1. Phonological Theory

JOHN GOLDSMITH

Subject	Theoretical Linguistics » Pholonogy
DOI:	10.1111/b.9780631201267.1996.00003.x

0 Introduction: Phonotactics, Alternations, Contrasts; Representations, Rules, Levels

In this first chapter, I would like to bring together the issues joined and the proposals encountered in the range of papers that follow. These papers discuss a broad range of topics, often cross-referencing each other, usually by way of support, though there is *some* controversy, which I will highlight in the course of in this chapter. The most effective way to bring out the general unity is to step back and formulate the questions that our current theories are intended to answer. I would suggest that the following three questions lie behind most of the work that we find in phonological theory:[1]

> 1 What constitutes a phonological word in a given language? Many of the things we do in analyzing the phonology of a language are part of the effort to answer this question: we characterize and make an inventory of the sounds in the language, how the sounds can be combined to form syllables and words, what the stress patterns are like in the language, and so on. Conditions on well-formed phonological words have traditionally been called *phonotactics*.
>
> 2 What is the nature of *alternations*, that is, the differences in phonological form that we observe in the realization of a morpheme in different contexts? From the phonologist's point of view, what we mean by "context" may be phonological or morphological, and both kinds of context are important in determining the phonological realization of various morphemes.
>
> 3 The final question lies at the doorstep of phonemic theory: What phonetic differences are *contrastive* in a given language? – that is, what sound differences can be used to mark a lexical or grammatical distinction? This may be the hardest of the three questions, and I will devote the greatest attention in this chapter to some of the current suggestions for how this question may be answered.[2]

Most of the everyday work of phonological theory focuses less on these three questions than on the conceptual tools that we employ in order to come to grips with the questions, and it is these tools that change far more rapidly than the questions themselves. The tools can be roughly divided into three groups as well. Phonological theory develops, in the first place, improved types of *phonological representations*; much of the difference in the look and feel of phonology over the past twenty years has come from the rapid changes that have occurred in the degree of articulation found in phonological representations. Second, phonological theory develops conceptions of *phonological levels*, with each level expressing distinct information and allowing distinct representations. Over the course of the development of phonological theory, the differences between the roles played by two or three basic levels in the phonology have always been central. This was true in the development of early phonemic theory, when phonological theory rested on a distinction between a phonetic and a phonemic representation. It is equally true today, when, for example, lexical phonology places certain

restrictions on the possibilities of representations in the lexicon, and quite different restrictions on the output of the postlexical phonology (see Mohanan's discussion of this, as well, in chapter 2 of this volume). Third, phonological theory employs the notion of a *rule*. Of these three, this is the most treacherous term to define in a way that can apply across all phonological theories; perhaps even the hope of achieving a satisfactory common description is unrealistic. But we would not be far off in characterizing phonological rules in the following way: they are the devices employed by the phonological theory to account for the relationship between representations at different levels. Certainly in simple cases this characterization applies in a straightforward fashion; within a structuralist phonemic analysis, a representation at the phonemic level /pat/ and a representation at a phonetic level [p^hat$^?$] will be related by the rules of the phonology of the language. The same can be said for an account within classical generative phonology (that of Chomsky and Halle 1968), though in this case the rules are organized so as to apply sequentially, constructing a derivation that links the underlying and the derived representation.

Putting these various notions together provides us with a nine-chambered grid, formed by the three traditional questions of phonological theory along one axis, and the three sorts of tools that are used along the other axis. In the next three sections, I will locate the papers in this *Handbook* along these dimensions, and discuss a few issues that are controversial at the present time.

	Phonotactics	Alternations	Contrasts
Representations	Autosegmental and metrical structure Prososdic hierarchy	Feature geometry and limits on kinds of assimilation Prosodic morphology	Underspecification theory
Levels	Licensing Abstractness Structure-preservation	Issues of stratal organization	Organization of the lexicon
Rules	Metrical theory Harmonic rule application Optimality theory	Equation of phonotactics and alternations: strict cyclicity	Structure-building versus structure-changing operations

1 Phonotactics

The most basic of the traditional goals of phonological theory has been to establish the means of specifying, for any given language, just what a phonologically well-formed word in that language is. This is the question of *phonotactics:* in what ways can the items of phonology be put together to make a well-formed word. We may, after all, wish to express the notion that [blik] is a possible word of English, while [bnik] is not. Among the possibilities that have been considered are the following:

a A well-formed word is one that is produced by taking an input string created by the morphological component, and applying the phonological rules of the language in the appropriate order.

b A well-formed word is one that consists of a sequence of well-formed syllables.

c A well-formed word is one in which all features (or autosegments) are associated to an appropriate skeletal position; all skeletal positions are associated with a syllable; and all syllables are associated with a foot.

d A well-formed word is one that simultaneously satisfies all the well-formedness conditions of the language (includign those given in (c)).

The first answer, (a), is roughly the position of classical generative phonology, that is, the theory

proposed in *The Sound Pattern of English* (Chomsky and Halle 1968, hereinafter *SPE*), an account which puts most of the burden of determining phonological well-formedness on the operation of a set of ordered rules. In this conception, there *are* no well-formedness conditions as such; a representation is well-formed by virtue of where it came from, not what it is.

But it has become widely accepted – to the point where we may say it is simply established – that the complete elimination of phonotactics in favor of rule operation misses important generalizations, and the first area studied in this way – and the best-studied one – is that of syllable structure, which is discussed in detail by Juliette Blevins in chapter 6. Well-formed words, it has been argued, consist at the very least of well-formed syllables: answer (b); the ill-formedness of [bnik] is due to considerations of sonority that are intimately involved in the definition of the well-formed syllable.[3]

Further thought and investigation has shown at least two considerations indicating that reducing phonological word well-formedness to syllable well-formedness is only a first approximation. First of all, the phonological material that can occur word-initially is not necessarily the same as what can appear syllable-initially, and the phonological material that can occur word-finally is not necessarily the same as what can appear syllable-finally; in either case, the word-initial or word-final positions may be more restricted or more relaxed than what would be expected if syllable well-formedness told the whole story.[4] Second, syllables with different prosodic prominence (stressed/unstressed syllables have different prosodic prominence) have different possibilities; a well-formed stressed syllable will often have characteristics that would not let it qualify as a well-formed stressed syllable (English is a typical example of a language with a greatly reduced class of vowel contrasts possible in unstressed position compared to stressed position).

