Chapter 9

The phonetics of sound change

John Ohala

1 Introduction

1.1 Strengthening the comparative method

One of the few solid scientific accomplishments of linguistics is the discovery of ways to reconstruct the history of languages via the comparative method. With it one first establishes sets of cognate words or morphemes in different languages and then posits an optimal route between them which consists of a hypothetical parent form and sound changes that transformed the parent forms into the attested cognates. Though not usually thought of in this way (but see Young 1819), the comparative method is a quasi-mathematical operation, involving implicit (and qualitative, not quantitative) estimations of probabilities of events and what is, in effect, the application of optimization theory. The posited reconstructed forms and the sound changes must also be within the bounds of the plausible, where plausibility is determined inductively, that is, by what the linguist has previously encountered in other human languages.

The quasi-mathematical character of the comparative method permits the linguist to treat words and their constituent sounds more or less as abstract algebraic entities without having to worry too much about their physical substance. They can be and have been manipulated like variables in an equation. Indeed, the leading historical linguists often emphasize that the reconstructed forms are simply parts of formulae for relating sets of cognates and should not be regarded as representing phonetic structure (Bloomfield 1914, 274ff; Meillet 1964, 39ff.). In addition, the inductive constraints on posited reconstructed forms and the sound changes that apply to them do not require that the linguist actually understand why languages are structured as they are or behave as they do, all that is necessary is to be aware of structure and behaviour frequently encountered. Thus, the actual phonetic substance of reconstructed forms and the mechanism of sound changes is usually of secondary importance.

In this chapter I argue that it is possible to do better historical phonology by taking into account the mechanism of sound change. This involves integrating phonetic studies with historical phonology. Of course, I am not claiming that 'complete' reconstructions of past languages are possible; just that the reconstructions can be improved by an attempt to understand the factors which give rise to sound change. In fact, my point is similar to ones made throughout the past few centuries: de Brosses (1765), von Raumer (1863), Key (1855), Osthoff and Brugmann (1878), Rousselot (1891) – to mention only a few – have all insisted and sometimes demonstrated that we could understand language change better by paying more attention to the phonetic and psychological aspects of change. This chapter, then, can be considered as an attempt to bring historical phonology up to date with current data and concepts in phonetics and psychology.

1.2 Delimiting the scope of the discussion

It is well recognized that changes in pronunciation can come about through many quite different factors, including some, such as spelling pronunciation, paradigm regularization, and fashion, which are language and culture-specific. In this chapter I consider only sound changes which have been attested independently in substantially the same form in many unrelated languages and which, therefore, are most likely to arise from language universal factors, i.e. physiological and psychological factors common to all human speakers at any time. Henceforth I will use the unqualified term 'sound change' to refer to these common, frequently encountered sound changes.

A further delimitation on the discussion to follow is that I am primarily concerned with the preconditions for sound changes, not their actual trigger and not their subsequent spread through the lexicon, through the dialect community, and from one speech style to another. Thus I will not be concerned with questions such as 'why did this sound change occur in such-and-such language at such-and-such time?' although I will suggest that such questions are for the most part fruitless pursuits.

2 Phonetic discoveries of importance to sound change

There are two major discoveries in phonetics which form the starting points for the subsequent discussion.

2.1 The first point: the infinite variability of speech

One of the major discoveries of phonetics over the past century is the tremendous variability that exists in what we regard as the 'same' events in speech, whether this sameness be phones, syllables, or words. Some variability was evident to phoneticians simply from ear analysis, for example, as exemplified by the work of Pāṇini, Amman (1700), Ellis (1889), and Sweet (1877), but more and more variation was noticed through instrumental analysis, starting in the late nineteenth century with kymographic studies and accumulating markedly with the onset of acoustic studies in the 1930s and 1940s. It is now accepted that there is infinite variation in speech – though still more or less lawfully determined. The 'same' sound is measurably different not only when spoken by the different speakers (which might be expected) but also when spoken by the same speaker in different phonetic environments or at different rates or levels of loudness (Lindblom 1963; Traunmüller 1981). The typically short list of 'allophones' given in traditional phonemic inventories of languages does not begin to give the whole story of the amount of variation present in speech.

2.2 The second point: the parallels between phonetic variation and sound change

That some of the synchronic variation in speech is similar to sound change has long been noted, for example, vowels are commonly non-distinctively nasalized before nasal vowels and that is the environment which most often gives rise to distinctively nasal vowels via sound change. But as instrumental studies and perceptual studies of speech accumulated, more and more points of similarity were noticed. In the next two sections I give some examples first from the domain of speech production and then from the acoustic-auditory domain.

2.2.1 Variation in the domain of speech production

2.2.1.1 Tonal development and consonantally induced FO differences on following vowels

Edkins (1864) and Maspéro (1912) and subsequently several other researchers noted that in East and South-east Asian languages[1] certain tonal distinctions developed out of former (subsequently neutralized) voiced vs. voiceless contrasts on pre-vocalic consonants; a higher tone developing after what had been the voiceless consonant and a lower tone after the voiced. An example from closely related dialects of Kammu, described by Svantesson (1983), is given in (1).

(1) Data showing that tone in Northern Kammu corresponds to a voicing contrast in Southern Kammu (from Svantesson 1983: 69).

Southern Kammu	*Northern Kammu*	*Translation*
klaaŋ	kláaŋ	eagle
glaaŋ	klàaŋ	stone

Parallel to this is the discovery documented via acoustic analysis for many diverse languages (see, e.g. Sato 1950; Hombert 1978; Hombert et al. 1979) that the fundamental frequency (FO) on vowels is higher following voiceless consonants than voiced; see Figure 9.1.

Löfqvist et al. (1989) have discovered the apparent physiological cause of this effect: during voiceless stops there is a higher contraction rate for the cricothyroid muscles in the larynx, the chief tensor muscle and the one most directly involved in regulating the FO of voice. This increased tension of the vocal cords may serve to insure the voicelessness of the stops by stiffening the vocal cords.[2] The concomitant FO difference on the following vowel is therefore probably a fortuitous consequence of this activity directed towards maintaining the voicelessness of the consonant, at least there is no evidence yet to suggest that the FO perturbation is purposeful.

