

2 Evolution in Language and Elsewhere

It is a natural principle that the script and the sounds of language differ according to time and place.

Chén Dì (1541-1617)

Darwin's (1859) theory of the evolution of lifeforms has been used as a metaphor for historical developments in other fields almost since its conception. Darwin himself made use of the language-equals-species comparison, and this has continued in the modern field of linguistics, being most apparent in the field of historical linguistics. Mother languages develop, split into daughter languages, which in turn develop and split, much like the speciation of biological organisms. As biological organisms become extinct, languages also die, never to be heard again. The gradual change of languages in isolation in contrast to abrupt shifts which may occur through social contact have been likened to the punctuated equilibrium characteristic of biological evolution (Thurston 1987).¹ Environmental conditions leading to linguistic diversity, including richness of natural resources and geographic isolation, are similar to those associated with extreme localism and higher rates of speciation in life forms.² And, reflecting the similarities between models of genetic inheritance and those of language change, the trees used to represent linguistic relationships have a similar form and interpretation to those used in modern cladistics. A mother-node (e.g. Proto-Indo-European) defines a genetic family of languages, with daughter nodes (Indo-Iranian, Balto-Slavic, Germanic, etc.) defined by shared linguistic innovations.

However, it has long been recognized that language change differs in significant ways from biological evolution, and that comparisons between the two systems must remain metaphorical.³ Historical linguistics, the study of language change, is limited to approximately the past 5,000-7,000 years. Ancient writing systems go back no farther than 5,000 years, and the rate of language change makes historical reconstruction difficult for time depths greater than about 7,000 years. At the same time, human fossil remains reveal brain and vocal tract structures suggesting that the modern human language faculty is at least 50,000-100,000 years old.⁴ The linguistic study of language change then, is not the study of the evolution of language in the human species, or the study of the biological evolution of the language organ. The study of language change over the past 5,000-7,000 years assumes a mature human language faculty, and must attribute changes in the sound and shape of languages to cultural evolution. Language is passed along, like other knowledge, from one generation to the next; where an individual is isolated from the culture, it is not passed along, and where whole cultures die, languages die as well. Throughout this book language evolution refers to documented and hypothesized changes in linguistic systems which constitute one focus of historical linguistics. Evolutionary Phonology is the study of sound systems as a function of language evolution in this historic sense.

In Evolutionary Phonology, the metaphor of biological evolution is borrowed to model a very specific aspect of language change: regular phonetically based sound change and regular phonetically based sound inheritance at the level of the individual.⁵ At the most general level, it can be observed that sound patterns and living organisms change over time. More narrowly, a comparison is made between errors in DNA replication and errors in sound replication. In both cases, changes resulting from errors in replication are random and non-optimizing. At the same time, the primary source of shared characteristics in biological organisms and language is direct inheritance. To my knowledge, no one has made use of the evolutionary metaphor at the level of speech production and perception to model sound patterns in precisely this way.⁶ What Evolutionary Phonology offers is a way of viewing sound change in terms of well understood biological concepts including parallel evolution, convergent evolution, direct inheritance, adaptation, and disadaptation. Many recurrent sound patterns are argued to be the result of recurrent sound changes, and do not merit inclusion in the class of properties attributed directly to the human genetic language faculty. The biological metaphor is particularly useful here, since human languages, are arguably already highly adapted to human learning capacities (Locke 1983, Lindblom 1986, Deacon 1997).⁷

While the evolutionary metaphor informs this study, it is worth stressing that no direct reference is made to evolutionary principles in the formal model of sound change presented in this chapter, in the sound patterns reflecting common sound change discussed in Part II, or in the leading ideas and results of the model summarized in 1.7. Where some linguists find it useful to compare synchronic rule systems to the rules of chess, or to imagine the human language faculty as a parallel distributed processor, in the domain of sound change, analogies with biological descent are revealing. Sound patterns are viewed as emergent properties of synchronic grammars in the sense of Lindblom et al.(1984). There is no need to encode the primary content of phonological representations and constraint systems in the human mind, since these properties can be shown to emerge from natural processes of language change inherent to the transmission of language from one generation to the next.

2.1 Language Evolution

Evolution is on-going in all forms of life, as it is in all languages. Life-forms change over time and are able to pass these changes on to future generations. Evidence for continuous change in living creatures is everywhere: in the varieties of domesticated plants and animals, in the varieties and distributions of wild creatures, and, most dramatically, in the fossil record (Darwin, 1859). Like living organisms, languages change over time and these changes may be passed on to future generations. Evidence for continuous change in languages also surrounds us.

Ancient writing systems serve as the fossil record, preserving on stone or paper what was once a living spoken language. Long before Lamarck and Darwin were documenting biological evolution in all its forms, the Chinese scholar Chén Dì, (1541-1617), used ancient Chinese texts to identify sound correspondences between Old and Middle Chinese. Chén Dì was lucky, since the *Shījīng* (ca. 500 BCE) included Chinese poems organised by tone and rhyme, which could be compared directly to the

Qièyùn, a rhyming dictionary of the 8th century. Prior to Chén Dì's discoveries, the differences in rhyming schemes between the two texts were viewed as liberties or laziness on the part of *Shījīng* poets, or as evidence for changes in the poetic rhyming rules over the centuries, from imperfect to perfect rhymes. Chén Dì, perhaps the first true historical linguist of the modern era, laid the foundation for all subsequent work on the history of Chinese by showing that differences between *Shījīng* and *Qièyùn* rhymes followed from the simple fact that languages change over time.⁸

If written texts are the fossil records of language, how does the metaphor proceed? What circumstances approximate "Variation Under Domestication", the first chapter of *The Origin of Species*? In this chapter, Darwin highlights the extreme degree of intra-species variation found in domesticated plants and animals, suggesting that it is the unusual circumstance of domestication which allows otherwise useless, cumbersome or even monstrous features to be passed along. In his case study of pigeon varieties, Darwin (1859:84-85) concludes that these breeds:

...though agreeing generally in constitution, habits, voice, colouring, and in most parts of their structure, with the wild rock-pigeon, yet are certainly highly abnormal in other parts of their structure: we may look in vain throughout the whole great family of Columbidae for a beak like that of the English carrier, or that of the short-faced tumbler, or barb; for reversed feathers like those of the jacobin; for a crop like that of the pouter; for tail-feathers like those of the fantail.

What explains the seemingly greater diversity and abnormalities of cultivated animals and plants when compared to their counterparts in the natural world? Darwin's hypothesis, now substantiated by modern genetics was that "this greater variability is simply due to our domestic productions having been raised under conditions of life not so uniform as, and somewhat different from, those to which the parent-species have been exposed under nature" (1859:71). Domestic plants and animals do not compete for nourishment as they do in the wild, as this is supplied. They are taken out of their natural habitats, rendering some of their adaptive features no longer useful. At the same time, features which might lead to quick demise in a natural world full of predators, are freely passed along to future generations when those predators are absent.

Potential linguistic analogues to the abnormalities of domesticated pigeons can be seen in the unnatural situation where language acquisition is directly influenced by prescriptive laws or by spelling pronunciations. In the naturalistic home or community setting, children go about acquiring language with very little help. Where they are corrected, corrections are typically ignored, and a word or phrase which receives too much attention, even in the form of praise, may go unused for some time. The primary ingredients of healthy language acquisition in children are healthy social interactions including language input, and healthy individuals. Provided these basic conditions are met, the bulk of language acquisition, including natural processes of sound change, takes place within the first three to five years of life, well before most literacy skills are acquired.⁹ However, unnatural conditions may impinge on the naturalistic setting. Prescriptive grammarians may decide to impose their views of the way language should be spoken on others. Or, literacy skills may give rise to pronunciations which are altered in line with orthographic representations. As a result of these influences, a natural sound change

may be arrested or reversed, giving rise to the sound-based equivalent of the pouter's crop or the fantail's bulky display.

Consider one small but striking case in the history of English in the context of cross-linguistic sound patterns. All languages have voiced sonorant consonants, where sonorants include glides, liquids and nasals but very few of the world's languages have a contrast between voiced and voiceless sonorants. Limiting our attention to labio-velar glides, it appears to be the case that no language in the world contains /ɱ/ , a voiceless labio-velar glide, without also having /w/, a voiced labio-velar glide, and very few languages tolerate a contrast between these two sounds.¹⁰ In the context of sound patterns at large then, the contrast between /w/ and /ɱ/ could be seen as *comparatively* useless or cumbersome, as opposed to the contrast between, e.g. /w/ and /j/, which is wide-spread in the world's languages.

British English at one time had a phonemic contrast between /w/, a voiced labio-velar glide and /ɱ/, a voiceless labio-velar glide. These two sounds were able to distinguish sound-meaning pairs like *weather* and *whether*, and still do in dialects of Scots English (Ladefoged and Maddieson 1996:326), where the original contrast is maintained, and in many varieties of American English (Labov 1994:314). In the English fossil record, there is much evidence for this contrast: words spelled with *wh* generally reflect old *ɱ while *w* reflects *w. Compare *whit* vs. *wit*, *which* vs. *witch*, *where* vs. *wear*, *whine* vs. *wine*, etc. However, in most dialects of English, there is no longer a /ɱ/ phoneme; the previously voiceless glide has become voiced [w] (e.g. *awhile*), or voiceless [h] (e.g. *who*), and there is only a single voiced labio-velar glide /w/.¹¹ The merger of *ɱ and *w in Modern English is not unexpected. Just as the fantail's surplus of tailfeathers might hinder flight and lead to decreased survival rates in the wild, there is evidence that the rarity of /w/ vs. /ɱ/ contrasts in the world's sound systems is a function of phonetically determined low survival rates. Voiceless sonorants like [ɱ] have very little acoustic energy, are hard to hear, and may be lost over time, or reinterpreted as contextual variants of their voiced counterparts. In English and many other languages, voiced sonorants like /w/ are commonly devoiced due to coarticulatory laryngeal gestures. Additional factors which may play a role in this merger are the low functional load of the w/ɱ contrast, and the fact that other voiced/voiceless sonorant pairs in English are non-contrastive.¹²

However, the merger of /w/ and /ɱ/ was seen as a minor catastrophe in the eyes of prescriptive grammarians who saw loss of a phoneme as a sure sign of language decay. In this case, exercises were devised to reintroduce the contrast to American school children who had lost it ("repeat '[ɱ]ich [w]itch is [ɱ]ich?', and 'I [w]onder [ɱ]ether the [w]eather [w]ill change?' 100 times please"), with spelling conventions making the task somewhat easier. For some speakers, the prescriptive enterprise was successful, and a contrast between /w/ and /ɱ/ was maintained, or reestablished (though never in all the places it occurred historically), buttressed by the archaic English spelling system. Though the contrast

between /ɹɪtʃ/ and /wɪtʃ/ may not conjure up images of bizarre domestic pigeon breeds aside their svelte wild cousins, an instructive parallel exists. Just as the English carrier pigeon's beak can be viewed as a strange deformity, the isolated low-functioning voiced/voiceless contrast between /w/ vs. /ɹ/ in American English can be classified as a phonological abnormality. In English, the natural demise of this contrast has been artificially delayed or reversed through unnatural means: enforced practise and repetition, combined with spelling pronunciation.

