variability in Chapter 4.

Taking a somewhat different approach to the question of gradient vs. categorical behavior, Browman and Goldstein (1992, p. 171) observe that "processes occurring during the act of talking will cause gradient changes that can ultimately be perceived as a categorically different gestural structure". Clements (1992, p. 12) notes that Browman and Goldstein's view suggests that "speech is produced in a gradient fashion, but perceived (and thus represented) categorically". This interpretation offers another explanation of why phonetics is inherently gradient, while mental representations (the phonology) are categorical.

The distinction between gradient and categorical behavior is intuitively attractive, yet what sorts of independent evidence might be adduced to support the claim that such behavior correlates with phonetic vs. phonological rule application, respectively?

As discussed above, within the framework of Lexical Phonology (Kiparsky 1982), a clustering of properties is used to determine the status of rules, lexical vs. postlexical. Although the distinction between lexical and postlexical does not correlate directly with phonological vs. phonetic, use of similar criteria should be helpful in this case as well. If a rule is shown to have characteristic behavior of a lexical rule (applying cyclically, applying
only in derived environments, applying within certain grammatical domains, respecting Structure Preservation, or ordered before such a rule), we can conclude that it is indeed phonological. Such criterion are used to corroborate evidence from nasal airflow data in the following chapters.

Some researchers have equated phonetic rules with postlexical rules. For example, Liberman and Pierrehumbert (1984) suggest that, except for phrasal rules, postlexical rules are phonetic implementation rules. Kiparsky (1985, p. 94) argues that this position is too strong. He observes that there are two types of postlexical rules, those with a categorical output (as a result of feature changing or default fill-in) and those which are essentially phonetic in nature, with gradient, variable output. (Kaisse's 1990 distinction between two types of postlexical rules, P1 and P2 can be seen to parallel Kiparsky's characterization.)

Although the distinction is mainly terminological, I take the position that postlexical phonological rules and phonetic implementation rules are distinct (rather than there being two types of postlexical rules), since the rule mechanisms involved are distinct. A logical argument in support of this position can be made from the process of historical change. One common type of historical change is that a phonetic characteristic (typically based on some sort of universal tendency) is exaggerated until it is perceived as distinct from the conditioning factor. This is what Hyman (1976) has termed phonologization. Consider, for example, the case of the shortening and eventual deletion of a segment. A segment might have a tendency to shorten in certain prosodic positions or at faster rates of speech. The most extreme version of shortening is deletion. As a phonetic process we would expect the results of this to be gradient and perhaps also variable; even if the segment is deleted in some cases, there is evidence through the gradient and variable nature of the process that the segment is indeed present underlyingly. Yet once the deletion becomes increasingly systematic, it could easily be reinterpreted by the hearer or language learner as not being present at all. For such members of the speaker community, the process would be realized as a categorical one of deletion. There is no reason to suppose that, at the precise point that the effects of a rule are reinterpreted as categorical, it becomes a lexical rule. It follows that there are either two distinct sorts of postlexical rules or as assumed here that phonetic implementation rules are distinct from postlexical ones.

While logically postlexical phonological rule and phonetic implementation rules are distinct, differentiating between the two may nevertheless be difficult. Insight in such cases can be gained by considering more clearcut cases first. If we can identify cases which are clearly phonological and others which are phonetic and show that realization of these patterns correlates with the expected plateau vs. cline distinction, this enables us to
use this distinction to evaluate less obvious cases. In the subsequent chapters, we will proceed in this way, first studying patterns of nasalization where the phonological status is independently ascertainable.

We will find that the patterns of nasalization investigated in this study are indeed amenable to characterization in these terms; but even in light of these results, it does not follow that all features and types of data will necessarily be interpretable in such a transparent fashion. A genuine counterexample to this characterization of phonetic vs. phonological would be a case where there is a rule, argued to be phonological, due to cyclic application or rule ordering, which is shown to apply in a gradient fashion or where an underlying contrast is realized gradually. One such possible exception is the case of realization of Hausa tone (Inkelas, Leben, and Cobler 1987) where each syllable must be specified for a tone and a sequence of high tones appears to result in a gradient output.

At this point it might be useful to consider a schematic representation of the grammar, presented in Figure 12.

```
lexical phonology
       ↓
postlexical phonology
            ↓
phonetic implementation: target evaluation interpolation phonetic constraints
```

**Figure 12.** Schematic view of the grammar.

The implication of this model is that there should be interaction between lexical phonological rules and morphological ones, but not between postlexical rules and lexical ones, nor between the phonetic implementation and the phonological rules. This means that rules manipulating discrete and timeless units are not interleaved with those that manipulate quantitative values. Although following the model of Pierrehumbert and Beckman (1988), where phonetic implementation adds information, rather than replacing it, it is not impossible that phonological rules follow phonetic ones. Some cases of what appear to be such ordering paradoxes have been cited in the literature, e.g. Anderson
(1975). The ordering of phonological and phonetic rules is an area that warrants further investigation and will be able to be addressed more systematically as detailed work on the phonetics-phonology interface continues.

4. Organization of the study

In this chapter, the background has been laid out for the remainder of the study. I have investigated a number of issues about phonological and phonetic rules and representations and their relationship. A target-interpolation model has been proposed as a starting point to account for the implementation of the feature Nasal. In the subsequent chapters, this model will be extended and modified to account for patterns of nasalization in French, English, and Sundanese.

The structure of the remaining chapters is as follows. In Chapter 2, more specific background for the study is presented and methodology is discussed, including description of the apparatus used for data collection and methods of analysis. In Chapter 3, we start with the facts of French, where a wide range of possible contexts of nasal-oral sequences is studied. In order to account for the observed patterns, certain modifications are proposed to the target-interpolation model, including assigning inherent duration to targets and developing the role of phonetic constraints. In Chapter 4, we turn to the data of English, where again a range of contexts of nasal-oral sequences are studied. It is observed that the facts are basically compatible with the model proposed in Chapter 3, although variability must be taken into account. In Chapter 5, we extend our study to Sundanese, a language which exhibits long-distance effects of nasalization. In Chapter 6, the concluding chapter, some comparison between the three languages is presented, the proposed model is discussed, and implications of the analysis are considered.

5. References

PhD dissertation, MIT.