2.2.1.2 Spontaneous nasalization

Bloch (1920, 1965), Turner (1921) and Grierson (1922), studying Indo-Aryan languages, have called attention to what they call 'spontaneous nasalization', i.e. the development of distinctive nasalization on vowels in words that never had any lexical nasal consonant (the usual source of nasal vowels). One type of segment that reappears in many of their examples is one characterized by high airflow, e.g. any voiceless fricative, especially [h], aspirated stops and affricates (Ohala 1983a); see (2).

(2) Examples of spontaneous nasalization (from Grierson 1922).

Sanskrit	*Prakrit*	*Old Hindi*	*Modern Hindi*	*Bengali*	*Translation*
pakṣ	pakkha	pākh	paŋkhā		a side
akṣi	akkhi-		ā̃kh		eye
uččaka-	uččaa-		ū̃čā	uɲča	high
satya-	sačča-	sā̃č			truth
sarpa-	sappa-		sā̃p		snake

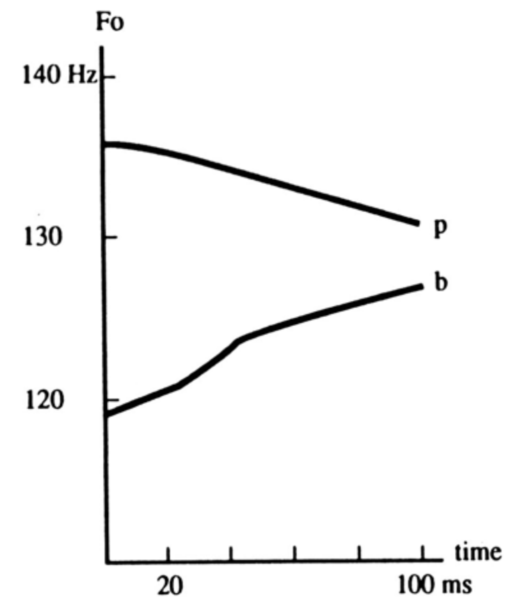

FIGURE 9.1 Average fundamental frequency values (in Hz) of vowels following English stops; data from five speakers. The curves labelled 'p' and 'b' represent the values associated with all voiceless and voiced stops, respectively, regardless of place of articulation. (From Hombert, Ohala & Ewan 1979.)

The same phenomenon exists in other languages; see Ohala (1975); Matisoff (1975). Ohala and Amador (1981, summarized in Ohala 1983a) studying both American English and Mexican Spanish, found that high airflow segments like voiceless fricatives have a greater-than-normal glottal opening which is partially assimilated by adjacent vowels which, in turn, creates an acoustic effect which mimics nasalization. Such pseudo-nasalization, they argued, is liable to be misinterpreted by listeners as actual nasalization and reproduced by them as such.

2.2.2 *Variation in the acoustic-auditory domain*

The preceding examples described variation found in the domain of speech production and in the conversion of articulation into sound. Variation has also been found in the perceptual domain.

2.2.2.1 [θ] and [f]
Sweet (1874: 8), among others, noted the frequent confusion and consequent substitution of [f] for [θ] in dialectal English, e.g. [θɪŋ] ~ [fɪŋ] 'thing'; [θɹu] ~ [fɹu] 'through'. Parallel to this is the frequent confusion of the same sounds in the consonant confusion study of Miller and Nicely (1955) under virtually all conditions of signal-to-noise ratio and filtering.

2.2.2.2 Labial velars and labials
Some confusions occur due to the acoustic similarity of sounds that are very different in articulation. The substitution of labial (or labialized) velars by labials is well known in the development of Classical Greek from Proto-Indo-European as well as many other unrelated languages. Examples are given in (3)

(3.a)	*Indo-European*	> *Classical Greek*	
	*ekwōs	hippos	'horse'
	*g^{w}iwos	bios	'life'
	*yek^{w}r̥	hepatos	'liver'
(3.b)	*Proto-Bantu*	> *West Teke*	
	*-kumu	pfuma	'chief'
	Proto-Yuman	> *Yuma*	
	*imalik^{w}i	mal^{y}pu	'navel'
	Proto-Muskogean	> *Choctaw*	
	*k^{w}ihi	bihi	'mulberry'
	*uNk^{w}i	umbi	'pawpaw'
	Sungkhla (free variants)		
	\khwai	\fai	'fire'
	/khon	/fon	'rain'
	Proto-Zapotec	> *Isthmus Zapotec*	
	*kk^{w}a-	pa	'where'

Again, there are striking phonetic parallels. Durand (1955) commented on the acoustic similarity of these two classes of sounds. Winitz et al. (1972), studying listeners' identifications of CV syllables (formed from the set C = [p t k], V = [i a u]), heard under a variety of conditions, found that confusions between the syllables [ku] and [pu] (where one can assume that the [k] was phonetically labialized before the rounded [u]) were among the highest of any of the CVs included.

2.2.2.3 Vowel quality shifts under nasalization
Although there are some complexities and puzzles still to be solved in the case of the phonetic and perceptual consequences of

adding nasalization to vowels, it is fairly well established that distinctively (i.e. non-contextually) nasalized vowels tend to be lower than their corresponding oral counterparts or the oral vowels from which they sprang (Beddor 1983). French presents an extreme case of this as evidenced by the cognate pairs in (4).[3]

(4)	Oral	Nasal
	linoleum [linoleym] 'linoleum'	lin [lɛ̃] 'flax'
	brune [bʁyn] 'brunette'	brun [bʁœ̃] 'brown'

Paralleling this is the finding by House and Stevens (1956) (with an important qualification added by House 1957) and supported by data presented by Fujimura and Lindqvist (1971) and Beddor et al. (1986)) that coupling the oral and nasal resonators leads to an elevation of the first formant [F1] in non-low vowels (which means a lowering of perceived vowel height). Wright (1986) and Beddor et al. (1986) confirmed that perceptually non-contextually nasalized vowels sound lower than their articulatorily equivalent oral counterparts.