As with living creatures, however, the varieties and distribution of wild or natural languages provides us with the most abundant evidence for constantly evolving linguistic structures. Every historical study of language is able to identify differences in speech across time. And every living language that has been studied involves variation within and across speakers. This variation may be random or predictable; it may be related to rate of speech or not; and it may be related to any number of social variables, including age, gender, socio-economic status, adolescent peer groups, or a multitude of other quantifiable factors (Labov 1994, 2001). Differences in sound patterns may be very small across generations, but they exist and provide evidence for the constant evolution of language. In the natural history of language, sounds change over time. Let us now look closely at how and why sound change occurs.

2.2 Sources of natural sound change

Evolution is a fact about life and language. But how does evolution take place? How do things change? Living creatures pass on their genetic blueprint through DNA. Errors occur when DNA is replicated, and these errors result in subsequent changes in the life-form. Languages have a much less precise inheritance mechanism than DNA. There is no miracle molecule at work. A child is born and learns natural language through exposure to that language as it is spoken in the surrounding speech community. Granted, there are neurological mechanisms involved which appear to make language learning very different from other learning tasks, but there is no direct replication of language in anything resembling a genetic blueprint. Languages are learned over the course of years, with a learning process of trial and error which is highly individualistic. No two children will experience the exact same linguistic input in their early years of life, and no two children, including identical twins, acquire language in precisely the same way. The world is a very noisy place, and it is in the context of this noise that language is transmitted. Given all of these factors, language change appears to be the rule, not the exception.

In this context, let us focus on language evolution in a very particular sense. First, we will confine our investigation to sound patterns and sound change. Second, we will look at language as it is transmitted from one individual to another, without considering additional social factors which also clearly play a role in language change (Labov 1994, 2001). With this narrow focus, we can return to the original observation, – that languages differ from living organisms in their fairly imprecise method of transmission. This process of transmission involves a speaker providing input to a listener, with the

listener attempting to internalize the speaker's grammar in order to understand speech. The process of transmission takes place in a sea of noise and starts from a point where the human infant listener has no knowledge of any sound-meaning associations in the speaker's language. We will assume throughout that the speaker's performance directly reflects his or her language competence, ruling out speech errors on the part of the speaker. Under these conditions, where can error creep in? What is responsible for the exact or inexact transmission of sounds across generations? What general models can be used to explain recurrent types of sound change in the world's languages?

The great majority of regular sound changes in evidence in the world's languages appear to be phonetically motivated.¹³ Evolutionary Phonology associates errors in transmission of sound patterns with the general typology of phonetically conditioned sound change in (1). This typology highlights three distinct natural phonetic sources of sound change which I refer to as CHANGE, CHANCE and CHOICE throughout this volume.¹⁴ One factor is the probability of an acoustic signal being misheard by the listener/learner in the course of language acquisition. If some signal A can be misheard as B, then a change of A > B is phonetically motivated on the basis of perceptual similarity. If a sound change has perceptual similarity as its primary basis, it is classified as an instance of CHANGE (3i). A distinct source of sound change involves the localization of non-local percepts. All speech involves some degree of coarticulation between adjacent segments (Hardcastle and Hewlett, 1999). If, in the course of language acquisition, a segmental representation is acquired, long-domain acoustic properties will give rise to ambiguities involving segmentation. If a sound change has ambiguous segmentation as its primary basis, it is classified as an instance of CHANCE (3ii).

(1) GENERAL TYPOLOGY OF SOUND CHANGE IN EVOLUTIONARY PHONOLOGY (S = speaker, L = listener)

- i. CHANGE: The phonetic signal is *misheard* by the listener due to perceptual similarities of the actual utterance with the perceived utterance.

Example: S says [anpa]

L hears [ampa]

- ii. CHANCE: The phonetic signal is accurately perceived by the listener but is intrinsically phonologically ambiguous, and the listener associates a phonological form with the utterance which differs from the phonological form in the speaker's grammar.

Example: S says [ʔaʔ] for /aʔ/

L hears [ʔaʔ] and assumes /ʔa/

- iii. CHOICE: Multiple phonetic signals representing variants of a single phonological form are accurately perceived by the listener, and due to this variation, the listener (a) acquires a proto-type or best exemplar of a phonetic category which differs from that of the speaker; and/or (b) associates a phonological form with the set of variants which differs from the phonological form in the speaker's grammar.

Example: S says [kakáta], [kǎkáta], [kkáta] for /kakata/

L hears [kkáta], [kǎkáta], [kakáta] and assumes /kkata/

A third source of sound change is the intrinsic variability of speech along the hyper- to hypo-articulated continuum (Lindblom 1990a). In all languages, speech varies according to rate. If, from the pool of variants, a listener chooses, as basic, a form which was non-basic for the speaker, sound change can occur. If a sound change has phonetic variation as its primary basis, it is classified as an instance of CHOICE (3iii).

Misperception or confusion of one phonetic percept with another under CHANGE is common and well documented. Laboratory experiments documenting perceptual biases range from the propensity to perceive [ki] as [tʃi] (Guion 1996, 1998), and [θ] as [f] (Eilers 1977), to the interpretation of a nasalized vowel [Ṽ] as [ṼN] (Kawasaki 1986; Lahiri and Jongman 1990). In the example illustrated in (1i), a speaker says [anpa], but the listener hears [ampa]. This is one instance of a much more general pattern where an intervocalic heterorganic sequence of nasal stop followed by oral obstruent is misperceived as an intervocalic homorganic nasal-obstruent sequence. What phonetic principles underly this misperception? The sounds [n] and [m] are produced at distinct points of articulation, alveolar and labial respectively; these two sounds differ acoustically, both in terms of the formant transitions from the preceding vowel into the nasal, and in terms of the values of their nasal antiresonances. Some languages do contrast [n] and [m] in the same context (e.g. Ngiyambaa *bunbil* ‘pillow, for beating time to singing and dancing’ versus *bumbil* ‘instrument for mimicking emu calls as a lure’), so it is clearly possible for humans to distinguish these sounds in this context. Nevertheless, misperceptions of this type occur with greater than chance frequency, and are argued by Ohala (1981, 1985, 1990, 1993) to result from the intrinsic weakness of place cues for the nasal in contrast to those of the following pre-vocalic stop.¹⁵ Notice that all instances of CHANGE, by definition, will involve changes in *pronunciation* between the speaker and listener. Whether these changes in pronunciation lead to phonological reanalysis on the part of the listener is not determined by the source of change itself. In a language where [n] and [m] contrast in other contexts, e.g. word-initially before vowels, or intervocalically, the sound change shown in (1i) will typically lead to phonological reanalysis of /anpa/ to /ampa/. If however, there are limited environments where [n] and [m] contrast, the listener may assume /n/, and incorporate place-assimilation into the grammar. In the first case, CHANGE gives rise to a change in both pronunciation and phonological representation; in the second case, it is instantiated by a change in pronunciation and a new phonological alternation. One phonetic source of sound change, then, is misperception on the part of the listener. Certain misperceptions, or perceptual confusions, are more likely to occur than others, and give rise to instances of CHANGE. While all cases of CHANGE involve a change in pronunciation, associated changes in phonological representation may, but need not, occur.

A final comment on CHANGE regards the difficulty of finding empirical evidence for it in children’s speech. As detailed in 9.1, there is good evidence that during the first three to four years of speech, the majority of a child’s mispronunciations reflect maturational constraints on production, not reflections of language competence. Fully mature segmental organization is generally not complete until the early

school years. During the early stages of acquisition, children's pronunciations can be quite far from adult norms. In the earliest stage (9-18 months), whole syllables are left unpronounced, CV syllables predominate, and only a small number of basic consonants are used. However, even by 3-4 years when closed syllables, consonant clusters, and fricatives appear, many reductive strategies are still found in children's speech, including unstressed syllable reduction ([æmz] for *animals*), final consonant loss ([fɔ] for *thought*), cluster reduction ([saimi] for *slimy*), consonant harmony ([lelou] for *yellow*), and a variety of segment substitution processes ([tau] for *cow*, [mu:b] for *move*, [siŋ] for *thing*, etc).¹⁶ If children mishear the speech around them, and as a consequence, attempt to pronounce misheard words, there may be no evidence of this change in the child's speech for the simple reason that the child is far from being able to coordinate the articulatory gestures involved in accurate production of adult targets. At the point at which a child finally does produce a string like [ampa] instead of adult [anpa], an attentive adult who actually perceives a consistent difference between these two forms is more likely to be delighted at hearing [ampa] as opposed to [pa], [aapa], or some other more distant variant, than engage the child in a mini-session of articulatory phonetics. Given this, during the point at which sounds are being misheard, the adult has no real clues as to a child's grammatical restructuring. There is good reason to believe then, that at the level of the individual, sound changes with sources in CHANGE can slip into a child's grammar almost unnoticed, providing the seeds for one source of phonetically based sound change within the wider speech community.