It has proven both helpful and insightful to synthesize these observations by means of a hierarchy of prosodic categories, as suggested by Selkirk, Nespor and Vogel, Hayes and others (see chapter 15 by Sharon Inkelas and Draga Zec on the syntax-phonology interface, and chapter 9 on prosodic morphology by John McCarthy and Alan Prince), stretching from the autosegment at its smallest end, through the skeletal position, the syllable, the foot, and the phonological word (and extending to larger categories than the word as well). Each unit of phonological organization is subject to restrictions governing what elements may (and what may not) make up such a unit. The most familiar case is the way in which segments are composed of skeletal positions associated with features; there are always restrictions (both lower bounds and upper bounds) on which features, and how many features, may associate to a single skeletal position. These restrictions amount to a statement of the segmental inventory of the language. The syllable is subject to well-formedness conditions as well regarding the complexity of each of its components, such as the onset and the rhyme (see chapter 5 by Ellen Broselow and chapter 6 by Juliette Blevins), just as feet are subject to well-formedness conditions on the complexity of syllables in their different foot positions: answer (c).[5]

Are there further well-formedness conditions? In general, the answer is positive. For example, Junko Itô and R. Armin Mester mention in chapter 29 a constraint in Japanese against single (i.e., nongeminate) *p* which holds in native and Sino-Japanese forms; in that dominant part of the vocabulary, only geminate *p*, or *p* following an *m*, may appear. While that additional constraint is best stated in terms that *use* the vocabulary of syllable structure, it is not (at least in any sense formulated to date) a statement about syllable structure perse; it is a statement about a particular combination of features, of syllable structure, and of intersyllable structure. Hence, (d) as a general statement that includes (c) plus other, language-particular generalizations, is the best formulation at this time.

Notions of licensing have been developed by a number of phonologists for several ends, of which the most direct is the need to express the fact that higher-level prosodic categories, such as the syllable, place restrictions on what kind of information can appear within them. A language may permit a segment within the coda of its syllable, for example, without permitting the full range of consonants to appear in that position. Restrictions on what features and segments a prosodic position can license, combined with the requirement that in a well-formed representation, all elements be licensed, results in many cases in the proper statement of what constitutes a well-formed word in the language.[6] David Perlmutter, in chapter 8 on phonological quantity and multiple association, discusses how licensing can account for the odd distribution of geminate obstruents, which in many languages can appear stretched over the coda of a syllable and the onset of the following syllable, even though obstruents may not otherwise appear in coda position in those languages.

During the last few years considerable interest has been generated by the observation that prosodic categories, such as the phonological word, can be subject to a *minimality* condition, i.e., in numerous languages no phonological word may be smaller than the prosodic foot in the language (which, in turn, is frequently bisyllabic). This restriction may override what are otherwise unobjectionable generalizations of the language; this matter is discussed at length by John McCarthy and Alan Prince in their chapter on prosodic morphology (chap. 9), and it arises in Marlys Macken's discussion of language acquisition as well (chap. 22).

The matter of establishing the phonotactics of a language can be approached by analyzing the problem into its component parts, and recognizing that different requirements or restrictions can be placed on representations at different levels in the grammar. It is helpful to bear in mind that the term *level* is used in two ways that may seem distinct, but which share a common origin. On the one hand, the traditional notion of a level derives from having a particular set of tools (syntactic categories, morphological categories, discourse categories, etc.) for analyzing each aspect of an utterance; levels of this sort (syntactic, morphological, discourse, etc.) could, in principle at least, be said to hold "simultaneously" of an utterance. On the other hand, derivational analyses of phonology posit an underlying and a derived representation, and these distinct and apparently incompatible representations are also referred to as belonging to different levels.[7]

In her discussion of underspecification (chap. 4), Donca Steriade explores some cases in which the distribution of phonological information seems to demand two distinct representations in the derivational sense. A typical case of this sort involves one set of conditions regarding where a phonological contrast can appear underlyingly, and a different set of conditions regarding where the contrast can be realized phonetically: a case, say, where nasality can be marked contrastively only on stressed vowels, but where it can be realized phonetically on any voiced segment. She argues that a two-level analysis can be replaced by a single-level analysis in which the notion of licensing is developed and applied.

In earlier versions of generative theory, considerable attention was given to analyses containing abstract segments in the underlying representation which were not part of the surface inventory of segments, that is, employing two different inventories of segments at two different levels. Much of the clamor behind the discussions of these analyses, pro and con, evaporated with the development of autosegmental analyses, in part because the reformulations as an autosegmental account removed the abstractness. That is to say, if an analysis posits a high, back, unrounded vowel that never surfaces in a language with back/front vowel harmony, that vowel is an abstract vowel even if its neighboring vowels assimilate to its [+back] specification. But if we posit a [+back] autosegment as part of a root that associates with affixes, though it fails to associate to one or more vowels in the stem, the autosegment is not abstract, since it *does* quite simply appear on the surface. Such an observation is independent of whether that analysis is factually correct; it does, in any event, cast a new light on what one means when referring to an abstract analysis.

Structure preservation is a concept that pertains to the study of phonotactics at different levels. Structure preservation is the name given to the observation that a large class of phonological rules apply in such a fashion that their output conforms to well-formedness conditions, generally understood to be well-formedness conditions on underlying structure; as a special case of this, these rules do not create any segmetn types that do not exist underlyingly (see Kiparsky 1982b, 1982c). But in view of the fact that languages generally allow a much wider range of segments on the surface than they do underlyingly, structure preservation is understood to deal with levels that are part of the lexical phonology as opposed to any post-lexical phonological level.[8]

The development of the metrical theory of stress rules, explored by René Kager in chapter 10 and by Morris Halle and William Idsardi in chapter 11, allows for an organization of stress systems in languages of the world that is simple and compact, though richer in important ways than that obtained from metrical theory as it was understood a decade earlier. The greater richness permits a more faithful description of the facts of known languages, while still remaining reasonably close to the abstract structure of earlier metrical theory. James Harris's discussion of stress in Spanish in chapter 32 illustrates the way in which current metrical theories of stress offer important resources for the analysis of languages.