3 The implications of parallels between variation in the synchronic phonetic and the diachronic domains

3.1 Synchronic variation = sound change?
I reviewed above just a few examples out of many possible[4] that demonstrate the parallels to be found between sound change, on the one hand, and synchronic phonetic variation, both in the production and in the acoustic-auditory domains, on the other. Does this mean that the phonetic variation is sound change? The answer to this question is somewhat complex.

3.1.1 *Variation in perception may equal 'mini' sound change*
In the case of the variation in perception, caused by the confusion of acoustically similar (but sometimes articulatorily different) speech sounds, such variation is potentially sound change. At least, we could refer to this as a 'mini sound change' – one that takes place in the interaction between a single speaker and a single hearer. All that would be necessary for such mini sound changes to become the 'maxi' sound changes that get recorded in the historical grammars of languages would be that some of the subjects in the Miller and Nicely (1955) or the Winitz et al. (1972) studies leave the laboratory and to start pronouncing the name of the thing that had been presented to them in the laboratory as [θa] or [ku] as [fa] and [pu], respectively, and then for other speakers to copy their pronunciation. Anecdotal

evidence suggests that the kind of misperceptions that these researchers found in the laboratory setting go on all the time in ordinary exchanges between speakers and hearers. What keeps the vast majority of these from leading to maxi sound changes are factors such as the following:

a) Pronunciation norms are redundantly represented in all speakers in the given language community. A listener who misapprehends a word spoken by one speaker has many opportunities to hear the word spoken by others and presumably would not be likely to mishear in all cases. In addition, others' puzzled reactions or amusement can alert a speaker to his odd pronunciation. In some cultures, orthography provides yet another guide to pronunciation.
b) It is probably a rare thing for one speaker's innovative pronunciation to spread via imitation to sizable numbers of other speakers such that it becomes a persistent and characteristic feature of a well-defined dialect community.

Nevertheless, some small fraction of misapprehensions of pronunciation can and do go uncorrected and can spread to other speakers and thus give rise to maxi sound changes. The essential aspect of misperceptions which make them equivalent to sound changes is that they constitute a *change of norms*: the listener forms a phonological norm that differs from that intended by the speaker.

3.1.2 Variation in production does not automatically equal sound change

As for variation in speech production, e.g. the FO differences on vowels after voiced and voiceless consonants, these do *not* by themselves constitute sound change precisely because they do not necessarily involve a change in pronunciation norm. There are several reasons for asserting this.

First, it is not clear that speakers intend primarily to produce these effects; rather they are consequences of other effects which the speaker intends. Of course, it is difficult to get an accurate picture of the speaker's intentions in speech production and, accordingly, opinions differ on this point (cf. Stevens et al. 1986; Stevens and Keyser 1989) but there is some evidence that at least some of the variation seen in the speech signal can occur without involving a corresponding modification in the speaker's underlying pronunciation target (Ohala 1981; Browman and Goldstein 1990).

Secondly, most of the types of variation have been found to occur widely in virtually all languages studied instrumentally. Phonetic features that are universal are likely to be physically caused and not maintained by culturally established templates.

Thirdly, in many of these cases it is possible to explain the occurrence of such variations by reference to physical phonetic theories. For example, Zue (1976) found that the centre frequency and the amplitude of noise at stop burst for [g] tends to be higher for front vowels than back vowels. This is explicable from basic acoustic and aerodynamic principles (Fant 1960; Stevens 1971). Other examples could be given, for example, the tendency for velars to devoice (Ohala 1983b, 1991).

Fourthly, and perhaps most importantly, there is evidence that listeners can normalize predictable variation: one of the principal acoustic differences between /s/ and /ʃ/ is the lower centre frequency of the latter. But the centre frequency can also vary due to contextual influences, for example, anticipatory assimilation of lip-rounding from a following rounded vowel (which serves to lower the frequency). Mann and Repp (1980) found that a synthetic fricative that would be identified as /ʃ/ before a following /ɑ/ is identified as [s] when the following vowel is [u]. Listeners presumably expect the lowering of the centre frequency of an /s/ before rounded vowels like /u/ due to anticipatory assimilation. Thus, when the sibilant has a lower centre frequency and the following vowel is heard as round, they apparently factor out the expected low frequency and 'reconstruct' a higher centre frequency characteristic of /s/. Similarly, Kawasaki (1986) presented evidence that listeners can factor out some of the nasalization on vowels when there is a nasal consonant nearby. Beddor et al. (1986) demonstrated that listeners can factor out the predictable lowering effects of nasalization on vowel quality when a nasal consonant is immediately adjacent to the vowel. Another way of describing such behaviour is to say that listeners *normalize* or *correct* the speech signal in order to arrive at the pronunciation intended by the speaker minus any added contextual perturbations. In the following discussion I will use the term 'correction' to describe this process of perceptual normalization.

Such perceptual correction of the speech signal by listeners serves to *prevent* sound change, that is, to prevent the change of pronunciation norms between speaker and listener even though the speech signal exhibits an incredible amount of variation. Thus, finding variation in speech production that parallels diachronic variation does not permit us to say that such synchronic variation *is* sound change.

3.2 The ubiquity of perceptual correction

Perceptual correction of a distorted, variable stimulus is not unique to speech and, in fact, is well known in other sensory modalities. Under the topic of 'perceptual constancies' it has been posited and studied extensively in the visual domain (Rock 1983). How is it that we are able to extract a constant percept of scenes that we view in spite of their showing variations in size, colour, shape, and so on, depending on the conditions under which they are seen? A person seen from a distance subtends a small angle in the visual field for simple physical reasons that are easily explained by geometry. Thus a person seen from a distance might appear physically to be as large as one's thumb. But the true size of the person can be reconstructed if the viewer can exploit any of a number of cues that the person is far away, e.g., parallax cues, interposition of other closer objects, perspective (convergence of the parallel lines), the texture of the scene (the degree of detail visible), etc.

In fact, as regards correction of the speech signal, there is evidence that listeners are able to do it in the case of distortions imposed by the channel over which it is transmitted, for example, with certain frequency bands filtered out or with noise superimposed (Ohala and Shriberg 1990; Shriberg and Ohala 1991; these are discussed below). Much the same ability must underlie how we learn to understand different dialects or foreign accents of languages we know.