Intrinsic phonological ambiguity under CHANCE is also well documented.¹⁷ The term CHANCE refers to the fact that, unlike CHANGE, there is no language-independent phonetic bias involved in this type of sound change; the signal is inherently ambiguous, though independent structural features of the language may give rise to higher probabilities for one phonological analysis than another. Some reanalyses of this sort are regular metatheses and regular dissimilations (Ohala 1993; Blevins and Garrett 1998, to appear; Ritchie 1999). Sound changes of this sort most commonly involve phonetic features with multisegmental domains, like the example of laryngealization in (Iii).¹⁸ The example in (Iii) shows the speaker producing [ʔaʔ], a short laryngealized vowel, preceded and followed by glottal closure. Vowel laryngealization is a phonetic feature associated with glottal stops, and glottal stops are often redundant properties of laryngealized vowels. If the listener assumes any phonetic redundancy within the string, then there are at least four potential phonological analyses of [ʔaʔ]: /ʔa/, /aʔ/, /ʔaʔ/ or /a/. If /ʔa/, /ʔaʔ/ or /a/ is chosen, a sound change has occurred.

There are three important ways in which CHANCE differs from CHANGE. First, CHANCE does not involve misperception on the part of the listener. Second, CHANCE does not involve any immediate change in pronunciation on the part of the listener. A gradual shift in pronunciation can evolve as a consequence of the new phonological analysis, but it need not. A third difference between the two sources of sound change involves inherent biases versus priming effects. In CHANGE the probability of a sound change occurring is generally related to biases in the human perceptual system; in CHANCE,

pre-existing sound patterns within the ambient language being acquired may produce priming effects which increase the probability of sound change.¹⁹

This last difference highlights an intriguing property of the general model of sound change in (1): sound changes are modeled as probabilistic, with higher probabilities of occurrence relatable to other phonetic features of the language in question, or to general strategies involved in extracting sound patterns from the acoustic signal. For example, if a language has a high token frequency of CV syllables, then a listener hearing [ʔaʔ] may be more likely to interpret this as an instance of /ʔa/ than /aʔ/. The listener makes a simple frequency-based inference: most syllables are CV, therefore [ʔaʔ] is most likely CV, therefore, posit /ʔa/. In chapter 6, the common structure-preserving nature of several different types of sound change is attributed to language-specific factors of this sort. Chapter 6 also suggests a general preference rule involved in phonological acquisition: all else being equal, a single segmental source is assumed for a particular phonetic/phonological feature. This principle will rule out /ʔaʔ/ as a potential underlying form in (1ii), since laryngealization is associated with two segments, giving preference to /ʔa/, /aʔ/, or /a/.

Both CHANGE and CHANCE involve idealizations of the phonetic world, since both take as input single invariant tokens of a particular speech form. However, all spoken languages involve a great deal of intra-speaker variability in the phonetic realization of phonological forms. Intra-speaker variability is arguably a manifestation of accordion-like transforms of the speech stream, from the forceful maximally segmentable stretched out profile of clear speech, to the lax coarticulated and highly compressed mumble of casual speech. In CHOICE, intra-speaker phonetic variability is the source of sound change. The term CHOICE refers to intrinsic choices offered to the language learner in observing phonetic variation: these include choices for the center or prototypical exemplar of a particular category, as well as choices as to which variant should be chosen when different variants suggest different phonological forms. Following Lindblom (1990a, 1998), intra-speaker variation is viewed as a direct manifestation of the continuum from hyperarticulated listener-oriented 'clear' speech (Moon and Lindblom 1994) to reduced, hypoarticulated 'casual' speech (Lindblom et al. 1992; Willerman 1994).²⁰ In clear speech, the speaker articulates more forcefully: segments are generally longer and ambiguities due to coarticulation are reduced. In casual speech, the opposite holds: segments are generally shorter, and coarticulation is extreme. Consider the pronunciations of /ka-kata/ 'to laugh' in Bellonese, a Polynesian Outlier language of the Solomon Islands. In hyperarticulated speech, there is a full vowel in each syllable: [ka.ka.ta]. However, in fast speech, the vowel flanked by adjacent consonants is not pronounced and we find [kkata]. Intermediate forms present the listener with a short vowel between the identical consonants: [kākata]. In this example, two related choices present themselves. First, which *phonetic variant* of the word 'to laugh' is the proto-typical one? Second, should this word, and others with the same patterns of vowel reduction, be represented phonologically with or without a vowel between the two identical consonant sounds? A naive view might be that speakers will choose /kakata/ due to the fact that two of the three basic variants contain a vowel.

However, there is growing evidence that token frequency plays an important role in the acquisition of sound patterns (Skousen 1989, Bybee 2001, Pierrehumbert 2001). If [kakata] has the lowest token frequency, occurring only in careful speech, ritual speech, or song, listeners may base their hypotheses on other, more common, everyday variants. Another complicating factor in this example, which I return to in 6.2.2, is that the variable vowel always has the same quality as the vowel of the following syllable, and hence is *predictable*. If the vowel in [kākata] is short enough, and of predictable quality, it may be interpreted, not as a phonological vowel, but as a transition associated with phonetic release of the /k/ sound.²¹ In this example, where the historical form is /kakata/, choice of /kkata/ by the listener is an instance of sound change. Many examples of CHOICE are illustrated in chapters 6 and 7, and sound changes which combine CHOICE with CHANGE or CHANCE are also discussed. The most important defining characteristic of CHOICE is its source in intrinsic phonetic variation. As a consequence of its source in variation, changes in token frequencies of particular phonetic variants are catalysts in CHOICE. To clarify this last point, it will be useful to lay out the model of sound change in somewhat more detail.

In (2), the model of sound change proposed in (1) is amplified by showing cases where no sound change occurs (2i), by integrating phonetic variation into all types of change, and by identifying universal aspects of the phonetics-to-phonology mapping. In (2), three independent phonetically governed sound changes are illustrated: fronting of the back vowel [u] adjacent to a dental/alveolar consonant under coarticulation, **ut > yt*, (Ohala, 1981); **θ > f* due to perceptual similarity of lamino-dental and labiodental fricatives [θ] and [f] (see 6.1.1) and **t' > ʔ*, the change of a voiceless glottalized (unreleased) stop to a glottal stop, which results from anticipation of glottal closure combined with absence of release (see chapter 5).

As is well known, the mappings between acoustic features and units of perception is of a non-trivial nature:

The basic elements of language, its consonants and vowels, are not objective, physical units, but rather psychologically defined entities that can be reliably identified only by a human listener. Speech is not a simple left-to-right sequence of discrete and invariant alphabetic segments such as we see on a printed page. The reason for this is that we do not speak phoneme by phoneme, or even syllable by syllable. Typical rates of speech – 10 to 15 phonemes/second – are possible because we coproduce, or coarticulate, the units. At each instant, our articulators are executing overlapping patterns of movement that may correspond to several neighbouring phonemes, including phonemes in neighbouring syllables. The result of this shingled pattern of movement, is, of course, a shingled pattern of sound, in which the acoustic structure of a given consonant or vowel varies from one context to another. Thus, the units of the acoustic signal do not correspond, one for one, with the units of perception. (Studdert-Kennedy 1998:169-70)

Nevertheless, assuming that speech, at some level, is represented in terms of segments and features, universal aspects of the mapping, listed in (3) can be defined. The numbered arrows in (2) refer to the steps listed in (3); double-shafted arrows in (2) indicated predicted directions of change given phonological restructuring.

(2) A formal model of sound change with examples

i. No sound change

Speaker	Listener
/ut/	/ut/
↓1	↑3
[ut,ut,yt...] _{vi}	2→ [ut,ut,yt...] _{vi}

iii. CHANCE: sound change via 'mis-application' of phonetic-phonology mapping

Speaker	Listener
/yt/	/ut/
↓1	↑3
[yt...] _{vi}	2→ [yt...] _{vi} ⇒ [ut,yt...] _{vj}

iva. CHOICE: sound change from phonetic variation (change in variant frequency)

Speaker	Listener
/ut/	/yt/
↓1	↑3
[yt,ut,...] _{vi}	2→ [yt,ut...] _{vi}
4↑	
[ut,yt,...] _{vj}	

ii. CHANGE: sound change via misperception

Speaker	Listener
/θa/	/fa/
↓	↑3
[θa...] _{vi}	2→ [fa...] _{vj}

ivb. CHOICE: sound change from phonetic variation (change in variant set)

Speaker	Listener
/ut/	/yʔ/
↓1	↑3
[yʔ... yt'] _{vi}	2→ [yʔ,...] _{vj}
4↑	
[yt', yʔ...] _{vi}	

(3) Universal decomposition of sound change

Step 1: Universal and language-specific phonetics give rise to a range of surface forms in natural speech production, abbreviated here by []_v (= [] and its phonetic variants), where subscripts indicate identity, or lack thereof, and are listed in order of decreasing frequency.²²

Step 2: Utterance (set) is perceived 'correctly' (i,iii,iv), or 'incorrectly' (ii).

Step 3: Utterance is associated with a phonological form on the basis of

- a. Universal knowledge (features, prosodic categories)
- b. General segmentation and pattern matching algorithms
- c. Default mapping principle: one-to-one association between features and segments
- d. Other ambient sound patterns of the language being acquired
- e. Relative frequency of variants

Step 4: Shift in variant frequencies from speaker as child to speaker as adult

In (2i), there is no sound change, and both the range of phonetic surface forms, their relative frequencies, and the phonological interpretation of the sequence are directly inherited.²³ The context-free *θ > f in (2ii) is an example of CHANGE with variation playing no role. In (2iii), CHANCE is illustrated by the dissimilatory change *yt > ut. In this example, the listener attributes the frontness of [y] to the affect of the following [t], and on this basis, posits an underlying form which lacks the shared

assimilatory feature. Because the speaker's [yt] was an instance of /yt/, the listener's /ut/ constitutes a sound change. Notice again that in (2iii), there is no immediate change in pronunciation. The double-shafted arrow in (2iii) shows the predicted direction of change for this particular sequence. Because the vowel has been analysed as /u/, it will likely move towards other phonetic tokens of /u/ in the language if they exist.

In (2iv), two cases of CHOICE are diagrammed. (2iva) involves a simple vowel change $*ut > yt$, where the high frequency of fronted variants of /u/ gives rise to a category shift. In (2ivb) a more complex example of CHOICE is illustrated by $*ut > yʔ$. Here, there is an overt change in pronunciation. Whereas the speaker has tokens [yt'] and [yʔ], due to the high frequency of [yʔ], the listener adopts this as the basic form from which the phonological representation is extracted. The interest of this type of example relates to loss of the phonetic conditioning factor for vowel fronting: due to a sound change of $*t > t' > ʔ$ (cf. a similar change in many dialects of English), the $*t$ which conditioned vowel fronting is lost. Variation plays a direct role in the changes in (2iva) and (2ivb), but not elsewhere.