The development of a more elaborate theory of phonological representations has modified our understanding of rules. Early in the development of these theories of representation the argument was made that rule formulations became simpler when we had recourse to such constructs as tiers and syllables, and the simplicity was taken as evidence that these rules represented true linguistic generalizations. Over time, however, the extent of what we could say about what constitutes a well-formed representation grew to the point where the actual rules needed to achieve the well-formed representations grew simpler and simpler; eventually, in a good number of cases, the rules became trivial (of the form "add an association line"), and perhaps no longer what we would want to call rules – that is, particular packages of language-particular information regarding how to shift a representation toward well-formedness – but rather very general principles that would be found in many, many other languages,[9] leading some to question the existence of phonological rules altogether.[10]

A number of different perspectives can be found in the field today. The effort to specify generalizations that are (more or less) true on the surface, and to use these generalizations to simplify language-particular rule formulation,[11] was inaugerated in Sommerstein (1974), echoing the intention of Kisseberth (1970). In some cases, the addition of a group of phonotactics is presumed to lead to a simpler overall grammar because the return on the rule simplification is great compared to the small (formal) cost associated with adding some simple phonotactic statements. This trade-off is motivated for two kinds of cases that seem on the face of it to be closely related: that of rules that fail to apply if their output violates a phonotactic of the language, and that of rules that only apply to structures that violate a phonotactic and in such a way as to create an output structure that does not violate that phonotactic.[12]

A more radical step is taken (and it is one which merges the two cases of phonotactic-driven rule application just mentioned) when the proposal is adopted that the well-formedness of a representation is a scalar (indeed, a quantifiable) notion, not simply a matter of *yes* and *no*. This allows one to propose that a rule applies if and only if its effect is to increase the well-formedness of the representation – or to put the matter in a less dynamic fashion, the rule's effect comes into play if and only if the well-formedness of the output is greater than that of the input. This is the proposal advanced in Goldsmith (1993)[13] under the rubric of harmonic phonology. Other phonologists have explored similar frameworks, emphasizing the hierarchization – that is, the relative violability – of constraints on a language-particular and, in some cases, a universal basis. Noteworthy in this context are Singh (1987), Paradis (1988), and especially the discussion in LaCharité and Paradis (1993) comparing several approaches.[14] A still more radical proposal, that made by Prince and Smolensky (1993) and McCarthy and Prince (1993) under the rubric of optimality theory, discussed in chapter 9 below, places such an emphasis on the character of the output representation that there is no significant role played by the notion of the rule. This optimality approach views the relationship between the input representation (or underlying representation) and the selected output representation as being not subject to language particular consideration except for the relative ranking of the universally given constraints. Associated with each input representation is an extremely large class of candidate output representations, a class so large as to include virtually any representation that may be obtained from the input representation by adding – randomly, so to speak – formal symbols of any kind. Filters are used then to evaluate candidate outputs, and a simple procedure selects the optimal one.

2 Alternations

How do we account for alternations of morphemes in distinct morphological and phonological contexts? The reader of this volume will find none of the contributors worrying about the possibility that phonological theory has been forced to deal with empirical problems that ought rather be taken care of by morphological theory, and this lack of worry is surprising, perhaps, in view of the degree of concern expressed on this score in a survey of phonological theory published about fifteen years ago (Dinnsen 1979). The motto "minimize allomorphy!" remains today's watchword, in the sense that in practice, morphology continues to be concerned with the linear order and constituent structure of words, and with making a choice of which morphemes are to be employed to realize a set of morphosyntactic features in a given sentence; but contextually determined variations in the

realization of a given morpheme will to the extent possible be accounted for phonologically.[15]

An important tradition in phonological theory associated directly with generative phonology is the search for formalisms that allow the statement of rules of assimilation were no simpler in form than rules of dissimilation, and an immediate benefit resulted from adopting autosegmental notation for assimilations, in that the addition of an association line could be easily isolated as one of the characteristics of especially simple phonological rules. The study of feature geometry, discussed in detail by G. N. Clements and Elizabeth Hume in chapter 7, is based in part on a commitment to the principle that the simple character of partial assimilations (assimilation in which more than one, but not all, the features of a segment are involved) is due to a fact about the way in which features are organized representationally. Their commitment to the principle that "rules perform single operations only" leads them ineluctably to the conclusion that an operation that seems to assimilate several features simultaneously is operating on a feature constituent node that in turn dominates a set of features. Thus this innovation can be viewed as an innovation both in our theory of representations and in our theory of rules.

Another striking development in recent work that combines the nature of phonological representation with the treatment of alternations is prosodic morphology, discussed in depth by John McCarthy and Alan Prince in chapter 9. As they observe, prosodic morphology has as its domain a range of processes which fuse together two facets of language that linguistic theory has often attempted to keep apart, phonology and morphology, for in these cases, a morphological process makes crucial reference to one or another prosodic – hence phonological – category.

One of the goals of the development of lexical phonology and related approaches is to elaborate the phonology sufficiently to allow it to deal with the alternations in purely phonological terms. But it is nonetheless all too often the case that different phonological results emerge from what appears phonologically to be the same material (for example, in many Bantu languages, a sequence of a low vowel followed by a high vowel merges to become a mid vowel if the first is in a verb root, whereas elsewhere the vowel quality of the first vowel is lost; similarly, in English, some instances of the unstressed vowel indicated orthographically by *-y* cause a softening of a stem-final consonant [*president, presidency*], while others do not [*flint, flinty*]). K. P. Mohanan explores in chapter 2 the vicissitudes of attempts to analyze these differences solely in terms of derivational levels within the phonology, suggesting that the difficulties encountered in such attempts are likely to be insurmountable.

The development of the theory of lexical phonology brought new life to a traditional question in phonological theory: to what extent can the phonological changes associated with allomorphy be reduced to statements about the phonotactics of the phonological stem and the phonological word? Lexical phonology takes a strong position, essentially identifying phonotactics (or something close to them) with the phonological rules that are responsible for allomorphy. It does this with finesse, to be sure. Post-lexical rules are located in a separate component, sufficient unto itself, and the remaining lexical phonological rules are distributed to the various strata that compose the lexical phonology. These lexical phonological rules[16] apply under two sets of conditions, and thus serve two different functions: they apply to fill in phonological specifications that have been left unspecified underlyingly because of under-specification considerations; we can expect in general that a good deal more than half of the distinctive features[17] would be left unspecified underlyingly in the derivation of a word, and these features will be filled in by lexical rules. But equally importantly (and in practical terms, more importantly), these rules will apply in a structure-changing fashion to modify phonological specifications when they apply across a morpheme boundary, and it is this latter class of modifications that forms what is traditionally understood as instances of alternation. The relationship between these styles of functioning is discussed at length by Jennifer Cole in chapter 3, and it is this relationship, when added to the representational theory of underspecification (discussed below and in chapter 4 by Donca Steriade), that yields a particular theory of alternations, a theory of the phonologicization of alternations, in effect.

