Part I

Preliminaries

> The formation of different languages and of distinct species [is] . . . curiously parallel . . . We find in distinct languages striking homologies due to community of descent, and analogies due to a similar process of formation . . .
>
> Darwin (1871: 465–66)

1 What is Evolutionary Phonology?

> In language, as in nature in general, everything moves, everything is alive and changing. Baudouin de Courtenay (1897/1972)

1.1 Relating sound patterns to sound change

The field of modern linguistics is conventionally divided into distinct subfields, defined by the questions they address and analytical methods they use to answer these questions. Establishing a clear separation between the components of a synchronic description has proved particularly useful in investigating the distinctive properties of sound systems (phonology), words (morphology), and phrases (syntax). Yet the strict segregation of synchronic and historical description has a rather different character, as it is very often the case that these subfields overlap in their coverage. The post-Saussurean tradition has tended to dismiss historical explanations as entirely irrelevant to the task of synchronic description, on the grounds that the speaker of a language cannot be expected to know the history of that language and "a segment does not know where it came from" (Lass 1984: 178). This essentially ahistorical perspective leads to considerable redundancy, as numerous commentators have observed, since many patterns with a well-understood historical basis or origin must be reencoded in synchronic accounts. The fact that such patterns typically lack system-internal motivation within the synchronic grammar has likewise expanded the inventory of "universals" that must be attributed to a linguistic system because they cannot in any way be deduced or inferred from other properties. The predictable effect has been a general extension of synchronic descriptions and mechanisms to encompass nearly all patterns and generalizations within a linguistic system, irrespective of their status or origin.

A central claim of the present work is that this "one size fits all" methodology invariably fails to explain – and often fails even to describe accurately – many of the sound patterns that recur in the world's languages. On the other hand, phonetically motivated accounts of the origin and

development of common sound changes provides the basis for a genuine explanation. This type of explanation offers perceptual or articulatory motivations for recurrent sound patterns in place of *a priori* claims that these patterns arise because they are intrinsic to the organization of the synchronic system.

The overlap between the subject matter of phonology and historical linguistics is particularly striking. Phonology is the study of sound patterns, their nature and use. In attempting to model sound patterns, linguists have studied the phonological systems of most of the world's languages and language families. This large corpus of descriptive studies has served as the primary database for work within phonological typology. Typologists have spent centuries cataloguing recurrent and unattested sound patterns across independent languages, language families, language areas, and language types. These sound patterns fall into many different categories. Among these we find studies of segment inventories, segment sequences, syllable types, stress patterns, tone patterns, feature distribution, and studies of common alternation types. Alternations between sounds in related words have been categorized as instances of assimilation, dissimilation, deletion, insertion, metathesis, coalescence, breaking, lenition, fortition, and neutralization, with further subdivisions within each of these.[1]

Historical linguistics is the study of language change and relationships among languages. In attempting to model patterns of change, linguists have studied sound change in many of the world's languages and language families. Sound changes fall into many different categories. It is noteworthy that the majority of commonly attested sound changes in the world's languages are mirrored by synchronic alternations of precisely the same type. Sound changes give rise to changes in segment inventories, segment sequences, syllable types, stress patterns, tone patterns, and feature distribution. Moreover, these sound changes are, precisely like synchronic alternations, categorized as instances of assimilation, dissimilation, deletion, insertion, metathesis, coalescence, breaking, lenition, fortition, and neutralization. At a much finer level of detail, the most common sound changes and the most common types of synchronic alternations are nearly coextensive. The many common sound changes which have direct parallels in synchronic sound patterns include velar palatalization, final obstruent devoicing, vowel nasalization before nasal consonants, and place assimilation of nasals to following oral stops.

[1] Phonology has been extended to the study of signed languages used by the Deaf, where visual image patterns replace sound patterns as the object of study. For an overview, of current issues in sign language phonology, see Brentari (1995). A brief discussion of the relevance of sign language phonology to evolutionary approaches can be found in 11.1.

The pervasive parallels between common types of sound change and common synchronic alternations, has suggested to many, notably the neogrammarians, that synchronic sound patterns are a direct reflection of their diachronic origins, and, more specifically, that regular phonetically based sound change is the common source of recurrent sound patterns. Evolutionary Phonology investigates this hypothesis and explores its consequences for phonological theory and models of sound change. However, before elaborating this approach, it is worth clarifying a guiding methodological principle.