A few remarks are in order regarding Step 3 of (3), where a listener extracts generalizations from the raw acoustic signal based on, at least, the factors listed. Universal aspects of phonetic and phonological knowledge include audible phonetic features, distinctive phonological features, segments, and prosodic structures, which can all be identified by human infants, as summarized in Werker and Pegg (1992), and reviewed in 9.1. General segmentation or pattern-matching algorithms are, to a large extent, unexplored. For the purposes of this volume, all that is important is that frequency of particular variants can play a role in the learner's development of phonetic and phonological categories. The approaches consistent with this model of sound change include: Skousen's (1989, 1992) analogical model; the WRAPSA (Word Recognition And Phonetic Structure Acquisition) model of Jusczyk (1992); Stemberger's (1992) connectionist view; Pierrehumbert's (2001) exemplar dynamics; Wedel's (2003) analogical modeling of sound patterns invoking self-organizing principles; and analogical learning confirmed by the experimental results Ernestus and Baayen (2003). In all of these approaches, token frequency may play an important role in sound patterns, and may initiate instances of sound change with sources in variation. Given that the set of utterances which any child hears in the course of language acquisition will be different from that of the next child, with different frequencies of, e.g. word variants, the model in (2)-(3) implies that every individual's grammar will be different at the level of phonetic implementation. This recognition of minute differences in individual grammars sets Evolutionary Phonology apart from other approaches where speakers of the same language are assumed to have the same grammar.

The importance of frequency in instances of CHOICE within this model is worth highlighting. In traditional neogrammarian treatments, gradual articulatory drift was attributed to 'mechanical' factors, which in Step 1 of (3) are the transforms of speech which occur at different rates, and which distinguish clear speech from casual speech. However, what remained unexplained within the neogrammarian account was *why* these mechanical shifts took place. Within the model of sound

change proposed in (1)-(3), there is an uninteresting, but testable answer given. Mechanical shifts in pronunciation take place at the level of the individual when the input the learner is exposed to gives rise to different relative frequencies of variants from that which gave rise to the speaker's grammar. That frequency effects play a role in shaping sound patterns is well demonstrated (Bybee 2001; Bybee and Hopper 2001). By allowing frequency effects to guide a language learner's phonological generalizations, shifts in articulation can be modeled as shifts in the perceptual-acoustic space defined by token frequencies.²⁴ CHOICE, as modeled in (1)-(3), incorporates change at the subphonemic level (changes in pronunciation) as well as changes which are often referred to as 'phonologizations', where there is a change in some aspect of the phonological system.

A final aspect of CHOICE which I return to in 8.3 is that it is the sole locus of what is sometimes termed 'teleology' in sound change. When a range of variants exists for a particular word, a speaker may choose one variant over another to maximize contrast with an otherwise homophonous word. I suggest in 8.3 that this type of choice on the part of the speaker is typically limited to paradigm-internal contrasts, and can account for certain cross-linguistically rare feature contrasts.

There are several important differences between the model in (1)-(3) and other treatments of phonologization (e.g. Hyman 1977, Ohala 1981, Hajek 1997). First, Step 1 of (3) assumes no fundamental distinction between universal phonetic rules and language-specific phonetic implementation rules. This is based on empirical findings questioning the universal/language-specific phonetics dichotomy. Kingston and Diehl (1994) suggest that much of phonetic interpretation is controlled or language-specific, and that far more articulations are directly controlled by speakers than was previously thought. Of particular interest is their finding that the lowering of F_0 on vowels after voiced consonants in English occurs even in contexts where the phonologically [+voiced] segment is *not phonetically voiced*, or in any obvious way phonetically distinct from allophones of the [-voiced] category in other contexts. If this lowering occurs in the absence of the phonetic feature which is universally associated with, then clearly, it is no longer the result of automatic phonetic processes. At the same time, studies of coarticulation suggest general trends related to other language-specific factors (Manuel 1999). Inter-speaker variation shows that some speakers have little coarticulation compared with others "who show relatively extensive temporal and spatial coarticulatory effects" (Manuel 1999:189). It is unclear how such intra-speaker variation can be accounted for without some blurring of the universal/language-specific boundary. As suggested by Manuel (1999:189), the role of contrast may "set a maximum limit on coarticulation, but has little to say about how it is further limited."

Another difference between the CCC-model and more circumscribed accounts of phonologization is that it incorporates context-free splits and mergers. This is not true of Hyman's (1977) model. Hyman's assumption is that the evolution of a new contrast implies or leads to the loss of a former contrast, - that there is an inevitable trading relationship between contrasts: "*the development of a phonological rule carries the seeds of its own destruction.*" (Hyman 1977:412). However, there are context-free sound changes (e.g. $*\theta > f$) which involve changes in phonological representation but no

trading relationship. And there are also processes like the phonologization of vowel harmony, where the trigger of the assimilatory change is not lost, though a shift from variable coarticulation to harmony needs to be modeled. Even if Hyman's model is restricted to the evolution of new phonological contrasts or phonological alternations, the descriptive generalization does not hold. In many languages, [s] is palatalized to [ʃ] adjacent to [i], but this does not typically result in a shift from /i/ to /ɨ/, when a contrast between /s/ and /ʃ/ evolves. A final difference between the CCC-model of sound change and those of Hyman (1977), Ohala (1981), and Hajek (1997), is the prominent role given to phonetic variation in instances of CHOICE, and the integration of frequency effects in determining the output of sound change under CHOICE.²⁵

The central role of phonetic variation in defining the choice space for sound change has been suggested by Passy (1890), Sievers (1901), Ohala (1974a, 1989), Kiparsky (1988, 1995) and many others. However, there are few accounts in which frequency effects are modeled in sound change. One notable exception is Watson (1999), who highlights the role of frequency with reference to sound changes like the one diagrammed in (2iva):

Ohala is inexplicit as to why listeners of one generation should have a different response to such variability from their forebears, suddenly failing to do the appropriate filtering out of contingent features. A cause of this may be a statistical change in the pattern of variability... (Watson 1999:160)

Watson provides an analysis for how a /VN/ sequence with a small amount of contextual nasalization, (a) below, shifts phonetically to [ṼN] with a greater amount of contextual nasalization, (b) below, which is phonologically reinterpreted as a nasalized vowel:

Time 1: [VN] with small amount of contextual nasalization (a) V [-nasal]
 Time 2: [ṼN] greater amount of contextual nasalization (b) V [+nasal]

His analysis makes direct reference to changes in the statistical frequency of variation:

The essential factor in bringing about this abduction is the individual variability already discussed in coarticulation... At time point 1, a majority of speakers in a language community produce vowel nasalization only to extents consistent withthat is with no phonological specification of vowel nasality, only [VN] coarticulation. Even at this stage, a minority of speakers might show significantly more nasalization... this need not effect the community's norms. However, this may be changed by a small sociolinguistic shift. If a higher proportion of speakers (or speakers from a higher status group) shift to a larger degree of coarticulation (which they can do without any necessary change in their own phonological representations), new generations of speakers will be led to infer phonological structure (b) rather than (a)." (Watson 1999:166)

Other recent models incorporating frequency effects include Bybee (2001) and Pierrehumbert (2001).

Though the CCC-model isolates distinct phonetic sources of ambiguity, it will often be the case that a single sound change has multiple sources. For example, nasal place-assimilation in VNTV is characterized as a canonical instance of CHANGE, though many languages allow for coarticulation of N and T in non-careful speech, suggesting that CHOICE may also be involved (see Chapter 5).

Coarticulation is a natural feature of speech at normal rates, and can give rise to inexact transmission of speech, as in the case of CHANCE illustrated in (1ii) and (2iii) above. At the same time, the speaker's ability to control a phonetic continuum from casual to clear speech can be understood as a potential check on the segmentation problem. But this intra-speaker flexibility results in variation, feeding CHOICE, which can also result in sound change. To summarize, CHANGE, CHANCE and CHOICE are intrinsic features of the inexact language inheritance mechanism and constitute primary sources of sound change at the level of the individual.

2.3 Non-optimal sound change

Though Darwin was unaware of DNA as the conduit for genetic inheritance, one of the most important aspects of his theory of evolution is the assumption that changes which occur in the course of evolution are random. The study of genetic mutations is in its infancy, yet, as far as we can tell, these mutations are non-optimizing in the sense that they do not necessarily result in a better organism (Gould and Lewontin, 1979). Like genetic mutations, the three phonetic sources of sound change defined in 2.2 are in no way goal-directed. Sound change happens, but it does not happen in order to make speech easier to articulate, easier to perceive or easier to transmit. Sound change happens, but it does not necessarily result in a more symmetrical, more stable or generally improved phonological system. Sound change happens, but for every case where it happens, there is a parallel case where it does not happen.

The non-optimizing nature of sound change finds empirical support in cross-linguistic studies of metathesis (Blevins and Garrett 1998, to appear). Many cases of CHANCE like $a? > ?a$ in (2ii) can be paired with mirror-image changes in other languages. For example, in Slavic $VR > RV$, where R is a liquid, but in Le Havre, a Romance language, $RV > VR$. Any claim for Slavic that RV is generally better (phonetically or phonologically) than VR will have to contend with the fact that RV sequences are eliminated in favour of VR sequences in Le Havre.²⁶ And, there are many languages in which VR and RV sequences have endured: in Standard French, most native words beginning in $prV...$ reflect Latin prV , while in many Pama-Nyungan languages, Proto-Pama-Nyungan $*VR$ is reflected as VR . In both cases, direct inheritance reflects the stability of these sequences over hundreds or thousands of years.