3 Contrasts

Phonologists find it crucial to be able to represent differences of sound that can be used in a language to distinguish distinct lexical items or distinct grammatical items and categories. It is

necessary to say that the differences of sound are used to refer to either distinct lexical *or* distinct grammatical items because not *all* differences need be distinguished in the formalism – or so traditional wisdom has had it (this, indeed, is the fundamental insight of phonemic theory). *Some* differences, that is, may perfectly well be part of a person's linguistic knowledge (using that term in a pretheoretic way), but fail to satisfy the criterion of being relevant to lexical or grammatical information. For example, while I may tacitly know that imitating an Italian accent involves placing a falling tone on all accented syllables, or that an urban New York accent involves affricating my *t*'s before non-high vowels, this knowledge does not contribute to distinguishing any lexical items, nor any grammatical items, and thus does not enter into a strictly phonological account (though it is linguistic knowledge, and it involves knowledge of sound systems).

In classical generative phonology, all contrastive information was reduced to distinctions involving features; in fact, all contrasts could be reduced to the differences between $+F_i$ and $-F_i$ for some finite set of features F_i. All this has changed in the years since. The advent of lexical phonology (discussed by K. P. Mohanan in chapter 2, Jennifer Cole in chapter 3, and Paul Kiparsky in chapter 21 below) in the early 1980s (see Kiparsky 1982) brought renewed interest in the reasons for which not all phonological differences are equal.

Let us review this problem, which can be expressed simply as the question of how we should treat the difference between two sounds (or phones), *x* and *y*, in the phonology of a particular language.[18] The simplest situation we might find is that *x* and *y* are allophones in complementary distribution or in free variation. In the former case, we find that there is no phonetic environment in which both *x* and *y* appear, while in the latter *x* and *y* may occur freely with no lexical or grammatical difference involved. While there may be a sociolinguistic difference noted in the use of *x* and *y*, either may be used in any context in which the other is permitted, and in these two cases, we have two phones which are allophones of a single phoneme, in structuralist terminology. In the terminology of lexical phonology, the difference between *x* and *y* is post-lexical, and the difference plays no role in the lexical phonology.

A slightly more complex situation involves cases in which the phones *x* and *y* are in free variation in a certain context, and in complementary distribution in all the other contexts in which both appear. Once again, *x* and *y* would be treated as allophones of a single phoneme in traditional structuralist accounts, or as related ultimately by post-lexical operations in the context of lexical phonology.

At the opposite end of the continuum from these cases of allophones of a given phoneme, we find the case where *x* and *y* are everywhere in contrast – that is, in every phonetic context in which *x* may be found, *y* may also be found, but in a fashion that produces a word that is grammatically or lexically distinct; and – again, in the simplest case – *x* and *y* differ only by the specification of a single feature, F. The contrast between [t] and [d] in English illustrates this case well, and this is a difference that plays a central role in the lexical phonology. But in between these two extreme cases – phones being allophones of a phoneme, and phones being in contrast – loom the more difficult cases. There are at least three sorts of cases that will be familiar to anyone who has worked on the phonology of a natural language; I will refer to these as the *just barely contrastive* situation, the *not-yet-integrated semi-contrastive* situation, and the *modest asymmetry* situation. Let us consider each of these in turn.

Just barely contrastive sounds: *x* and *y* are phonetically similar, and in complementary distribution over a wide range of the language, but there is a phonological context in which the two sounds are distinct and may express a contrast. A typical example of this is the case of the "tense" A/lax æ in many forms of American English, discussed by Paul Kiparsky in chapter 21.[19] These sounds are largely in complementary distribution, and in free variation in a smaller set of words. In this writer's speech (and simplifying slightly), *A* occurs before tautosyllabic *m, n*, and *b*, and æ elsewhere: we find *sAm, pAn*, but *sæng* (and not **sAng*). However, one sharp contrast exists: the noun *can* with tense *A* [kAn], and the modal verb *can* with "lax" æ. In such cases, it is certainly not true that "anything goes"; a novel word with a lax æ before a tautosyllabic *n*, for example, seems quite impossible, and pronunciations such as *mæn, sæm* seem quite impossible, too. Thus while a contrast exists, a stubborn one which will not disappear under scrutiny, the contrast occurs in an extremely small part of the range of contexts in which the sound is found. The contrast is a lexical one, but only just barely.

Not-yet-integrated semi-contrasts: In this case, a language has two sounds, *x* and *y*, which may well be in contrast in some environments, but which in a particular environment show a sharp asymmetry, in that *x* appears in large numbers, while *y* appears in small numbers in words that are recent and transparent borrowings. In that environment, the contrast may be one that is being built up (perhaps through language contact) or worn down (through the vicissitudes of analogy and grammar simplification). Junko Itô and Armin Mester cite examples from several languages in chapter 29, discussing material from Japanese in detail, and Jerzy Rubach discusses two closely related cases in Russian and Polish, in chapter 31. In English, we might place in this category the contrast between *s* and š, used word-initially before a consonant other than *r*, as in words like *stick, shtick, sick, Schick*; while *shtick* is a possible English word (indeed, an existing English word), it remains in the periphery of the contemporary phonology of the language, and is for now a transparent borrowing.

Modest asymmetry cases: This involves pairs of sounds, *x* and *y*, which are uncontroversially distinct, contrastive segments in the underlying inventory, but for which in at least one context there seems to be a striking asymmetry in the distribution of the segments, judging by the relative number of words with the one and words with the other, or by some other criterion. The clearest examples of this modest asymmetry are the traditional cases of distribution which involve neutralization and which motivated the class of archiphonemes.

Less clear is the case of a contrast such as vowel length in English in the context of what has come to be known as Trisyllabic Shortening, that is, before two syllables, of which the first must be unstressed. It has long been noted that there are alternations involving vowel length differences in such pairs as *divine/divinity*, and this is no doubt closely related to the fact that even when we look at morphologically simple words, there are many more cases in which a short vowel appears in the context (__σ σ), like *Canada*, than there are in which a long vowel appears in such a context, like *nightingale, stevedore*, or *Oberon*. As we noted above, within the framework of lexical phonology, lexical rules that apply in a derived environment also apply in a nonderived environment as rules that specify the default or expected value of a feature in that environment; again, within the context of lexical phonology, any rule that functions in such a way would be considered in this broad category of "modestly asymmetric" relations between segments.

These five kinds of contrast naturally form a cline of the following sort:

1 Contrastive segments

2 Modest asymmetry case

3 Not-yet-integrated semi-contrasts

4 Just barely contrastive

5 Allophones in complementary distribution

The sense in which these form a cline – a single dimension of variation – is that the phonological system exerts varying amounts of force on the specification of the feature F (the feature that distinguishes the segments *x* and *y* in question). At the bottom, in (5), the specification of F is determined entirely by the grammar – by the essentially unbreakable regulation of the rule of allophony, and thus it is the grammar (as opposed to the specific lexical item) that has complete control over the feature specification. As we turn to the higher cases, the principles of the grammar place weaker and weaker restrictions on F's specification. In (4), the principles regarding the distribution of the values of feature F are almost unbreakable, but in a relatively small phonological and/or lexical class, a phonological contrast exists. In (3), an undeniable contrast exists, but it is one that acts like it is at the limit of what is permissible, and there is a strong tendency to push the one feature value to the other value, so as not to offend the native sensibilities of the language in question. When we turn to (2), we find a featural difference in which some asymmetries may be seen – for example, in statistical proportions – but nothing else suggests an asymmetry between the values of F, and finally, in (1), there are no perceptible differences between the functioning of the two values of the feature F, and there is no pressure, synchronic or diachronic, giving rise to a preference for one feature value or the other.