3.3 How variation in production can lead to sound change: hypo-correction

If the variation observed in the speech production domain does not constitute sound change, then what are we to make of the parallels between such variation and the diachronic variation? The answer is simple: if the listener fails to correct the perturbations in the speech signal, then they will be taken at face value and will form part of his conception of its pronunciation. Via such 'hypo-correction', as I call it, the phonetic perturbations, originally just fortuitous results of the speech production process, become part of the pronunciation norm. This, presumably, is what is meant by the term 'phonologization'.

Why would a listener fail to correct a perturbed speech signal? First, it must be emphasized again that such failure represents a very small fraction of all the interactions between speaker and listener. That said, it is possible to speculate about several possible situations where it might happen. First, the listener may not have the experience to enable him to do such correction.

Children in the process of acquiring the phonology of their language are in this position as are adult second-language learners. A propos of this, it is interesting to note that in linguists' phonological descriptions of languages, one frequently finds longer lists of allophones of phonemes when the linguist is *not* a native speaker of the language as opposed to when he is. Could it be that the non-native speaker linguist notices more variation because he lacks the experience the native speaker has which allows the perceptual factoring out of predictable variation?

A second reason for hypo-correction is that a listener may, for various reasons, fail to perceive or to attend to the phonological environment which causes, or as phonologists usually put it, 'conditions' the variation. Thus it is frequently the case that sound changes which are the result of assimilations and other effects assignable to the domain of speech production take place with the simultaneous loss of the conditioning environment, as in (1), where the tonal development was accompanied by loss of the voicing distinction. Many (but certainly not all) of the most common sound changes have this feature, e.g. the development of distinctive nasal vowels is generally correlated with the loss of the nasal consonant which conditioned the change from oral to nasal in the vowel. This is an important aspect of many hypo-corrective sound changes and more will be made of this.

We may also speculate that the farther away the conditioning environment is from the conditioned change – that is, the greater the temporal gap between cause and effect – the more difficult will it be for the listener to be able to establish the causal link between the two and use this link as the basis for correction.

3.4 Shorthand notation of sound changes

It may be helpful to summarize some of the essential points of sound change by hypo-correction using the kind of informal shorthand representations phonologists have become used to. The development of distinctively nasal vowels from loss of a post-vocalic nasal consonant has usually been represented as in (5).

(5) $\mathrm{VN} > \tilde{\mathrm{V}}$

Alternatively, this has sometimes been broken down into two stages, as in (6).

(6.a) $\mathrm{V} > \tilde{\mathrm{V}}$ / ___ N

(6.b) $\mathrm{N} > ø$ / ___ $\tilde{\mathrm{V}}$

But stage (6.b) is frequently unattested and is assumed simply because the process in (6.a) failed to dispose of the N.

According to the mechanism I propose this should be represented as in (7).

(7)	Time 1		Time 2
(7.a)	/VN/ > [ṼN]	(7.c)	/VN/ > [Ṽ(N)]
(7.b)	[ṼN] > /VN/	(7.d)	[Ṽ] > /Ṽ/

Here, as usual, forms between slashes represent the intended pronunciation; those between square brackets, the phonetic form. At Time 1 before any change has occurred, according to (7.a), an intended sequence of a /VN/ is produced with the vowel phonetically nasalized. This is not a rule of grammar; it is a constant, timeless process that owes its existence to the physical constraints of the vocal tract. At this time, according to (7.b), listeners are able to take the [ṼN] sequence and reconstruct the intended /VN/ sequence. (7.b) is a rule of grammar; it is what I have been calling 'correction'. At Time 2, which represents the change itself, the process in (7.a) also occurs as (7.c) with the modification that the final N is weakly implemented such that it is difficult to detect or to associate with the preceding [Ṽ] (here the conventional notations do not serve us well). Like (7.a), (7.c) is also not a rule of grammar. (7.d) represents the action of the listener who detects only [Ṽ] and thus can only reconstruct /Ṽ/ as the intended pronunciation.

3.5 Advantages of viewing certain sound changes as hypo-correction

An advantage of (7) is that it indicates that the end product, the Ṽ, was present phonetically before the change. In fact, for all sound changes of the type called hypo-correction it is the case that the 'after' state is present in an incipient form in the 'before' state and can be studied in any present-day human language having the appropriate structure.[5] The only difference is that in the before state listeners effectively discount the predictable variations of sounds. Nevertheless, (7) is still no proper substitute for the full story which must derive (7.a,c) from first principles and demonstrate (7.b,d) via perception tests. The above-cited work by Kawasaki and by Beddor et al. (1986) constitute the required perceptual demonstration.

According to generative phonology sound change occurs because of a change in the grammar. In contrast, in the account I give, the only type of grammatical rule that changes anything, e.g. (7.b), has the function of preserving a pronunciation norm, not changing it. The change itself is the result of an unintended failure of the perceptual process.

One of the most important aspects of this account of sound changes due to hypo-correction is that it explains in a very simple way why these 'natural' sound changes parallel the kind of variation found in ordinary speech production. The parallels can be found in the segments affected, the conditioning environments, and the direction of the variation. Even so it does not equate sound change with synchronic variation; a mini sound change emerges from them only when a listener fails to normalize or correct the variation.

But there is an important class of sound changes that is not 'natural': changes in this class may involve similar segments and conditioning environments to those discussed above but the direction of change is the reverse. These are dissimilative sound changes, and are discussed in the next section.

4 Dissimilation

4.1 Dissimilation is 'correction' erroneously implemented

Dissimilation is defined as the loss or change of one or more features, including whole segments, when the same feature is distinctive at another site within a word. It can be manifested in several ways: as dissimilation on adjacent segments (called 'contact dissimilation' here), dissimilation involving non-adjacent segments (called 'distant dissimilation'), the prevention of an otherwise regular sound change that would have resulted in a sequence of sounds subject to dissimilation ('preventative dissimilation'), and as a constraint against the co-occurrence of similar segments. Examples of dissimilation are given in (8).