A different argument against optimizing sound change is that a single sound change may simplify articulation at the cost of perceptual saliency or vice versa. In Bellonese where /kakata/ has at least three perceptually distinct surface forms, [ka.ka.ta], [kākata], and [kkata], the listener must choose between at least two phonological representations /kakata/ and /kkata/. The casual speech variant, [kkata], is, one might argue, easier to articulate, since it involves less articulatory effort, with one less vowel gesture than the careful speech variant. Ease of articulation, then, might favour /kkata/. At the

same time, positing /kkata/ means that the language has a new phonological contrast between /kkata/ and /kata/. As noted in chapter 7, pure durational contrasts like [kk] vs. [k] are difficult to perceive in word-initial position, and often lead to neutralization of the length contrast in this position. One problem then, is that where ease of articulation favours /kkata/, maintenance of perceptual contrast favours /kakata/. While it is conceivable that tradeoffs of this sort may be calculated for any pair of phonetic forms, a serious problem is the fact that phonological representations like /kakata/ have *multiple phonetic variants*.²⁷ Intra-speaker variability itself makes the calculation of phonetic optimality of phonological forms intractable. In order to calculate ease of articulation or perceptibility values for /kakata/ vs. /kkata/ we must know precisely how each will be phonetically implemented by the articulators. But, as already demonstrated, intra-speaker variation provides a range of alternatives which differ precisely in these values. In fact, as a case of CHOICE, at the point in time where /kakata/ is reanalysed as /kkata/, there is no difference in the articulatory space associated with each of the two forms.

Finally, we turn to the most interesting challenge for a model of non-optimizing sound change, the case of CHANGE. Recall the observation that [anpa] may be misheard as [ampa], but [anpa] is not misheard as [anta]. The unidirectionality of this misperception smacks of optimization. Ohala's (1990) general explanation for this sound change is the weakness of perceptual cues to place of articulation in the post-vocalic pre-obstruent environment, and their strength in pre-vocalic position. In functional terms, /n/ vs. /m/ before /p/ is hard to hear, so the contrast is eliminated.²⁸ However, Ohala's experimental evidence indicates that speakers simply *do not hear* [np] as [np] a good part of the time. No principle of optimization is necessary, since the sound change follows from the way our auditory system processes speech. In other cases of CHANGE, an optimization-approach is disconfirmed by the same sort of bidirectionality exhibited by metathesis sound changes. An instructive case is the perceptual affinity between nasality and aspiration, known as rhinoglottophilia (Matisoff 1975:265).²⁹ Breathiness and nasalization have similar acoustic effects on the vowel spectrum (Ohala 1975:303), and these spectral similarities can result in misperception of breathiness as nasalization (Ohala 1980, 1987; Klatt and Klatt 1990). In the natural world of sound change, nasalization appears as a reflex of aspiration in Bzehedukh and Shapsegh, two Caucasian languages, while the Owerri dialect of Igbo shows aspiration as a reflex of nasalization (Blevins and Garrett, 1993). The inevitable conclusion is that sound patterns, like organic forms, are "not an array of optimal adaptations to their immediate surroundings, but complex products of history, not always free to change in any direction that might 'improve' them." (Gould 1990:67).

If the three basic types of sound change in 2.2 are non-optimizing, then what factors are responsible for the sound patterns which repeat themselves in one language family after another? Why do so many unrelated languages have word-final devoicing, with word-final voicing a rarity? Why does word-final place-neutralization of /p,t,k/ occur with greater-than chance frequency, while similar word-initial neutralizations are unattested? Why do consonantal length contrasts tend to exist in the middle of words and not at word edges? And why is it possible to come up with near-exceptionless phonological

generalizations? If changes which give rise to these patterns cannot always be seen as improvements on the sound-shape of language, what does give rise to these striking similarities in form?

2.4 Sources of similarity

When a biological character in one organism resembles that of another, there are four logical sources for the observed similarity. The first and most usual explanation is that similar characters reflect *direct genetic inheritance*. In this case, the species are genetically related and the feature in question has been acquired by descent from a shared ancestor which also had this feature. The importance of inheritance as an explanation for shared biological features is stressed by Gould (1983:80), replying to Dobshansky's (1951) application of adaptive landscape to the question of discontinuous distribution of species:

...surely the cluster of cats exists primarily as a result of homology and historical constraint. All felines are alike because they arose from a common ancestor shared with no other clade. That ancestor was well adapted, and all its descendants may be. But the cluster and the gap reflect history, not the current organization of ecological topography. All feline species have inherited the unique cat *Bauplan*, and cannot deviate far from it as they adapt, each in its own particular (yet superficial) way. Genealogy, not current adaptation, is the primary source of clumped distribution in morphological space.³⁰

The acceptance of direct genetic inheritance as a prime component of linguistic structure was fundamental to the neo-grammarians of the 19th century, but generally abandoned by the Chomskian generative tradition.³¹ Though the generative view is perhaps not as far-fetched as Dobshansky's model of species distribution, they share a methodological fallacy: a mechanism (like adaptive landscape or innate properties of universal grammar) which can explain *some* aspects of regular distribution, is taken to explain *all* aspects of regular distribution, despite the existence of alternative explanations, including direct inheritance.³² Within Evolutionary Phonology, the source of similarities across languages is the subject of empirical investigation. Where a common ancestor possesses a sound pattern found in one of its descendants, and where it can be shown that there is a direct path of inheritance between these languages and sound patterns, the sound pattern of the descendant language is explained through descent, and synchronic principles proposed to derive it are likely to be superfluous. All Indo-European languages have closed syllables of the form CVC. Should this be surprising, given that Proto-Indo-European also had closed syllables, and that many of these were directly inherited by daughter languages? In language, as in life, the primary source of shared characteristics is direct genetic inheritance.

A second account of similar characteristics in different organisms highlights the fact that the characteristics are not as similar as they look. If the characteristics are only superficially similar, having arisen through different developmental pathways from different ancestral conditions the

development is referred to as *convergent evolution*. A well studied biological example involves the eyes of vertebrates and the eyes of cephalopods (cuttle-fish, nautilus, octopus, etc.). These eyes have superficial similarities, but are very different in design and function, and have evolved independently. So, in language, superficially similar sound patterns may have arisen in different ways from different kinds of sound change.

Convergent evolution in language is visible at many different levels. Words can be very short, consisting of two or three sounds, and sound inventories can be quite small (20-30 phonemes), so the probability of two words with a similar sound – meaning relationship occurring among the 4,000 or so languages, all with vocabularies of tens of thousands of words is very high.³³ An example of convergence at the lexical level is the word for 'dog' in English and in Mbabaram, an Australian Aboriginal language of the Cape York Peninsula. In both languages, one sound sequence associated with this general meaning is [dɔŋ]. Though the Mbabaram word and the English word have similar sounds and meanings, we know that the general relationship between sound and meaning in language is an arbitrary one.³⁴ In this case then, the resemblance between the English and Mbabaram words for 'dog' is either the result of convergent evolution (e.g. an accident), or a result of borrowing. Comparative work on the languages of Cape York allows us to identify this as a case of convergent evolution: Mbabaram /dɔŋ/ derives from proto-Paman **gudaga* 'dog', by four regular sound changes: **gudaga* > *gudɔga* > *udɔga* > *dɔga* > *dɔŋ* (Hale 1964, 1976a,b). The superficial relationship between the English and Mbabaram words is a spurious one, and unlikely to illuminate fundamental aspects of sound patterns or sound change.

More relevant to this study are cases of convergent evolution at the level of sound patterns and sound inventories, independent of meaning. In one relevant case, an inherited feature of one language is non-inherited in another.³⁵ Convergent characteristics are found in Japanese and Gilbertese, a Micronesian language. In both languages, the only possible word-final consonants are nasals.³⁶ Though these look like superficially similar constraints, they have arisen in different ways. In Pre-Gilbertese, all inherited forms with final nasals were once vowel final. Loss of word-final voiceless vowels has left the nasal exposed in final position (Blevins, 1997). In Old Japanese, all words also ended in vowels. However, the development which led to final nasals was not one of sound change, but of language contact (Shibatani 1990: 121-22). Borrowings from Chinese languages containing nasal-final words are the source of nasal-final words in Middle Japanese. Again, a superficial relationship, in this case, between Japanese and Gilbertese, cannot be explained by a single phonetic or phonological principle.³⁷

Another case of convergent evolution is the emergence of similar sound patterns or contrastive oppositions which arise by very different types of sound change. For example, consonantal length contrasts may evolve through vowel loss between identical consonants, as in the **kakata* > *kkata* change in Bellonese, or, through post-tonic lengthening of an earlier short consonant, as in Buginese *takke* 'stick' from proto-Austronesian **təkən*. Chapter 7 details seven distinct general pathways leading

to the evolution of consonant length contrasts. The importance of convergent evolution in this case study is that distinct evolutionary pathways give rise to distinct cases of geminate behaviour within synchronic phonologies.

A final case of convergent evolution is the existence of identical sound changes with distinct evolutionary sources. Consider a sound change involving vowel loss: $V > \emptyset$. Vowel loss can arise from CHANGE, in the case where, for example, a short final voiceless vowel is not perceived due to its weak intensity. But vowel loss can also result from CHOICE, as in the Bellonese **kakata > kkata* example. Whether misperception or CHOICE is the source of sound change has implications for the form of synchronic grammars, as demonstrated in chapter 7. Another example of a multi-source sound change are common assimilations like **np > mp*. As already discussed, this sound change can result from CHANGE, as described in 2.2, where [anpa] is misheard as [ampa]. But the same sound change can also arise from CHOICE: in many languages, including English, /np/ sequences are produced as [np] in clear speech, but can be pronounced as [mp] in casual speech (e.g. [ɪnpʊt], [ɪmpʊt] for 'input'). The sound change **np > mp* may be an instance of CHANGE, CHOICE, or a combination of the two. Where two common sound changes converge on a single sound pattern, such sound patterns are expected to be more common cross-linguistically than cases where convergence is not a factor.