All of these cases are familiar and important to the phonologist. How are they distinguished from

each other in current phonological theory? By and large, it is the case of "allophones in complementary distribution" that is kept qualitatively distinct from the other four, while among the other four cases (contrastive, modestly asymmetrical, not-yet-integrated semi-contrasts, and just barely contrastive) the differences are largely minimized, and all are treated as "lexical contrasts."

3.1 Underspecification Theory

The questions we have been discussing have been addressed in recent years as in large measure a matter of phonological *representation*. Underspecification theory, the subject of chapter 4 by Donca Steriade, has been concerned with determining whether, and to what extent, feature distinctions should appear in a phonological representation not as a choice between +F and −F, but rather as a choice between +F and no marking at all. In her contribution here and elsewhere, Steriade is at pains to distinguish what she calls trivial (or inherent, or permanent) underspecification from nontrivial underspecification; if a feature is allowed to take on only one value (+, say) at every level in the grammar (and this sort of case is not controversial), its underspecified character is trivial (though determining that this is the case may not be an easy matter). Such features are called privative, or monovalent, and the consequences for the treatment of phonological representations of such features is discussed in detail by Colin Ewen in chapter 17 and by Sanford Schane in chapter 18. Perhaps the most fruitful area of research on this issue has been that of vowel harmony systems, where asymmetries between the behavior of the two values of the harmonic feature have been studied in detail, as Harry van der Hulst and Jeroen van de Weijer report in chapter 14. Only if the feature is restricted to a single value at a deep level of representation, and finds itself expanded to two values at a more superficial level, is a feature nontrivially underspecified, in Steriade's terminology.

A great deal of work has been invested over the past ten years regarding the fashion in which underlying specifications of featural differences and contrasts are to be represented formally, with an emphasis on whether and when a feature must be treated as a privative, monovalent feature. The reader will find several different positions taken regarding radical underspecification theory in the chapters that follow. Paul Kiparsky offers the strongest defense, while Donca Steriade presents a compelling case against it. Let us consider some of the issues that are involved.

There are questions that lie behind this disagreement that go even beyond the character of underspecification theory. Kiparsky's investigation of featural underspecification brings one to a conclusion that we might summarize in the following way: what we have traditionally referred to as "lexical diffusion" is nothing more nor less than our observation of the grammar of the language attempting to balance off the relative complexity of rule statements and of lexical entries, the two major components of the grammar. The argument is subtle, and worth reviewing in some detail.

Kiparsky proposes, first of all and in line with widely understood principles of lexical phonology, that in any particular phonological context, there will be, for any given feature F, an expected, or unmarked, value; this will be either +F or -F. Whether the unmarked value is +F or -F depends very much on the phonological context C, and if the unmarked value of F in context C is +F, then the marked value will be -F. How are the decisions made as to what is the appropriate context to consider, and which value should be chosen to be the unmarked value? By and large, radical underspecification (and Kiparsky's position is squarely in this camp) has left this an open question, but it points in two directions for the eventual answer. The first way to explain the choice of unmarked values is to appeal to universal principles, such as the observation that the unmarked value for the feature Voice is + with no context specified, while the unmarked value for the same feature on obstruents is -. Such observations have served in the first place as phonological universals, and may now serve in the breach as universal default feature markedness principles. The second way, which is much more important for our purposes, is to look at the relative complexity of the information packed into the rule component and the lexical component of the grammar.

Let us recall that from the traditional generative point of view, a phonological grammar is divided into two major components: a set of rules, broadly construed, and a lexicon. The lexicon contains (among other things) the underlying phonological representation of all of the words of the language. The traditional generative concern has been to select the least complex grammar consistent with the data of the language. However, we – linguists – must constantly bear in mind that there is typically a trade-off between the complexity of the rule component, on the one hand, and the lexicon, on the other, in the sense that the same set of data can be treated in several different ways. Some will show greater

complexity in the rule component and less complexity in the lexicon, and others greater complexity in the lexicon and less complexity in the rule component. Let us briefly review how this can arise, and what the consequences are for markedness and markedness reversal.

Imagine that we found a language much like English, in which there was a category of words relevant to the phonology (let us call the category "Proper-Nouns," for purely expository purposes; in a real case, it might be nouns, prepositions, or what have you) in which far more words began with the voiced alveopalatal affricate ǰ than with its voiceless counterpart č. In our example, this would mean that the language contained more ProperNouns like *John, Jim, Jerry, James*, and *Geoff* than it did like *Chuck* or *Charles*. While on general phonological grounds as we noted just above, č is unmarked for the feature Voice (it is a voiceless obstruent, after all), the fact is that there are more ǰs in this particular morphophonological environment. In traditional generative terms (and this conception is unambiguously embraced by radical underspecification theory), the measure of complexity of a representation is just the set of actual marks required to specify the underlying representation and up until now, it has served us well to mark "+" on the feature Voice when we encountered a ǰ or any other voiced obstruent, leaving the "–" off when we encountered a voiceless obstruent. But in this particular context – word-initially in ProperNouns – this economy serves us badly; we would be better off putting down a – on the feature Voice at the beginning of *Chuck* and *Charles*, and leaving the voiced ǰs unmarked. In order to achieve this economy *in the lexicon*, however, we must set down a rule in the rule component of the form:

(1)

$$\begin{bmatrix} \text{uVoice} \\ \text{+Strident} \\ \text{+Delayed Release} \end{bmatrix} \rightarrow [\text{+Voice}]/\# \quad \begin{bmatrix} \text{+Coronal} \\ \text{——} \end{bmatrix}_{\text{ProperNoun}}$$

and this rule will, of course, cost the rule component something. Thus we find a trade-off between rules and representations, and we (or the grammar) will choose to reverse the markedness of a feature just in case there is a good enough trade-off involved – for that is what we have just done; we have said that the markedness relationship flipped about in a particular morphophonological context (ProperNoun-initially). When is the trade-off *good enough*? Even phonologists committed to this perspective have not the slightest idea; perhaps one must count up the number of formal symbols in the new rule that must be financed (in the case above, the rule seems to cost about 11 units, counting (i) u (ii) Voice (iii) + (iv) strident (v) + (vi) Delayed Release (vii) + (viii) Voice, (ix) ProperNoun (x) +, and (xi) coronal). How many features must we be able to save in the lexicon to pay off the price of the rule? Perhaps eleven; perhaps one hundred eleven. We do not know.