All else being equal, simpler grammatical models are usually preferred to more complex ones. More specifically, any model which duplicates explanations, within or across domains, is in some basic sense more complex than one which does not. Hence, if we can demonstrate that principled diachronic explanations exist for particular sound patterns, considerations of simplicity would seem to dictate that explanations for the same phenomena should not be imported into, or otherwise duplicated within, synchronic accounts. In all cases where clear diachronic explanations exist for a particular synchronic pattern, this diachronic explanation makes a synchronic account redundant, since the optimal description should not account for the same pattern twice. To take just one concrete example, if it can be shown that final obstruent devoicing is a common sound change with a principled phonetic explanation, then we can recognize many synchronic prohibitions against voiced obstruents in final position as the direct result of this common sound change (see chapter 4). The resulting patterns of voiced versus voiceless obstruents must be describable within synchronic grammars, but the grammars do not need to explain the absence of voiced obstruents in final position. Any cross-linguistic statements of this sort merely duplicate an explanation which already exists independently in the diachronic domain. A central premise of Evolutionary Phonology, then, is that principled diachronic explanations for sound patterns replace, rather than complement, synchronic explanations, unless independent evidence demonstrates, beyond reasonable doubt, that a separate synchronic account is warranted.

The idea that common sound patterns reflect common sound changes, and that historical explanations have priority over synchronic ones, are views that have been expressed many times in the history of linguistics. In the study of sound change, the neogrammarians were renowned for their belief that the formal and functional status of an element within a synchronic system could be understood only in terms of its diachronic origins (see chapter 3). Though the neogrammarian view is no longer quite so widely assumed as a consequence of Saussure's reorientation of the field, it is nonetheless endorsed by many general grammarians, historical

linguists, typologists, phonologists, and phoneticians of the twentieth century. Consider, for example, Jesperson's (1924) position that to truly understand a linguistic system "we must know how it came to be." A similar position is taken by Greenberg in his discussion of phonological typology. He suggests that diachronic processes explain frequently occurring sound patterns, and stresses the complementary nature of diachronic and synchronic considerations in characterizing universals of sound patterns (Greenberg 1966a, 1978). Sampson (1970: 618–19) likewise questions the status of morphophonemic rules in synchronic grammars, and the general overlap of historical and synchronic explanations:

> If morphophonemic rules may indeed be thought of as reconstructions of history, we are then free to ask whether this part of the phonological component, as distinct from the MS [morpheme structure constraints: JB] or phonotactic rules, has any place in a description of linguistic competence . . . the latter could simply be interpreted as universal constraints on the nature of possible sound-changes; and as for the former, if regularity of alternations may be explained as due to a historical sound-change having applied to a situation not exhibiting alternation in the relevant respect, *there is certainly no need to give a second explanation of the same facts in terms of synchronic linguistic competence*. (emphasis added: JB)

In modern phonological theory, the precise locus of explanation has been variable (see chapter 3). Nevertheless, in one of the few careful evaluations of the role of extra-grammatical factors in shaping synchronic sound systems, Anderson (1981: 497) suggests that "we can only determine that some property is to be attributed to the essential nature of language *if it does not seem to have an account in more general terms*" (emphasis added: JB). In studies of the phonetic basis of sound change, it is taken for granted that the majority of recurrent sound patterns can be explained in terms of phonetically natural processes. Consider, for example, the introductory remarks to Hombert et al.'s (1979) detailed phonetic study of tonogenesis: "Sound changes or sound patterns that are attested in diverse, widely-separated languages cry out for an explanation by what is common to all speech communities: the physical apparatus which humans use to produce and perceive speech. One such sound change that reveals many striking common patterns is the development of tone."

Despite the consensus that frequent sound patterns reflect common phonetically motivated sound change, there is no single work which demonstrates this point comprehensively and in detail. This gap in the literature is noted by Ferguson (1990: 59–60), in his essay on one particularly common sound change, that of $s > h$:

> One of the most powerful tools in the armamentarium of linguists engaged in the study of diachronic phonology is the often implicit notion that some changes

are phonetically more likely than others. Thus if a linguist finds a systematic correspondence between [g] and [dʒ] in two related language varieties, it will be reasonable to assume that the stop is the older variant and the affricate the younger one until strong counter evidence is found. The linguist makes such an assumption because experience with many languages has shown that the change of [g] to [dʒ] is fairly common and tends to occur under certain well-documented conditions whereas the reverse change is unusual and problematic. This line of argumentation has been employed, either explicitly or implicitly, since the earliest days of modern historical linguistics.

Because of the importance of this methodological tool, one might expect that general treatises and introductory textbooks on historical linguistics would devote considerable space to a presentation of the relative probabilities of various possible sound changes, as well as explanatory factors accounting for them. Also, because of the centrality of alternations and processes to the field of phonological theory, one might expect that general treatises and introductory textbooks in phonology would devote considerable space to this topic. Unfortunately, authors of books on historical linguistics or phonological theory have a great deal of other ground to cover, and this simple but important concern tends to be neglected.