Convergent evolution contrasts with a third logical possibility for the evolution of like characters: *parallel evolution*. In biological systems, parallel evolution describes similar developmental modifications which evolve independently. Under parallel evolution a feature with more than one evolutionary origin comes into existence by similar developmental processes from similar ancestral conditions.³⁸ For example, within the family of lizards, toepads have evolved independently in three lineages: *Iguanidae*, *Scincidae* and *Gekkonidae* (Larson and Losos, 1996). If we view toepads in functional terms, then in all three cases, a similar feature has evolved which enhances the clinging ability of lizards in their arboreal habitat.³⁹ Parallel evolution is also apparent in the world of sounds. For example, word-final devoicing of obstruents has evolved independently in at least three different subgroups of Indo-European: Romance (Catalan, Friulian), Germanic, and Slavic. Or, to take another example, consider the sound change **np > mp* which was used in 2.2 to illustrate CHANGE as a source of sound change. Many of the world's languages show homorganicity constraints on intervocalic nasal-stop sequences. In chapter 5 I show how these constraints result from the independent occurrence of CHANGE in unrelated language families. Since CHANGE has solid phonetic foundations, and can be simulated under laboratory conditions, attempts to attribute this homorganicity constraint to features of the synchronic grammar alone seem misguided. In sum, a central finding of Evolutionary Phonology is that the changes which give rise to truly similar sound patterns in genetically unrelated languages are most often the result of parallel evolution.

A final source of similarity across features are physical constraints on the form and function of the organism. Within biology, these are usually associated with aspects of an organism's growth and development, and may include anything from laws of physics relevant to embryonic development, to

biochemical processes regulating common patterns of spots and stripes on felines and shellfish (Stewart, 1998). Properties which derive from *synchronic constraint systems* are the primary object of study in most modern schools of phonology where elaborate theories of rules and representations are constructed to account for these properties.⁴⁰ Languages which prohibit heterorganic nasal-stop clusters are accounted for by featural licensing constraints on place-of-articulation. The absence of epenthesis into geminates is attributed to the structural representation of geminates as linked structures. Synchronic metathesis is expressed in terms of competing phonotactic constraints in distinct morphological environments. Though all of these analyses are claimed to reflect innate properties of synchronic grammars, as general accounts they are both too weak, and too strong. In Part II I offer alternative historical explanations for these same sound patterns based on the CCC-model of sound change.

The purpose of this book is to demonstrate that many of the similarities in sound shape across languages are best explained in terms of parallel evolution or direct genetic inheritance. Examples of parallel developments abound in the world of sounds, as they do in the natural world. Where regular sound change is involved, it is shown to have one of the three sources detailed in 2.2. With this foundation, synchronic grammars are liberated from the burden of explaining most cross-linguistic similarities in sound patterns, and can be modeled to best describe attested patterns and alternations. Synchronic constraint systems are minimal in form. They specify phonological categories, both featural, segmental, and prosodic, and they specify possible relationships between these categories. Synchronic constraint systems do not express the occurrence of similar sound patterns across languages, when these sound patterns can be shown to have their source in direct inheritance, convergent evolution, or parallel evolution. In sum, the phonological landscape is much like that in evolutionary biology: genealogy, not synchronic constraint, is the primary source of non-random distribution. Within the realm of inherited features, parallel and convergent evolution play a significant role in shaping sound patterns. Convergent evolution may give rise to similar sound patterns with distinct phonological behaviour, while parallel evolution is the primary source of phonotactic regularities in feature-distribution.

2.5 Natural selection in a world of sounds

No discussion of Darwinian evolution would be complete without mentioning natural selection. Darwin's theory of *natural selection* is built on the premise that life forms which survive to pass along their genes to future generations are those which are well adapted to the environment in which they live. An *adaptation*, in the technical sense, is a character resulting from natural selection to serve a particular biological role. As emphasized by Gould (1991:12):

The essence of Darwin's theory lies in his contention that natural selection is the creative force of evolution – not just the executioner of the unfit. Natural selection must construct the fit as well; it must build adaptation in stages, by preserving, generation after generation, the favorable

part of a random spectrum of variation... variation must be random, or at least not preferentially inclined toward adaptation. For, if variation comes prepackaged in the right direction, then selection plays no creative role, but merely eliminates the unlucky individuals who do not vary in the appropriate way.

We have already established that, like genetic variation, sound change at the level of the individual is non-optimizing, or, in Gould's terms "not preferentially inclined toward adaptation". If this is so, is there an analogue to natural selection in the world of sounds? Are sound patterns which are passed down from one generation to the next well adapted to their environment? What would adaptation look like in the world of sounds?

Before answering these questions, let us look at a clear case of adaptation in the reptilian world. A prime example of adaptation through natural selection is the case of lizard toepads mentioned earlier (Larson and Losos, 1996). The ancestors of all lizards were terrestrial and had claws, not toepads. Accurate genetic grouping of lizards and their lineages allows us to determine that toepads evolved at least three separate times in the lizard family: in *Iguanidae*, *Scincidae* and *Gekkonidae*. All lizards with toepads appear to have enhanced clinging ability, suggesting that this may be an adaptation to life in the trees (Irschick et al. 1997). Support for this view comes from the fact that in each of the three cases studied, toepad evolution is simultaneous with evolution of use of arboreal habitats. The chance probability of each instance of toepad morphology and enhanced clinging ability occurring on an arboreal lineage is $P = 0.028$, allowing Larson and Losos (1996:206) to conclude that "expanded toepads in lizards have evolved as an adaptation to increase the clinging ability in arboreal situations."

Now let us return to the linguistic questions posed above. Are sound patterns which are passed down from one generation to the next well adapted to their environment? What would adaptation look like in the world of sounds?⁴¹ One analogue to natural selection at the phonetic level are neutralizing cases of CHANGE, like the **np > mp* discussed in 2.2. In this case, a contrast in place-of-articulation for nasals is neutralized in pre-obstruent position.⁴² In this instance, misperception is a function of the higher probabilities of misperception of certain sounds and sound sequences than others which result from aspects of the human auditory system. If a contrast between two sounds is just barely perceptible in a particular phonetic environment, its chances of survival in a noisy world are slight. Notice that we are talking about contrastive categories, not individual sounds. Furthermore, contrasts must be considered within their phonetic environment. In reconsidering the case of CHANGE where [anpa] is heard as [ampa] it makes very little sense to compare the sounds [n] and [m] outside of the specific environment in which they occur. In the same sense that the usefulness of claws and toepads cannot be assessed outside of particular physical environments in which they occur, there is no sense in which /n/ is a better or more useful nasal consonant than /m/ or *vice versa*.⁴³ Adaptation occurs with respect to a specific phonetic context. In attempting to account for recurrent phonotactic patterns, then, a linguistic counterpart of natural selection may play a role precisely where misperception results from the weakness of a particular perceptual contrast in a specified phonetic environment.

The same mechanism of natural selection responsible for syntagmatic sound patterns could govern paradigmatic phonological contrasts and their phonetic instantiations. Categorical perception is the phenomenon whereby, given a range of acoustic tokens along a gradient scale, a listener identifies those at the periphery of the scale with different categories, while those in the middle have no more than chance probability of being classified into one category or the other. To take a concrete example, in many languages, what distinguishes a [p] sound from a [b] sound in the syllables [pa] vs. [ba] is voice onset time, - the time between the release of the lips for the bilabial stop and the onset of voicing for the following vowel. By presenting listeners with a range of voice-onset times, two categories, and a rough category boundary is established. Categorical perception for a wide range of phonological features (laryngeal, place, manner, tone) has been established for infants only a few days old, and has also been demonstrated for chinchillas and quail (see 9.1).⁴⁴ Humans are born with a pre-set mechanism for discriminating speech sounds, and the consonant and vowel categories of natural language can be viewed as adaptations to these settings. Phonetic instantiations of phonological categories are a subset of those distinguishable at birth, and a range of studies demonstrates that language-specific overrides of universal categorical contrasts, not their acquisition, is the hallmark of early phonological learning (Jusczyk 1992, Werker and Pegg 1992). An explanation, then, for recurrent phonetic instantiations of phonological categories is that these instantiations are an adaptation to the phonetic space defined by categorial perception. Category boundaries, where misperception is the norm, are avoided. Though I will continue to assume, based on evidence from infant perception studies, that phonological features and categories constitute part of universal grammar, arguments along these lines lead one to explore the possibility that these aspects of linguistic structure are also emergent.

At an altogether different level, the listener's acceptance of intra-speaker variation (the source of sound change in CHOICE) can be viewed as an adaptation to the limits of human vocal tract anatomy and muscular control. Coordination of distinct articulatory gestures is highly inexact in speech. Implementing a phonological sequence like /ns/ requires transition from a physical state where the velum is lowered, and there is complete oral closure produced by the tongue tip touching the alveolar ridge, to a following state where the velum is raised, and the tongue tip is slightly lowered to create turbulent noise as the air passes between the tongue and the alveolar ridge. Exact coordination of velic movement and tongue tip movement would give rise to [ns], where the velum is raised precisely at the same time as the tongue tip is lowered. This sort of precise coordination of articulators, however, is not found in natural speech. Instead, tongue tip lowering may slightly anticipate velic lowering, producing [n̥s̥], or velic lowering may slightly precede tongue-tip lowering, producing [n̥s̥]. An example of this second kind is found in English. Consider the often indistinguishable pronunciations of the words *prince* and *prints* as [prɪn̥s̥]. If listeners were unwilling to accept [n̥s̥] as a phonetic variant of /nt/, the communicative function of language would be greatly impaired. It follows that the general facility listeners have in associating careful and casual speech forms can be viewed as an adaptation to physical constraints on speech production.

Finally, it is possible that sound patterns are adaptive in terms of the mechanisms of language acquisition.⁴⁵ If certain phonological systems or generalizations are learned faster than others, and if a first workable hypothesis allows the learner to dispense with other possibilities, then phonological systems which are learned first will have the best “fitness”. In the example of CHANCE discussed in section 2.2, a phonetic string [ʔaʔ] occurs as a speaker's phonetic realization of the phonological sequence /aʔ/. However, given that the creakiness of the vowel may be a redundant feature of an adjacent glottal stop, and that creaky voiced vowels can give rise to redundant preceding and following glottal stops, the listener has a choice of phonological analyses. Did the speaker mean to say: /ʔa/, /aʔ/, /ʔaʔ/ or /aʔ/? If certain phonological generalizations are acquired more quickly than others, the choice-space may be reduced. For example, imagine that a learner has already posited a phonological constraint, based on other sound patterns, which forbids closed syllables (syllables in which a non-vocalic element follows the vowel). A grammar incorporating this constraint eliminates the closed syllables /aʔ/ and /ʔaʔ/ from the choice-space due to a pre-existing hypothesis about the form of the grammar (even if this hypothesis is ultimately not the right one.) In this case /ʔa/ and /a/ have better fitness with respect to the evolving grammatical system in the mind of the speaker.