Kiparsky's contention is that what we call *lexical diffusion* consists, first of all, of a dynamic in which the rule component pushes each phonological rule in the direction of generalizing the context in which it applies. If generalizing the context consists of leaving off formal symbols in the rule, then this makes the rule simpler and less costly; but if the language is to generate the same output after "simplifying" – i.e., generalizing – one of its rules, then it must add additional specifications to a new set of lexical items. Which lexical items? – those that satisfy the generalized form of the rule but that did not satisfy the rule in its earlier form. That is, around the *core* set of words to which Rule 1 applies, there is what we might call a *penumbra* of words that do not satisfy Rule 1, but would satisfy it if Rule 1 were generalized to Rule 2, according to which the markedness reversal holds not only for affricates but for fricatives as well.

(2)

$$\begin{bmatrix} \text{uVoice} \\ \text{+Strident} \end{bmatrix} \rightarrow [\text{+Voice}]/\# \quad \begin{bmatrix} \text{+Coronal} \\ \text{——} \end{bmatrix}_{\text{ProperNoun}}$$

In the earlier stage, when Rule 1 was present, the initial consonant in words like *John* was unmarked for Voice, while the initial consonant in *Shane* and *Sam* was unmarked for Voice and the initial consonant in *Zachary* was marked for Voice. However, if Rule 1 generalizes (simplifying the rule component) to become Rule 2, words like both *Shane* and *Zachary* will now move into the core of the rule, and the marking on their initial consonants will be reversed (*Shane's* is now marked, and *Zachary's* is unmarked). This does indeed sound like a move that will buy simplicity for the rule component, but it will cost the lexicon dearly in terms of markings on the feature Voice. Although this is not remarked upon, it is crucial to an account such as Kiparsky's that such an extension (cheap for the rule component, costly for the lexicon) will actually occur, and not infrequently. Once it has occurred, though, the lexicon is not only free to change, it is encouraged to do so, by the economics of markedness, and all of the words that were in the penumbra of the older Rule 1 will now be pushed toward dropping their newfound Voice markings; they will drop their new markings, and become unmarked with respect to the new, generalized Rule 2 (and hence will become voiced consonants word-initially). As they do this on a lexical element by lexical element basis, we will observe the change, and we will call it the lexical diffusion of the rule of ProperNoun-initial voicing.

Paul Kiparsky makes a strong case that this is a reasonable understanding of the process of linguistic change known as lexical diffusion. Looked at more closely as an account of underspecification, however, it seems to me that in certain respects the argument is too weak, while in others it is too strong, though the insight at the center regarding the relationship of synchronic phonological theory and our understanding of diachronic processes remains attractive and even compelling.

The insight at the center is that lexical diffusion is a set of leap-frogging simplifications, first in the rule component, and then in the lexicon, each driven by the urge toward total simplification of the grammar. The sense in which the argument is too weak is that this view in no way supports the view of radical underspecification employed in the description given above. That is, radical underspecification insists that in the underlying representations, even among the phonologically contrastive features of the language, there will be far less than fifty percent of the features in the underlying representations actually specified, because in any given environment, one feature value will always be the unmarked value, and that value will not be explicitly filled in underlyingly, but rather will become filled in during the derivation. Radical underspecification lends itself very naturally to the hyperformal understanding of "naturalness" according to which a representation is more natural or simple (or preferred, or less marked) if it has fewer formal symbols.

Radical underspecification theory assumes the presence of a formal device (let us call it *D*) which determines for each position, and for each lexical entry, which feature specifications may, and thus must (for that is the logic of underspecificationism), be left unspecified. Intuitively speaking, *D* determines (given the phonological rules of the language and univeral default rules) which feature specification(s) may be left out of a given underlying representation and still return the correct surface form. Radical underspecification theory uses *D*'s computation to leave those feature specifications out of the underlying representation. Once they are left out, they cannot and do not add anything to the total complexity of the word's lexical entry.

Once we make explicit what is going on, though, it should be obvious that radical underspecification – the requirement to leave these redundant features unspecified in underlying representation – is by no means the only way of getting the result that we were looking for above. The result that we need is that the more features there are underlyingly that are "unmarked" – that is, the result of patterns of expectation – the "simpler" the representation is underlyingly. But this can be achieved with device *D* and *no* theory of underspecification. That is, given a fully specified underlying representation, let device *D* be run, and let it paint *blue* all of the feature values that are unmarked values in their context; now we may define the complexity of that underlying representation as the sum total of all the markings that device *D* did not paint blue. It would thus be an error, I believe, to think that fundamental concerns with simplicity commit one (or even lead one) to underspecification.

But there is also, as I suggested, a sense in which Kiparsky's argument is too strong, a sense in which if the argument were valid as it stands, then it would equally support conclusions that are palpably false. The problem is this: Kiparsky's contention is that there is pressure (due to the drive for simplification of the grammar as a whole) to shift lexical features from their marked to their unmarked values, regardless of the feature.

But if the pressure to shift from marked to unmarked values in the cases which Kiparsky explores (the *just barely contrastive* sorts of cases, as I have labeled them) derives in essence from the drive to simplify grammars by simplifying representations, there are no grounds for distinguishing among the types of features that will feel this pressure, and the theory will then predict that all features will equally feel the pressure to conform, to take on the unmarked value. But this is never found in the case of what I called above *fully contrastive* features, and imagining what would happen if it were makes this point clear. Take, for example, the difference in English between the pairs of obstruents *d* and *t*, and *g* and *k*. The voiced obstruents bear, we may agree, the marked value of the feature Voice, while the voiceless ones bear the unmarked value. If there were pressure on the feature Voice to take on the unmarked value, then *dog* [dog] would be in danger of shifting lexically to *talk* [tok], with *dawk* or *tawg* as possible intermediate stages in this shift. But this prediction is not borne out by the facts: plainly, for a healthy center of our lexicon, there is no evidence of pressure of any sort at all to change one segment to another, despite the fact that radical underspecification theory, used as a tool to account for lexical diffusion and change, *treats fully contrastive pairs of segments in the same way as barely contrastive pairs of sounds*.