(8.a)	*Proto-Indo-European*	>	*Sanskrit*	
	*b̤end̤		band̤	'blind'
	Proto-Quichean	>	*Tzutujil*	
	*k'aq		k'ʲaq	'flea'
	Ancient Chinese	>	*Cantonese*	
	*pjam		pin	'diminish'
	Proto-Quechumaran	>	*Quechua*	
	*t'ant'a		t'anta	'bread'
(8.b)	*Slavic*			
	stoj + ā	>	stojɑ̄	'stand'
(8.c)	*Sanskrit*			
	sarsrāṇa (for expected *sarṣrāṇa)			

(8.d) *Yao*

*Cʷ V C
[lab]

English

* #C w – ; * #C l –
[lab] [stop]
[apical]

Arabic

* C V C
[phar] [phar]

These are 'unnatural' sound changes in the sense that, first, we are unable to invoke any principles of speech production that would predict changes in this direction. In fact, in the case of the backing of front [a] to back [ɑ] in (8.b) above, we would predict the opposite. Even if we move further 'upstream' and attempt to invoke putative cognitive principles of speech production, we are unable to make a good case (see below and Ohala 1991). Meringer and Mayer (1895) conducted their classic study of speech errors in order to determine whether they were the cause of dissimilative sound changes (Cutler 1982). They concluded that the types of variation were too different to be related to each other.

Secondly, this is not what we find in ordinary speech production. Broad and Fertig (1970) found the interactions between initial and final consonants in a CIC syllable to be small and what interactions there were, were positive, not negative or inhibitory (however, see Weismer 1979).

How can we account for sound changes which go in a direction that is the reverse of the 'natural' direction, that is, from what would be expected given speech production constraints? In fact, we have already discussed such a mechanism earlier: the listener's ability to undo or reverse the predictable perturbations found in speech. I propose that dissimilation arises due to the listener's mis-application of these corrective processes. I call this 'hyper-correction'.

For example, considering the case of Latin /kwiŋkwē/ > */kiŋkwē/ (and subsequently to Italian /tʃiŋkwe/, etc.) where initially lip rounding existed distinctively at two sites in the word. In addition, the lip rounding no doubt was evident on the intervening vowel given that it was flanked by lip-rounded segments; in fact, non-distinctive lip rounding was probably often present even on vowels just followed by /kw/ (according the Latin grammarians; see Devine and Stephens 1977: 37–42). A listener then, could have been confused as to whether the lip rounding detected at the beginning of this word was distinctive or a non-distinctive perturbation caused by the lip rounding on the second syllable. Some listeners apparently guessed wrong.

Perceptual confusions such as those posited as the basis of dissimilation are not unique to speech. This is, in essence, the same mechanism underlying camouflage, that is, hiding something by making it similar to its surroundings. The white arctic hare surrounded by white snow is 'hidden' from predators by being visually indistinguishable from the snow. The distinctive labialization on the initial stop of /kwiŋkwē/ is 'hidden' by being adjacent to and thus equated with labialization attributable to the second stop. In both cases the camouflaged thing can be *physically* detected by viewers/hearers but the percept – the *cognitive representation* of the scene/word formed by them – lacks the hidden element.

4.2 Supporting evidence for analysis of dissimilation as hyper-correction

4.2.1 Which features do and do not dissimilate

(a) Dissimilation restricted to 'stretched out' features

A logical consequence of this account of dissimilation is that it should only involve features which manifest themselves over fairly long temporal intervals, that is, which can encroach on adjacent segments and thus create an ambiguity as to where the feature is distinctive and where fortuitous. Examples of such 'stretched out' features are labialization, aspiration, retroflexion, pharyngealization, the voice quality called 'glottalization',[6] and place of articulation. It would not involve features such as 'stop', 'affricate' which do not stretch over long temporal intervals. In general, this prediction is borne out.

This prediction regarding the features that should and the features that should not be subject to dissimilation needs further refinement as to what is meant by the term 'feature'. This word is sometimes used to mean 'classificatory label', in which case the term may be completely devoid of phonetic substance, e.g. FORTIS, LENIS, or partially lacking in phonetic substance, as is the case sometimes with the feature VOICE. 'Feature' is also used variously to refer to articulatory events, e.g. NASAL, LATERAL, or to acoustic properties, e.g. STRIDENT, GRAVE. What matters in the present case are the *acoustic-auditory cues* which serve to differentiate linguistic signals. There are typically multiple cues differentiating what are called 'features'. For example, the cues to the feature NASAL (in the case of consonants) include the abrupt discontinuity in amplitude and spectrum with respect to adjacent vowels or other sonorants and the increased bandwidth of the formants of adjacent vowels (i.e. vowel nasalization). The first of these cues manifests itself rapidly, within some 20 to 30 msec.; the second cue manifests itself over at least 50 msec, sometimes longer. The major cues to features like LABIALIZED (lowered resonances and overall amplitude) occur over a time-span that may be 100 msec or more. Cues then may be characterized according to the time window which they require to be heard. This, in turn, determines the extent to which they overlap with cues for adjacent segments. In general, manner of articulation, e.g. stop vs. fricative vs. nasal (consonant), tend to be conveyed primarily by fast cues whereas so-called secondary articulations and modifications, e.g. labialization, retroflexion, velarization, pharyngealization, glottalization, aspiration, tend to be cued by slow cues. Some of the place of articulation cues are intermediate on this continuum: the formant transitions cuing place may range from 30 to 80 msec or more (Lehiste and Peterson 1961; Kewley-Port 1982). The modified prediction regarding the phonetic elements subject to dissimilation is, then: features requiring a long time window for their perception are more likely to be subject to dissimilation than are those requiring short time windows. In the case of segments like nasal consonants which have some rapid cues and some slow cues, it will depend on which of these cues dominate the percept. (See Massaro and Oden 1980 and Abramson and Lisker 1985 for studies which attempt to explore the interaction of various cues.) In addition, the relative salience of various cues may differ from one language to the next. How various cues function in speech is ultimately an empirical question but since few studies on this topic have been carried out, the best I can do at present is to argue for the plausibility of my claims.

It seems clear that at the two extremes of the time-window continuum, stops and affricates seem not to be subject to dissimilation whereas secondary articulations like labialization and retroflexion are. But there are interesting apparent counter-examples.

(b) Dissimilation of laterals

According to phonetic arguments given above, laterals would seem to be segments that should not dissimilate since laterality cannot easily spread to adjacent segments; nevertheless, [l]s are subject to dissimilation in Latin as in (9).