Evolutionary Phonology addresses itself directly to this basic but central concern. This study fills a gap in the literature by providing a sustained argument demonstrating that a broad range of phonological phenomena can be explained in terms of common phonetically motivated sound change. Evolutionary Phonology constitutes a concrete and comprehensive attempt to explain the majority of the world's recurrent sound patterns in terms of well-understood instances of phonetically motivated sound change. As a concrete model, it incorporates current models of articulatory phonetics, speech perception, and language acquisition. As a comprehensive model, it summarizes a great deal of work in experimental phonetics, typology, variation, and theoretical phonology, and relates this to centuries of work modeling sound change and sound patterns. As an explanatory model, it locates the domain of explanation for many recurrent synchronic patterns in the diachronic dimension.

1.2 A formal model of sound change

One important component of Evolutionary Phonology is a model of phonetically based sound change which is broad enough to handle all the various types of attested change, and constrained enough to rule out unattested changes. A precise model of sound change is proposed in chapter 2, and incorporates two observations regarding human language which should be fairly uncontroversial. First, *all spoken language is characterized by a wide range of phonetic variation, some of which is language specific, and some of which is determined by physical properties of the*

human vocal apparatus. One dimension of this continuum of variation is that determined by careful versus casual speech, a dimension sometimes quantified in terms of articulatory effort. Another dimension of this continuum is fast versus slow speech, which may involve articulatory compression or expansion along the temporal dimension. Other relevant dimensions of variation in this model include frequencies of particular phonetic variants: frequencies are expected to be highly variable across speakers and across time, since they are determined by a wide range of factors including lexical frequency, frequency of contexts involving careful versus casual forms, and social variables (age, gender, class, etc.) associated with particular phonetic variants. A second observation is that, though language transmission from one generation to the next is constrained by perceptual, articulatory, cognitive, and social factors, *language transmission is, by its very nature, indirect and imperfect*. Within this imperfect system of transmission, sound change may be viewed as the norm, not the exception. Since every individual will have slightly different early childhood experiences, every individual will, by definition, form a grammar based on distinct sets of surface forms.

The range of sound patterns investigated in part II support the general model of sound change in three respects. First, the suggested typology of sound changes with sources in misperception, ambiguous segmentation, and ambiguity due to variation is descriptively adequate. Second, where sound changes appear to defy this typology, they can be shown to have non-phonetic origins. Third, and most strikingly, the general model of sound change makes predictions regarding phonetic preconditions of change which find general support in experimental and typological studies. Implications of this particular model of sound change are explored in part III. One implication of the model is that most aspects of sound patterns constitute learned language-specific information. A corollary of the model is the regularity hypothesis: phonetically based sound change is typically regular because sound change is a subcase of normal acquisition of phonological contrasts and categories.

1.3 Types of explanation: historical, phonetic, formal, and non-teleological

This locus of explanation in Evolutionary Phonology places it at some distance from other phonological models, where explanation is attributed directly to synchronic principles. The working hypothesis supported throughout this volume is that *recurrent synchronic sound patterns have their origins in recurrent phonetically motivated sound change*. As a result,

there is no need to directly encode the frequent occurrence of these patterns in synchronic grammars themselves. Common instances of sound change give rise to commonly occurring sound patterns. Certain sound patterns are rare or unattested, because there is no common pathway of change which will result in their evolution.

What are the frequent sound patterns exhibited by the world's languages? What are the recurrent sound patterns which phonological theory attempts to explain? Many generalizations have been discovered in different domains. In (1) I list a sample of these, organized in terms of sound patterns they attempt to characterize. Some of these will be mentioned later in the book. The purpose of this brief discussion is to exemplify significant ways in which explanation within Evolutionary Phonology differs from other approaches to sound change and sound patterns.

(1) Examples of generalizations over sound systems of the world's languages

i. SEGMENT INVENTORIES
   a. If a language has only three vowels, it will usually have /i, u, a/
   b. All languages have voiced sonorants and voiceless obstruents in their segment inventories.
   c. In the series of voiced stops /b d g/, /g/ is most likely to be missing.
   d. No language contrasts voiceless laryngealized obstruents with their voiceless ejective counterparts.

ii. STRESS PATTERNS
   e. There are languages in which stress falls consistently on the first syllable of the word, or the last syllable of the word, but there are no languages in which stress falls regularly on the middle syllable of the word (e.g. the second syllable of a three-syllable word, the third syllable of a five-syllable word, and the fourth syllable of a seven-syllable word.)
   f. There are languages in which stressed syllables must be separated by single unstressed syllables, and others where stressed syllables must be separated by two unstressed syllables, but there are no languages where stressed syllables must be separated by three unstressed syllables.
   g. There are languages with long vowels and short vowels where all long vowels must be stressed, but there are no languages with long and short vowels where all short vowels must be stressed.