The conclusion that we must draw from this, it seems to me, is that we have not yet reached a satisfactory understanding of the nature of the binary contrasts that are found throughout phonology. Neither marking all contrasts as +/− nor marking all contrasts as +/Ø is sufficient; we need some more refined means for linking the notion of specification. The pressure to shift may well exist for contrasts of one or more of the categories we discussed above, but the pressure is not found in *all* of the categories. We need to find, at a minimum, a way to recognize that some markedness distinctions carry with them the impetus to eliminate the marked value: for example, in the case of the not-yet-integrated semi-contrasts, we do expect the marked feature value to give way to the unmarked feature value. But this is not the case in fully contrastive pairs of segments (like the *t/d* of English). We must determine how the line is drawn (if there is a clearcut line to be drawn) between these cases, and we must find a way that the two cases (or the several cases) can be dealt with without wrongly assimilating the one case to the other.

A careful study of a language will often raise some questions as to just which phonological differences are contrastive, in various parts of the vocabulary. I touched on this briefly just above, suggesting that there is a cline to be found between differences that are allophonic and those that are full-fledged distinctions in a language, but other kinds of divergence can be found, as Junko Itô and Armin Mester observe in chapter 29. They point out that a fourway distinction between Yamato, Sino-Japanese, mimetic, and foreign forms is observationally motivated; the question then arises as to how these differences are to be organized in the phonology. Itô and Mester argue convincingly that if some "peripheral" parts of the Japanese vocabulary should be thought of as being subject to weaker constraints than those parts lying at its "core", the difference "has nothing to do with 'early' vs. 'late' in the derivation: the periphery is just as underlying as the core." The burden is on phonological theory, then, to shift the style of analysis to allow us to separate properly the influences of distinct constraints and phonotactics in a nonderivational fashion so that we can account for the range of differences found in different parts of the vocabulary of a single language.

There is a central and important question that draws together underspecification theory and the nature of rule application and which comes up several times in the course of the chapters that follow, most notably in the chapters on vowel harmony, underspecification, and historical change; this question is whether rules can apply in a feature-changing fashion directly, or whether what appears to be a feature-changing process is composed first of feature-delinking followed by a feature-filling process. As van der Hulst and van de Weijer note, there is considerable controversy as to whether the range of known vowel harmony (and nasal harmony, we might add, too; see note 21) systems requires feature-changing analyses. The development of theories in this area is at a point where it is no longer profitable, as it might once have been, to determine that the best theory, a priori, is one which allows the most tightly constrained class of languages, for as much progress has been achieved in recent years by enriching the formalism as has been achieved by constraining it, as I observed above in connection with metrical theory.[21]

4 Theories and Languages

Not all of the chapters in this book fit neatly into the three by three schema presented at the beginning of this chapter. The study of tone languages, for example, has been extremely fruitful for the development of recent phonological theory, but it is the clarity and the complexity of the tonal systems rather than anything peculiar to them that has made their study so important. The chapters by David Odden on African tone languages (12) and Moira Yip on Asian tone languages (13) cover two of the best-studied areas of tonal systems, though regrettably other areas of the world with important tonal systems such as Meso-America and Southeast Asia are not discussed here.

Theories of syllable and skeletal structure have profited enormously from the study of Semitic languages and of Afro-Asiatic languages more generally. In their surveys of Chadic (chap. 26), Ethiopian (chap. 27), and of Semitic languages (chap. 30), Paul Newman, Grover Hudson, and Robert Hoberman, respectively, present the aspects of the phonology of these languages that are most significant to current work in phonological theory. Australian languages have played an equally important role in the development of metrical theory, and Nicholas Evans in chapter 25 provides a detailed account not only of the metrical structure of Australian languages, but of other aspects of their phonology which have broad consequences for phonological theory. Bernard Tranel, in his analytic overview of liaison and latent consonants in French (chap. 28), offers an up-to-date account of an important area of research for theories of skeletal positions and prosodic categories.

The study of signed languages, such as American Sign Language, promises to have a profound effect on phonological theory, and perhaps ultimately on our understanding of what a human language is. The possibilities that emerge from a linguistic system not constrained by the resources of the vocal tract exploit capacities that had until recently been hidden from linguists' view, and the broadened vista that we have today may in retrospect be as significant for the development of linguistics as was the impact on the Western tradition of the study of non-Indo-European languages. In chapter 20, Diane Brentari discusses some of the salient phonological properties of American Sign Language, itself only one of a large number of signed languages of the world, but for the moment the best studied from a linguistic point of view.

5 Conclusion

In the chapters that follow, the reader will have the opportunity to read thirtyone accounts of the ways in which current phonological theory treats the central phonological problems faced by linguists today. We have seen enormous progress over the past several decades, and it is my hope, and expectation, that this rate of progress will continue in the years to come.

1 I would point out that these questions change very slowly in time as well, but will not discuss that issue here.

2 This briefest of summaries leaves out entirely two important subjects: how hypotheses and theories are evaluated (that is, what counts as justification for claims in a given theory), and how phonological theory is linked to other theories, both within linguistics (syntax, phonetics, morphology, etc.) and without (neurobiology, perhaps theories of lexical access, of speech production and perception, though much of the second group is arguably better viewed as being part of linguistics, not external to it). My remarks in the text above should be thought of as limited to matters *internal* to phonological theory, though there is discussion of subjects external to phonology in a number of the following chapters (chapter 2 by K. P. Mohanan and chapter 9 by John McCarthy and Alan Prince on the relationship of morphology and phonology, chapter 15 by Sharon Inkelas and Draga Zec and chapter 16 by Elisabeth Selkirk on phonology and syntax, and chapter 22 on language acquisition, chapter 23 on language games, and chapter 24 on experimental phonology by Marlys Macken, Bruce Bagemihl, and John Ohala, respectively). It is truer today than ever before that the studies bridging components and methodologies are critical for the testing and refinement of phonological theory.

3 The rise in sonority between *b* and *n* is not great enough, while the rise in sonority from *b* to *l* is. See chapter 6 by Juliette Blevins below; for a different view, with an explicitly computed notion of sonority, see Goldsmith (1993a) and Goldsmith and Larson (1992).

4 English and Arabic are often-cited examples of languages in which we find longer strings of consonants word-finally than are found syllable-finally. English allows for certain strings of three consonants, as in

Pabst or *midst*, while at most two consonants can appear at the end of a syllable word-internally, as in the first syllable of *vintner* or *Thornley*. Modern Standard Arabic permits no more than one consonant in the coda of a word-internal syllable, but two consonants are permitted word-finally.

5 On this, see, for example, Bosch (1991).

6 Itô (1986, 1989), and Goldsmith (1990).

7 See Goldsmith (1993a) for an extended discussion of this, and further references. The treatment of rule ordering plays an essential role in this discussion; see chapter 19, by Gregory Iverson, in this volume.