(9) Latin dissimilation of /l/ in stems containing /l/

liber-alis BUT famili-aris

mort-alis BUT popul-aris

But laterals do have very long transitions of Formants 2 and 3 (Lehiste 1964; ch. 2; Javkin 1979: 38ff.) and these are probably prominent cues. In fact, there are well-known sound changes where /l/'s become just glides, palatal or labial velar, depending on whether the laterals themselves were originally palatalized or velarized (von Essen 1964); see (10).

(10) Examples of laterals changing to glides.

Late Latin pleno- > Italian pjeno

Late Latin alterum > French autre

Furthermore, O'Connor et al. (1957) (and several studies since) demonstrated that /l/'s and the North American English /r/ can be synthesized as endpoints of a continuum which involves only variations of Formant 3, the lateral having high F3 and /r/, low. Thus it seems that laterals do have cues that require a long time-window for their perception and so their involvement in dissimilation is not puzzling. In addition, the study of O'Connor et al. (1957) suggests why /l/ generally dissimilates to /r/ rather than some other segment.

(c) Dissimilation of voicing?

The feature [voice] would seem to be a prime candidate for dissimilation since it has some reputation as a feature that spreads to adjacent segments (especially intervocalically) and there are a number of cases reported as voicing dissimilation: the dissimilation of the voicing of intervocalic fricatives in Gothic (Thurneysen's Law), the dissimilation of voicing in stops in Sakuma (Bantu), and the well-known co-occurrence constraint against successive voiced stops in Proto-Indo-European roots. However all these cases have been reanalysed – in some cases not without controversy – in a way which casts them as other than voicing dissimilation (Flickinger 1981; Bennett 1967; Hopper 1973; Gamkrelidze and Ivanov 1973). Furthermore, based on my understanding of the principal cues for voicing, I would predict that [voice] should not be subject to dissimilation. Although one might conclude that there are 'prosodic' or long time-window cues to voicing from such often-encountered claims of the sort 'preceding vowel duration is a cue to consonant voicing', in fact there is no evidence supporting such claims. What one has instead is evidence that vowel duration is sufficient to cue contrasts that may *also* be cued (or were *once* cued) by consonantal voicing. There is no evidence that a listener judges a voiceless consonant as physically voiced after a long vowel.[7] Rather, the primary cue to a segment being voiced is the generally robust cue of periodic pulsation in the lower frequencies. This cue operates in a relatively short time-window and does not manifest itself by colouration of adjacent segments; therefore it should not be susceptible to dissimilation. This matter is hardly settled but I would predict that the eventual resolution of the putative cases of voicing dissimilation cited above would show that [VOICE] was not involved.

(d) Robust vs. non-robust cues in creating languages' segment inventories

The characterization of cues by the amount of time needed to perceive them was inspired by Stevens' (1980) differentiation of features into robust, short ones and less robust, long ones. (The only modification I have made is to focus not on phonological features per se but the acoustic-auditory cues which signal them.) He pointed out that when constructing their segment inventory, languages generally use up the robust features first before exploiting the less robust ones. Thus only languages with very large consonant inventories make use of labialization, retroflexion, pharyngealization, glottalization, etc., and only languages with a relatively large vowel inventory make use of nasalization or front-rounding and back-unrounding. It seems evident that those sound types which appear only in large segment inventories are approximately the same ones which are most subject to dissimilation. Furthermore, a dissimilated segment, that is, the result of dissimilation, is generally one from the 'robust' class rather than the original less robust set. The link between the phenomena of segment inventory constraints and those of dissimilation is the long time interval needed for the perception of the acoustic modulations which cue certain features: this makes them auditorily less salient and makes them 'run into each other' when they appear on nearby segments thus leading to camouflage.

Having made the prediction in 4.2.1.1 above, I have tried to marshal some of the available evidence in support of it and to account for apparent counterexamples. Of course, other counter-examples exist (Ohala 1981). It is apparent that these (as well as the cases which are apparently consistent with the prediction) have to be examined carefully. Nothing should be accepted at face value. The outcome of conflict between the prediction and

the counter-examples will be one of the following:

(a) the prediction will be abandoned or modified:
  (i) the prediction has so many counter-examples that it has to be abandoned;
  (ii) the general prediction is correct but the specifics as to which cues make a segment subject to dissimilation requires modification;
  (iii) dissimilation arises due to different mechanisms; the prediction only covers one of these.

(b) the counterexamples can be reanalysed in a way that eliminates their conflict with the prediction.
  (i) they may not be cases of dissimilation;
  (ii) they may not involve the feature originally proposed.

4.2.2 *Hyper- and hypo-correction have different constraints on conditioning environment*

As was discussed above, a common although not invariable characteristic of sound changes due to hypo-correction is that the conditioning environment is lost at the same time as the conditioned change occurs. In fact, failure to detect the conditioning environment is a direct cause of the listener failing to implement correction of a contextually caused perturbation and therefore taking it at face value. In contrast, from the account just given it would be predicted that in sound changes due to hyper-correction the conditioning environment may *not* be lost at the same time as the conditioned change. Hyper-correction involves the listener blaming (so to speak) the conditioning environment for an imagined perturbation; obviously, then, the conditioning environment must be detected. (11) gives examples of a hypothetical version of Grassmann's Law in Sanskrit that should not occur.

(11) bhandh > ban (where the 'dh' drops just when the 'bh' is dissimilated)

I am unaware of any exceptions to this prediction.

4.2.3 *Hyper-corrective sound changes do not result in 'new' segments*

Although sound changes due to hypo-correction may often result in the introduction of new segments, for example nasalized vowels from loss of a post-vocalic nasal consonant, sound changes caused by hyper-correction apparently do not. The end product of dissimilation seems to be a segment drawn from the same set that the language had before the sound change. This behaviour follows from dissimilation being the result of the listener applying normalization processes to the speech signal: normalization involves recovering a (presumed) standard sound from a signal that differs in some way from the standard.

4.2.4 *Dissimilation-at-a-distance blocked in certain environments*

In Tzutujil the dissimilation of back velars fails to occur if the intervening vowel is [o]; see (12, from Campbell 1977).