Some specific features of sound patterns may be adaptive in more than one dimension. An interesting feature of all documented natural spoken languages is that there is never a phonological contrast between a released stop (e.g. [k[◦]]) and an unreleased stop (e.g. [k[̚]]) in word-final position. A released stop is one where oral closure is released while air is still flowing, and before intraoral air pressure has subsided. An unreleased stop involves release of oral closure after intraoral air pressure has decayed. Though this feature cannot partake in phonological contrasts, it can be controlled and specified at the phonetic level. In some languages, like Marshallese, word-final consonants are typically released. In other languages, like Cantonese, they are not. If release vs. non-release can distinguish the surface phonetics of word-final stops in different languages, why is this contrast not co-opted into phonological systems? I suggest that phonological specification of release is a highly nonaptive feature. First, consider the fact that words are most commonly produced in contexts where they are followed by other words. In a string of sounds, where a word-final stop is followed by a word-initial vowel, the smoothest articulatory trajectory involves releasing the stop into the following vowel. On the other hand, where a word-final stop is followed by another stop, the smoothest articulatory trajectory involves an unreleased stop followed by a released stop. Second, intra-speaker variation can be viewed as a continuum from hyperarticulated clear speech to hypoarticulated casual speech as already discussed. If clear speech serves the purpose of eliminating potential ambiguities created by the coarticulatory properties of casual speech, then clear speech is a domain where release can serve a very useful function. Since unreleased stops lack phonetic release features which may instantiate phonologically contrastive features of aspiration, glottalization, and major place of articulation, among others (see chapters 4 and 5), limiting the release/non-release opposition to the phonetic domain allows

for recovery of potentially contrastive phonological features in hyperarticulated speech. In actual usage, this is the norm: "His name is [mæt̚] not [mæk̚]!", where [mæt̚] is produced with an audible release of the /t/. The absence of phonological contrasts between released and non-released stops can be viewed as an adaptive feature of sound systems in at least two respects then. First, it is adaptive in articulatory terms, since the smoothest transitions between stops and following vowels involve release, while those between stops and following stops do not. If release is specified independent of the sound which follows, ease of articulation will be hampered. The second way in which purely phonetic release is adaptive is in terms of its discriminating function in hyperarticulated speech. All spoken languages contrast stops at major points of articulation, and many make laryngeal contrasts as well. Though phonetic cues for these features may be lost in contexts where ease of articulation leads to absence of release, they can be recovered in careful speech where release is introduced.

Just as certain sound patterns look highly adaptive, so others smack of *nonadaptation* or *disadaptation*. In evolutionary biology, if a character cannot be distinguished from its genetic antecedents with respect to conferring an advantage to the organism possessing it, it is called a nonadaptation. Disadaptation refers to the case where the genetic antecedent has greater utility to the organism than its evolved counterpart (Baum and Larson, 1991; Arnold 1994).⁴⁶ Both of these categories suggest characters whose evolution is inconsistent with natural selection, and there are few if any convincing examples of this in the biological literature. In speech, we can look for such cases at the limits of perceptibility, where nonadaptive features may arise fleetingly, and just as quickly, disappear. One recently verified case is the contrast between two degrees of nasalization in Palantla Chinantec, an Otomanguean language. In this language, a contrast exists between oral vowels, lightly nasalized vowels, and heavily nasalized vowels. Palantla Chinantec is the only reported language in the world in which degree of nasality (lightly vs. heavily) is phonologically contrastive. Interestingly, speakers have difficulty perceiving differences between lightly and heavily nasalized vowels, and this contrast might also stretch the limits of velic control. Nevertheless, the contrast exists. It was first reported by Merrifield (1963:5):
"...certain idiolects of Palantla Chinantec exhibit two degrees of nasalization which occur in identical environments, thus defining lexical contrasts." Subsequent confirmation comes from Ladefoged (1971) and Merrifield and Edmondson (1999) who demonstrate a contrast between oral syllables, heavily nasalized syllables, and lightly nasalized syllables. The contrast between heavily and lightly nasalized syllables involves two quantifiable parameters: in heavily nasalized syllables, nasal airflow begins shortly after the syllable-initial consonant, and continues at a relatively steady level through the syllable, while in lightly nasalized syllables, low volume nasal airflow begins roughly halfway through the nuclear vowel and gradually increases through to the end of the syllable. While Palantla Chinantec is the only language known to contrast degree of nasality, not all speakers of the language exhibit two degrees of nasalization in their speech. For some speakers, oral and lightly nasalized vowels fall together as oral, while for others, lightly and heavily nasalized vowels merge into a single nasalized category. The contrast is dying, confirming the suspicion that it is disadaptive, or at least nonadaptive. But, if natural selection does play a role in language evolution, how did this contrast arise in the first place? Two grammatical relationships may have played a role. First, heavy nasalization is associated with

animacy in some lexical pairs. Second, light nasalization evolved in some verbs from the inflectional person-number suffix /n/, via the sound change *VN > lightly nasalized vowel. In this case the phonetic feature of light nasalization may have been temporarily bolstered by its association with a specific morpho-syntactic feature.⁴⁷ As these grammatical distinctions erode, so does the phonological contrast: for many speakers of Palantla Chinantec today, there is only a binary contrast between oral vowels and nasalized vowels, as in so many other languages.

In this section I have suggested some ways in which natural selection may play a role in the evolution of sound patterns. The clearest cases with biological parallels are neutralizing cases of CHANGE and recurrent phonological partitioning of the phonetic perceptual space. Intra-speaker variation defined in terms of hyper- and hypo-articulated speech can also be seen as an adaptation to the imprecise muscular control involved in articulation of speech sounds. I have also pointed to the possibility that sound patterns could be adaptive with respect to grammatical generalizations arising in the course of language acquisition. A feature which might be viewed as adaptive in its ease-of-articulation function and maximize-contrast function at the phonetic level is the apparent universally non-contrastive status of stop-release. In this case, importation of release from the phonetic to the phonological domain would arguably result in a nonaptive feature. As far as we know, this has never occurred. However, nonadaptation may be visible in linguistic systems. The Palantla Chinantec distinction between three degrees of nasalization is arguably nonaptive in contrast to simple nasal/oral contrast from which it descends. In this case, associations between these phonological categories and semantic or morpho-syntactic features may have given rise to strengthening of the weak feature. However, subsequent decay of these associations has left the degree of nasalization contrast on the verge of extinction.

As Rose and Lauder (1996:9) point out in the introduction to their volume on modern views of adaptation:

Adaptation is no longer something that can be safely assumed by evolutionary or other biologists. Indeed, the more one examines the concept, the more it comes to resemble a newly landed fish: slippery, slimy, obstreperous, but glittering with potential. There it is, flapping about, full of energy, but the significance of all the commotion is not clear.

In the same way, adaptation cannot be assumed within Evolutionary Phonology. It may further our understanding of CHANGE, but be irrelevant to sound change arising from CHANCE or CHOICE. The question is an empirical one, and throughout this volume, we will consider its glittering potential and, at the same time, question its significance.

Throughout this chapter concepts in evolutionary biology have been used as metaphors for language change. I do not claim that language evolution is structurally identical to biological evolution, nor that Darwinian principles of natural selection can be applied to living organisms and languages in precisely the same way. One purpose of this comparison is to make a clear distinction between the mechanisms of change spelled out in 2.2, their non-optimizing character outlined in 2.3, and possible instances of

natural selection discussed above. Another important role of this comparison is to emphasize the four logical sources of resemblance between sound patterns laid out in 2.4. Similarities in sound patterns are often the result of direct genetic inheritance. In other cases, recurrent sound patterns may reflect convergent or parallel evolution. Finally, there is the possibility that regularities in sound patterns are a direct result of synchronic constraint systems.

From this point on, the biological metaphor will, for the most part, be left behind. While it is useful in understanding why certain patterns are common across languages, its usefulness breaks down when key mechanisms of change are compared. There are no direct counterparts to genetic mutation, biological inheritance, or natural selection in the sound changes which characterize language evolution for the last 7,000-8,000 years. Instead, it is more instructive to study the phonetic basis of sound change, the transmission of a learned code across generations of individuals, and the aspects of listener bias and learning algorithms which might limit the phonological choice space. Sound patterns have their own natural history, and it is this history which is documented in Part II.

Notes to Chapter 2

¹ Thurston's model of linguistic gradualism and punctuated equilibrium as applied to the languages of Northwestern New Britain differs from that popularized by Dixon (1997).

² The split of a mother language into distinct dialects, and eventually, distinct languages, correlates most closely with geographic and social isolation. This is evident in much of Amazonia, New Guinea, and equatorial Africa which have more languages per square mile than most other parts of the world, or did, until times of contact. Contact-induced language death has had a significant effect on linguistic diversity in these geographic areas, as elsewhere.

³ See Dalby (2002, chapter 1) for a succinct summary of differences between biological evolution and language change, with special reference to language contact.

⁴ Fitch (2000) reviews evidence suggesting that Neanderthals may have had greater capacities for speech than attributed to them by Lieberman (1984).

⁵ Whether or not these sound changes spread throughout the speech community depends on a range of social factors which will not concern us here. See Labov (2001) for a comprehensive survey of the issues involved.

⁶ A non-metaphorical evolutionary model of the emergence of sound patterns can be found in Steels (1998) and De Boer (2001). Jakobson (1929) and Kiparsky (1995) propose similar models, suggested, but with the selectional mechanism of language transmission heavily determined by phonological markedness principles which are not assumed within Evolutionary Phonology.

⁷ See also Briscoe (2000) on the coevolution of language and the language acquisition device.

⁸ See Chapter 2 of Norman (1988) for a concise history of Chinese historical phonology. The quote at the beginning of this chapter is taken from Norman (1988:42).

⁹ This is not to say that grammars change little after the age of five. Vocabularies continue to increase over a lifetime, and significant grammatical changes can take as well, but are typically associated with sociolinguistic variables.

¹⁰ Ladefoged and Maddieson (1996:326) list only three other languages with an underlying /w/ vs. /ʍ/ contrast: Klamath, Yao, and Aleut.

¹¹ For documentation of this sound change in progress in American English, and a similar analysis, see Locke (1983:206-08).