8 Borowsky (1986) presents an analysis of English syllable structure from this point of view; Wiltshire (1992) argues for an alternative, nonderivational analysis.

9 I should perhaps state the point a bit more carefully. From a classical generative point of view, the basic elements that are used jointly to formulate rules are quite simple as well (names of features, feature values, etc.). The claim implicit in classical generative phonology is that these items in the vocabulary of rule formulation will coalesce into rules in a more or less random fashion; there are no preferred ways for the items to group together among all the conceivable ways that are syntactically well-formed, that is, obey the basic principles of the syntax of phonological rules (e.g., there is an arrow that points to the right, a slash further yet to the right that marks the beginning of the statement of the environment, only one dash marking the environment, etc.). If there were specific ways in which the elements of phonological rules preferred to come together, then counting up the total number of formal symbols would not be an adequate measure of the complexity of a rule. But that is what we find: some simple rule formulations arise frequently ("delink the first association line of a doubly-linked autosegment") while others do not occur at all frequently ("add an association line to an element that is doubly associated"). It follows that an evaluation metric for the rule component that does no more than count formal symbols in the rules is not adequate.

10 E.g., those working in government phonology (see Kaye 1990, and other papers cited there).

11 As opposed to using these generalizations to rule on what was a possible underlying form, as was done in natural generative phonology, for example (see Hooper 1979, and references cited there).

12 A widely discussed case of a process which applies in order to repair ill-formed syllable structure is epenthesis; see, for example, Ito's (1989) discussion of this. The case of rules that do not apply when their output would violate a phonotactic is discussed, for example, in McCarthy (1986b) and Yip (1988a). A typical example of this sort is the case of a vowel deletion which fails to apply if the output of the rule is two successive, identical segments, a violation of the Obligatory Contour Principle (on the OCP, see chapter 12 by David Odden). Yip makes the observation – a crucial one, in this writer's opinion – that the OCP (functioning as a measure of well-formedness condition like any other such measure) is "not an absolute rule *trigger*" (if a language has no rule to improve an OCP violation, then there's nothing to be done about it), "but it is an absolute rule *blocker*" (Yip 1988a, p. 75), for the phonology always has the option of *not* applying a rule that it contains, so to speak. In all cases, a violation of the OCP is a worse-formed structure than an otherwise parallel structure that does not violate the OCP, and phonological rules applying in a harmonic fashion (see below in the text) will only apply to improve the well-formedness of a representation.

13 And elsewhere; see Goldsmith (1993b), as well as Goldsmith (1990, chap. 6), Brentari (1990b). Bosch (1991), Wiltshire (1992). The principle of harmonic application has been summarized by the phrase, "Always change, but only for the better."

14 LaCharité and Paradis (1993) is an introductory paper in an issue of the *Canadian Journal of Linguistics* (Paradis and LaCharité 1993) devoted to constraint-based theories on multilinear phonology, where the term "constraint" is equivalent to phonotactic.

15 I do not mean to be defending this position, but simply to point out that this has served as a background working principle for a background working principle for a good deal of the work in current phonological theory.

16 Or at least those in a cyclic stratum, and not a noncylic stratum. Proposals have been made that would mark strata or even rules as noncylic, freeing them from the operative constraints on the application of cyclic rules, See Jennifer Cole's discussion in chapter 3, as well as chapter 11, by Morris Halle and William Idsardi, and also chapter 32, by james Harris.

17 More than half, because for any given deal less than 50 percent of its occurrences in underlying forms will be unspecified. It is obvious that the percentage of occurrences of a given feature will be less than 50 percent if it were greater, we could simplify the grammer by making the unmarked value the other value. Further recognition of markedness reversals, as discussed in the text, for example, will ensure that the number of occurrences of a feature that bear a specification are considerably less than 50 percent.

18 In this discussion, I focus on the issue of contrast between phones that exist in a language; but a similar question can arise with respect to the question of what the permitted segments are in the lexicon of a language. This is closely connected to the discussion above, in the sense that some existant phones in a language (say, the flap [D] in English) are taken *not* to be in the inventory of the English lexicon, because they are integrated into English phonology outside of the English lexicon, and among the "post-lexical" (nonlexical) phonological rules. But the status of segments is often unclear, in much the way in which the status of contrasts is unclear, as the hierarchy in the text below illustrates. We know that certain sounds are not part of the inventory of a given language; English has no clicks, and Ndebele does. English has no nasalized vowels, and French does. English has a velar nasal, and French does not. Or does it? Sampson (1992) reviews the case for considering the velar nasal as a segment of contemporary French (see Walker 1975 and Walter 1983 on this as well) and argues (1) that the range of cases in which n appears in French words – generally borrowings from English originally – is such that it cannot be ignored in the analysis of contemporary French; (2) that there is no simple case to be made that the n is the phonetic result of some combination of uncontroversial French segments in some particular context – say, / ng / in word-final position; and (3) that it would be an error to accept that angma has a status in the phonology of French according it a freedom of distribution anything like that of the clear cases of consonants in the language. Sampson suggests that the emergence of n along with other overt nasal consonants in French (not English!) words such as *lunch, cross-country*, etc, is best understood as forming a significant subgroup in the French lexicon. In this fashion, he suggests, rather than trying to incorporate obviously aberrant pronunciations into an otherwise homogeneous phonology, one could take into account an otherwise irregular phonological existence.

19 See Kiparsky's contribution, chapter 21, to this volume. The facts cited here reflect this author's(JG) New York dialect.

20 Proper names are notorious for stretching the boundaries of permissible sound combinations, and in this case, too, we find contrasts like *Schneider/Snider* which do not seem to faze English speakers, or to be subject to any kind of nativization.

21 See Noske 1993 on feature changing and three-valued features. In the domain of post-lexical rules, it appears to me to be established that there must be both rules that apply in a structures-changing fashion and, in other languages, rules that apply only in a structure-building (i.e., feature-filling) fashion. One example of each is given in chapter 1 of Goldsmith (1990): the former is illustrated there by San Miguel El Grande Mixtecan and the latter by Sukuma.

Cite this article

GOLDSMITH, JOHN. "Phonological Theory." *The Handbook of Phonological Theory*. Glodsmith, John A. Blackwell Publishing, 1996. Blackwell Reference Online. 31 December 2007 <http://www.blackwellreference.com/subscriber/tocnode?id=g9780631201267_chunk_g97806312012673>

Bibliographic Details

The Handbook of Phonological Theory

Edited by: John A. Glodsmith
eISBN: 9780631201267
Print publication date: 1996