(12)

Proto-Quichean	Tzutujil	Translation
ke:x	kʲe:x	'horse'
k'aq	k'ʲaq	'flea'
BUT:		
koxl	kox	'cougar'
k'ox	k'o:x	'mask'

This otherwise puzzling behaviour makes sense if we take into account the phonetic manifestation of the back velar perturbation on vowels and the inherent properties of back rounded vowels like [o]. Back velars cause a lowering of the F2 of adjacent vowels (Klatt and Stevens 1969; Ghazeli 1977). Back rounded vowels, especially [o] and [ɔ] have the lowest possible F2 of all vowels. Apparently, then, dissimilation-at-a-distance is blocked if the intervening segments are *saturated* in the acoustic property (in this case low F2) that would otherwise manifest the consonantal perturbation. Not being able to detect any consonantal colouring of the vowel, the listener has no basis to imagine that such a perturbation migrated to distant sites in the word. (This account assumes, of course, that there are other cues to the feature BACK-VELAR than the lowered F2 transitions, otherwise the [o] itself might cause dissimilation of these segments. Presumably the spectrum of the stop burst would be sufficient to cue the BACK-VELAR feature even in the absence of distinctive modulations of the F2 transitions.)

I am unable to say how generally this 'saturation' phenomenon blocks dissimilation-at-a-distance but if found more widely it offers support to the account of dissimilation offered here.

4.2.5 *Experimental evidence of the perceptual basis for hypo-correction and hyper-correction*

Perhaps the most persuasive support that can be given for the above account of sound change in terms of hypo- and hyper-correction is experimental results. I cited above experimental evidence for listeners' ability to 'correct' the speech signal. Many of the same experiments implicitly duplicate the phenomenon of hypo-correction, for example Kawasaski's result that listeners judge a vowel from a syllable like [mɪm] to be more nasal when the flanking nasal consonants are removed (than when they are present) and Beddor et al.'s (1986) result that listeners hear more [æ] vowels on a [ɛ]–[æ] continuum when they are nasalized than when they are either oral or when the nasal vowels are followed immediately by a nasal consonant. But this latter study also found evidence for hyper-correction: the nasal vowels that were successfully normalized or corrected back to their equivalent oral quality when followed by a nasal consonant were over corrected, i.e., hyper-corrected, when the degree of nasalization was slight.

Ohala and Shriberg (1990) demonstrated hyper-correction in a vowel confusion study. As revealed in the traditional vowel chart where height correlates inversely with F1 and frontness with F2, front and back vowels of equivalent height have similar F1 values but are differentiated by F2. When short context-independent vowels are low-pass filtered in a way to eliminate F2, listeners naturally make many errors where front vowels are confused with back vowels, whereas the reverse confusion is rare. However, when these filtered vowels are embedded in a sentence context which has also been subjected to the same filtering, two things result: first, the overall level of errors decreases significantly (meaning that 'correction' takes place) and, secondly, there is a disproportionate increase in the number of back vowels confused with front (indicating hyper-correction). Taking their cue from the filtered sentence which was obviously missing acoustic energy at the higher frequencies, the listeners in effect 'added' high frequency energy to a few tokens of the back vowels and thus made front vowels of them. Although the particular pattern of perceptual errors studied in this case does not show close parallels to any known sound change (since low-pass filtering is not a common distortion of speech under normal conditions), it does show that listeners may hyper-correct the speech signal.

5 Overview

5.1 What can happen between a speaker and hearer: a basic taxonomy

To sum up the preceding two sections (3 and 4), there are four outcomes from an exchange between speaker and listener where the signal may be potentially ambiguous regarding the speaker's intended pronunciation.

A. Correction
B. Confusion of Acoustically Similar Sounds
C. Hypo-correction
D. Hyper-correction

In A there is no sound change; the listener successfully recovers the speaker's intended pronunciation from the less-than-perfect signal. B, C, and D all result in sound change – B and C due to failure to correct an ambiguous signal and D due to 'correcting' a signal that didn't require it. B could be included under C; the only basis for differentiating B from C is whether the disambiguating cues that could have been used by the listener (but were not) are temporally co-terminous with the ambiguous part or whether they are not, respectively. For example, at least one of the cues differentiating the stops in the syllables /gi/ and /di/ is a sharp peak in burst spectrum around 3 KHz; in other respects the spectra are quite similar (Ohala 1985b). On the other hand, in the case of the confusion of high vowels with lower vowels when nasalized, one disambiguating cue is the presence of an adjacent nasal consonant. Many, perhaps most, sound contrasts have both temporally simultaneous and temporally sequential cues which differentiate them from other sounds. Collapsing B and C yields the three categories

A'. Correction
B'. Hypo-correction
C'. Hyper-correction

5.2 Constraints on the explanation of the mechanism of sound changes

I have attempted to present a general paradigm for explaining how certain sound changes start. But a casual reading of this presentation may give the impression that it is unconstrained, that hypo-correction and hyper-correction can be played like 'wild cards' whenever needed: finding a sound change where A > B one labels it hypo-correction whereas given another where B > A, one simply labels it hyper-correction. If this were true the paradigm would offer no advance over the kind of explanation for sound changes where paradigmatic pressures, syntagmatic pressures, symmetry, asymmetry, simplicity, etc. – and these from any level: phonological, morphological, lexical, syntactic, and semantic, socio-pragmatic – can all be invoked with little or no independent justification.

It is inevitably true that the passage of time does result in loss of information – a rise in entropy – that makes reconstruction of past events ambiguous. Detective stories often feature events with an ambiguous history: the single shoe print in the sandy beach – did it happen due to someone stepping there or was the shoe in place for a long time while the surf caused the sand to rise around it and then the shoe removed? But after acknowledging the existence of ambiguities in reconstructing the past, there are still cases where some of the ambiguities can be resolved and it is these constraints on invoking hypo- and hyper-correction as a source of sound change that I review now.