¹² See Labov (1994:328-29) for multiple variables which may play a role in mergers. In Scotts English, the maintenance of the contrast may be related to the feature of pre-aspiration which is absent in dialects where merger occurred.

¹³ For discussion of some potential cases of regular sound change without identifiable phonetic explanations, see Blust (2003).

¹⁴ Small capitals are used to distinguish these technical terms from their common usages. When referring to the formal model of phonetically based sound change in (1), I will use the term ‘CCC-model’ where CCC abbreviates CHANGE, CHANCE and CHOICE. Also, from this point on, I will use ‘sound change’ to refer to phonetically motivated sound change, referring specifically to other types of change (e.g. lexical diffusion, analogy, contact-induced change) where necessary.

¹⁵ This imbalance leads to asymmetries in misperception: [anpa] is frequently misheard as [ampa], but not as [anta]. See Chapter 5 for further discussion of related asymmetries.

¹⁶ These examples from different children are taken from Appendix 9.1 of Vihman (1996:238-239).

¹⁷ Any phonetic string in isolation is ambiguous with respect to multiple phonological representations. However, context-dependent regularities in phonetic realization of phonological categories within and across words and phrases typically allow the language learner to zero in on a phonological analysis which matches that of the speaker. We are still at any early stage in understanding precisely how such generalizations are extracted from the speech stream. Relevant studies of speech perception are cited throughout this volume, and a range of perspectives is presented in Ferguson et al. (1992).

¹⁸ Difficulties in feature localization may also arise as a result of auditory decoupling which may occur in the perception of sibilant and click noise (Blevins and Garrett, to appear).

¹⁹ See 6.4 and 9.2 where these priming effects are attributed to Structural Analogy.

²⁰ This is part of the H&H theory of Lindblom (1990a). Evolutionary Phonology adopts the H&H *description* of intra-speaker variation as a function of the hypo- to hyper-articulated speech continuum. However, unlike H&H theory, it does not view sound change as ‘improving’ speech along this continuum. Rather, a stochastic model is adopted in which changes in frequency of particular phonetic variants result in shifts along the hypo-hyper-articulation axis.

²¹ The non-contrastive status of release in the world's languages is discussed in 2.5 in the context of nonapertive features of speech.

²² Within some models of phonologization, a distinction is made between universal phonetic properties, and language-specific ones. For example, Hyman (1977) identifies a state where a language has automatic lowering of F_0 following a voiced consonant due to universal phonetics, and a subsequent state where lowering of F_0 is too great to be attributed to universal phonetics. Hajek (1997:21) also distinguishes language-specific phonetic rules from universal phonetic rules. In the model proposed in (2), there is no formal difference between variants determined by universal phonetics and those determined by language-specific properties; in both cases, a set of variants is defined from which the listener *qua* learner must acquire language-specific phonetics, and at the same time, extract from acoustic data significant sound patterns and contrasts. Within this model, the enhancement of F_0 lowering following a voiced consonant is no different in nature from the enhancement of a pre-existing language specific feature, like the extreme vowel lengthening found in English preceding voiced consonants. The model in (2) is meant to capture subphonemic sound change as well as phonologization.

²³ For evidence supporting the learned status of detailed phonetic knowledge, including patterns of variation, see Pierrehumbert (1994, 2000, 2001), Bybee (2001) and Warner et al. (2002).

²⁴ Kiparsky (1988:373), citing Jespersen (1886/1993) and Nyman (1978) attributes the causal link between frequency and sound change to redundancy: "frequent items are more easily guessed by the hearer, so the speaker can afford more reduced pronunciations of them, which then may be lexicalized." In other words, reduced variants of high frequency words will be more common than reduced variants of lower frequency words, but reduced variants of lower frequency words may still occur. In order to account for the 'lexicalization' referred to by Kiparsky, frequency effects are incorporated in the model of sound change.

Many factors can give rise to changes in the frequencies of phonetic variants from one generation to the next, from changes in use of lexical items to increased use of particular genres which occupy one extreme of the hypo-hyper-articulation continuum. These frequency changes give rise to the seemingly non-deterministic pathways of change (e.g. lenition vs. fortition) observed through this volume.

²⁵ Hajek (1997:21) acknowledges variation in production as a constant which "may result in gradual, subtle but ultimately cumulative shifts in articulatory and acoustic targets" but does not build this directly into his model or typology of sound change.

²⁶ The proponent of optimizing sound change can always revert to the position that *RV* is better in Slavic, while *VR* is better in Le Havre. The problem, discussed further in 10.3, is that from the same proto-language, different languages evolve, while from the same language, different dialects emerge.

If *VR* is better in Le Havre, why isn't it also better in French? See 6.3 and 9.2.2 for a non-teleological account of structure preservation in sound change.

²⁷ Lindblom (1986, 1998) suggests an equation for optimal systems of phonetic contrasts, balancing articulatory ease with perceptual ease. However, in his equation the size of the inventory is pre-determined. In natural language learning, we cannot assume that the listener arrives at inventory size before segmental contrasts are discovered.

²⁸ An articulatory argument might also be made, since [np] requires two articulatory gestures, while [mp] has only one. However, as pointed out by Ohala, this account cannot explain why sound changes like **np > mp* involving regressive assimilation are extremely common, while shifts of **np > nt*, involving similar articulatory simplification but progressive assimilation, are rare.

²⁹ It is instructive in the phonological domain, because both nasality and aspiration are usually viewed as marked properties of vowels. And it is instructive in the phonetic domain because both aspiration and nasality shrink the perceptual vowel space.

³⁰ Physical environment appears to play a very minor role in shaping language. Multiple languages (so far as one can tell, any possible subset of natural languages) can be acquired by a single child in any part of the world. This context-free gift for multilingualism signals the essential independence of language structure and transmission from broad aspects of the physical environment and the human genetic code. Nevertheless, there are interesting speech modes which appear to have evolved relative to aspects of the physical environment. One example is whistled speech, where words are whistled instead of spoken. In Pirahã, an Amazonian language, whistle speech is used during rainforest hunts (Everett, 1986). The rainforest is arguably one of the noisiest natural environments encountered, and so, it is not surprising that a speech mode with better overall acoustics than the spoken word has arisen in its depths.

³¹ See chapter 3 for a brief general history of explanation in phonology.

³² Within generative models, direct genetic inheritance has relevance only to the periphery of the grammar, and no relation to its core, which includes most, if not all, commonly recurrent grammatical properties. Regularity in form is not the result of inheritance. Only peculiarities of grammar are viewed as "historical residue". As Lightfoot (1999:13) puts it: "We shall look to history...to understand the quirks of the modern language."

³³ See Ringe (1999) where the problem of random cognate sets for CVC-roots is detailed.

³⁴ Exceptional lexical classes with non-arbitrary sound-meaning correspondences include sound-symbolism, baby-talk, and onomatopoeia.

³⁵ Of course, there is no direct analog to borrowed features in the biological world, though hybridism and new technologies in inter-species organ transplants come close.

³⁶ In Japanese, the only nasal found word-finally is a nasal glide. In Gilbertese /m, n, ŋ/ are all licit word-finally.

³⁷ Though the argument is made that instances of convergent evolution reflect a general cross-linguistic preference for nasal codas. See, for example, Herbert (1986, 6.1.4), and the discussion in 6.6.

³⁸ In the field of evolutionary biology, parallel evolution contrasts with convergent evolution. In convergent evolution two independently evolved features which are superficially similar arise by different developmental pathways from different ancestral conditions. As mentioned earlier, eyes of vertebrates and eyes of cephalopods (octopus, squid, etc.) are a case in point. See Larson and Losos (1996) for precise phylogenetic definitions. I make a similar distinction between convergent and parallel evolution in language.

³⁹ At the structural level, the toepads are only superficially similar, and constitute a case of convergent evolution (Larson and Losos, 1996:192).

⁴⁰ The terms 'constraint' and 'rule' are used interchangeably in reference to phonological theories, unless the precise issue of constraint systems vs. rules systems is the focus of discussion.

⁴¹ Recall that I am limiting myself to study of language evolution over the past 7,000 years or so. For discussion of language as an adaptive feature at the level of the human species, and 19th century views on its evolution, see the collection of essays in Harris (1996). For a wide range of modern perspectives on language evolution and the emergence of phonology and syntax see Hurford et. al. (1998).

⁴² Observe that not all cases of CHANGE are neutralizing. For example, with rhinoglottophilia, misperception of breathiness as nasalization or nasalization as breathiness does not alter the number of phonological contrasts in the language. In simply shifting from one phonetic feature analysis to another, there is no principle of natural selection at work. Breathiness and nasalization are both well adapted to the phonetic environments in question, as indicated by the retention of these features as a set. They are simply so perceptually similar that, without further phonetic enhancement, they are easily confused.

⁴³ The claim has been made that coronal consonants are phonologically unmarked, and that /n/ is better than /m/ in some absolute way. In 5.4 I argue against this position.

⁴⁴ Categorical perception looks biologically ancient. So does left hemisphere specialization for vocal communication, which has now been demonstrated in monkeys, mice, hamsters, birds, and frogs (Studdert-Kennedy 1998:173).

⁴⁵ The obvious fact that all natural languages are learnable (by children before the critical age) could also be attributed to adaptation, in the form of co-evolution, though it could just as well be accidental. There is no evidence that the process of language acquisition has changed significantly over the past 7,000 years, nor is there any obvious external source capable of introducing unlearnable systems into the domain of natural language. So, for recent history, there is no plausible way of testing this hypothesis. However, calculations like those of Lindblom (1986) and Lindblom and Maddieson (1988) are highly suggestive of this type of co-evolution as a constant force shaping vowel and consonant systems respectively.

See Bever and Langendoen (1972) on language change as an interplay of ease of learning and ease of perception.

⁴⁶ In biological models, it is important to distinguish primary and secondary nonadaptation and disadaptation. Primary nonadaptation and disadaptation are cases where a character replaced one of equal or superior utility at its origin; secondary cases occur by evolutionary changes in selective regime (Baum and Larson, 1991; Arnold 1994).

⁴⁷ Relations between phonological contrast and the functions of these contrasts within the grammar as a whole are one focus of the Firthian school of prosodic phonology (e.g. Firth 1948, Henderson 1949). In 8.3 the potential role of paradigms in contrast maintenance is explored.