iii. PHONOTACTICS

h. In nearly all languages, each consonant in a syllable-internal obstruent cluster must agree in laryngeal features.
i. In many languages, each consonant in an obstruent cluster must agree in laryngeal features.
j. In many languages, there is no possible laryngeal contrast for obstruents in pre-obstruent position.
k. In languages where there is no possible laryngeal contrast for obstruents in pre-obstruent position, laryngeal contrasts are neutralized in this position in derived environments.

One question that has driven research in phonology and phonetics is *why* generalizations like those in (1) exist. How are they best explained? Are they best stated in terms of phonological primitives, as reflections of phonetic properties of speech, or a mix of the two? What are their origins? Are they tied to intrinsic properties of synchronic grammars, or is their appearance a consequence of the historical development of language? Do they serve a clear function in making speech easier to perceive or pronounce? Are certain properties, like those in (1h–k) related? Should some of these properties be viewed as accidental, or are there clear deterministic pathways in the course of language evolution? Of course, for each sound pattern in (1), there could be a different combination of answers to these questions, and a single generalization could also have multiple overlapping explanations.

The range of explanations offered for recurrent sound patterns can be illustrated with reference to the examples in (1). Consider, for example, the generalization in (1a), which suggests a general preference in three-vowel systems for the vowels /i, u, a/. In the twentieth century, this typological generalization seems to have been first discovered by Trubetzkoy, who wrote, in a letter to Jakobson in 1928:

> In the meantime I have started working on something else which fascinates me. I have compiled all vocalic systems I knew by heart (thirty-four in all) and tried to compare them . . . I will continue my work on them until I have collected about one hundred languages. The results are extremely strange. All systems can be reduced to a small number of types and can always be represented by symmetrical diagrams . . . There are some laws about the "formation of systems" which can be seen without difficulty . . . I believe that the empirical laws discovered in this way will be of great importance . . . (Jakobson 1975: 320)[2]

[2] Trubetzkoy's impressions are consistent with the findings of Maddieson (1984: 153–54), based on the UPSID database of 317 languages representative of the world's major

Trubetzkoy (1929, 1939) expands on his notion of the "formation of systems." In this work, he proposes one of the first feature systems for vowels which attempts to encode generalization (1a) by combining two phonological features, one of aperture (or sonority), and the other of timbre (place of articulation). His claim was that, with very few exceptions, aperture and timbre features were basic oppositions in all vowel systems. If the vowel system was triangular, it would involve a single vowel specified with the maximal degree of aperture /a/. The contrast between /i/ and /u/ was the addition of the timbre contrast to the close class of vowels. Trubetzkoy's approach classified the generalization in (1a) as an essentially phonological one, arising from constraints on the combinatory properties of phonological features. However, as our understanding of the physical properties of speech has deepened, other explanations for this universal tendency have been proposed. Within the quantal theory of speech (Stevens 1972), where quantal signals are those for which a distinct acoustic signal is achieved through a relatively imprecise gesture, [i], [u], and [a] are more quantal than other vowel sounds, and therefore better phonological categories. Another account of the same facts invokes the principle of vowel dispersion. This principle, first proposed by Liljencrants and Lindblom (1972), and elaborated by Lindblom (1986), suggests that vowels are evenly and widely distributed in the psychoacoustic vowel space. In other words, vowel systems are preferred to the extent that the perceptual space between vowels is maximized (independent of the ease or difficulty of the gesture). Under their account, for a three-vowel system, perceptual distance alone predicts the phonological categories /i, u, a/.[3] The quantal and dispersion approaches can each be viewed in either synchronic or diachronic terms. If vowels require too precise a gesture to generate distinctive categories, or, if they are too close in the perceptual vowel space, a synchronic distinction is impossible. At the same time, given that speech is transmitted by articulators which can lack precision, and that transmission occurs in a generally noisy environment, the accuracy of transmission of utterances from one generation to the next will depend to a great degree on the ease of articulation and ease of discrimination of different sounds. Language evolution will tend to converge on quantal vowels, or on vowel systems which obey the dispersion principle.

Maximal perceptual distance has also been suggested to account for the apparent universal in (1b): all spoken languages make use of sonorants

genetic groupings. Lindblom (1986) and De Boer (2001) treat similar generalizations as emergent properties of sound systems.

[3] Though see Maddieson (2003) on problems in extending this account to four-vowel systems.