5.2.1 *Constraints on using hypo- and hyper-correction in historical reconstruction*

First, one of the overriding constraints is that sound changes due to hypo-correction are those which are consistent with known properties of the speech production mechanism. That is, in an ideal case one would have empirically-based theories or models of speech production, for example anatomical (Harshman et al. 1977), aerodynamic (Rothenberg 1968; Stevens 1971; Ohala 1976; Muller and Brown 1980), the mapping from vocal tract shape to sound (Fant 1960; Carré and Mrayati 1990), and one could refer to these to see how the speech output is constrained. For example, an aerodynamic model would indicate that the maintenance of voicing on back-articulated stops like velars and uvulars would require special gestures; maintaining voicing on *any* stop longer than some 65 msec would require special gestures. Lacking these special gestures, devoicing is likely to occur (Ohala 1976, 1983b; Ohala and Riordan 1979; Westbury 1979). Even in the absence of a comprehensive model of speech production, we may obtain generally reliable information on natural tendencies of speech production by inductive means, that is, by surveying phonetic behaviour in several diverse languages. For example, cross-language studies suggest that vowels are generally longer before fricatives than corresponding stops (Peterson and Lehiste 1960; Delattre 1962) but to my knowledge this pattern has not yet been explained by reference to known properties of the speech production mechanism. Sound changes attributed to hypo-correction, then, would involve listeners copying at face value those details of speech that originally owe their existence to the influence of physical phonetic properties of the speech production system. It must be allowed, however, that after a sound change has occurred – that is, the listener misinterprets the function of these phonetic details – the shape of these phonetic events and features may be different, perhaps exaggerated *vis-à-vis* their original state. Thus, for example, distinctive vowel nasalization on vowels that used to stand next to nasal consonants may be considerably greater in amplitude and temporal extent than it was when the nasalization was a predictable phonetic feature (Solé and Ohala 1991).

Sound changes due to hyper-correction do not conform to known constraints of speech production. I will consider in detail one proposed counter-example to this principle.

5.2.2 *Dissimilation of aspiration: a counter-example?*

There has been speculation that Grassmann's Law in Sanskrit and Greek occurred due to the greater 'effort' required to make two aspirates in a row:

> An aspirate requires great effort . . . beginning from the abdominal muscles and ending in the muscles that open the glottis to its widest extent. It was in order to economize this muscular energy that the tenuis was substituted for the aspirate (Müller 1864: 179–80).

> Aspirated consonants are . . . costly in that they use considerable respiratory energy. A word with two such sounds is very costly, and an obvious candidate for pruning in any attempt to reduce the overall effort required for an utterance (Ladefoged 1984).

In fact, however, although 'ease of articulation' is often appealed to as a cause of sound change, no one has found a way to measure total energy expenditure in speech in order to establish convincingly its relevance. However, the amount of air under pressure in the lungs can easily be measured but it is not clear that aspirated consonants in speech push the respiratory system to its limits (Ohala 1990b). In general, speakers have a considerable respiratory reserve at their command and seldom exhaust it. Moreover, there are inherent aspects of the operation of Grassmann's Law (and similar processes in other languages; see de Reuse 1981; Turner 1923–5; Grammont 1933: 316; Jhā 1958: 142; Allen 1957) which are inconsistent with the hypothesis which invokes 'economy of articulatory energy': Energy cost is presumably cumulative from beginning to end of utterances and one would suppose that the urgency to reduce energy expenditure would be greater later in the utterance rather than earlier. Yet it is generally the first aspirate of two which is de-aspirated, not the second. Furthermore, if the energy expenditure mattered one would suppose that dissimilation of aspirates would occur any time two or more of them followed each other in speech not just when the two were found within the boundaries of a single word, as seems generally to be the case. (See pp. 263–4 for an account of why sound change generally takes place within word boundaries.) Finally, if articulatory energy figured in dissimilation it should not matter whether the two sounds involved are similar or not; any two energy-costly sounds, even those very different from each other would be likely to dissimilate. But again, this seems not to be what is found. There is therefore no evidence that Grassmann's Law is initiated by the speaker trying to conserve articulatory energy.

5.2.3 *Further differences between hypo- and hyper-correction*

As argued above, hypo- and hyper-correction can also be differentiated by the criteria in (13).

(13)

Criterion	Hypo-correction	Hyper-correction
Loss of conditioning environment?	Possible	No
Results in new segments?	Possible	No
Change from robust segment to less robust (by Stevens' criteria, Stevens et al. 1986)?	Possible	Usually no

5.2.4 *The ultimate check: duplicating sound change in the lab*

Perhaps the ultimate check on any hypothesis about the cause of a particular sound change is to test the hypothesis in the laboratory. If particular sound changes are posited to have a phonetic basis then one should be able to duplicate the conditions under which they occurred historically and find experimental subjects producing 'mini' sound changes that parallel them. It is because of the posited phonetic character of sound change that a laboratory study is possible: were the initiation caused by grammatical and cultural factors, this would be more difficult or perhaps impossible. I have referred above to several studies which, in effect, simulate sound change in the laboratory: some, like Kawasaki (1986) and Wright (1986) were designed with diachronic questions in mind, others, like Winitz et al. (1972) and Mann and Repp (1980), are relevant to diachronic issues even though that was not their original purpose.

6 Discussion

I have presented a general plan to explain the initiation of sound changes found in similar form in diverse languages. Certain specific sound changes were discussed in detail and for some of these experimental evidence was cited in support of the hypothesized explanation. Obviously, much more research is needed to flesh out this plan; however, I hope the outline given here in addition to the experimental papers cited lend it plausibility. This plan has the following potentially important characteristics which differentiate it from other accounts of sound change.

6.1 Sound change is non-teleological

There is a long tradition of teleological accounts of sound change where speakers are claimed to choose (no doubt unconsciously) a novel pronunication in order to optimize some aspect of communication: to make speech easier to pronounce, easier for the listener to hear, or easier to process by making the grammar simpler (Müller 1864: 176ff.; Whitney 1867: 69ff.; King 1969: ch. 4). But the mechanism of sound change I propose above is non-teleological: there is no intention by either the speaker or the listener to change pronunciation. Indeed, the whole purpose of the listener's interpretive activity is to attempt to deduce the pronunciation intended by the speaker, i.e. to *preserve, not to change*, the pronunciation norm. It is when the listener makes mistakes in this interpretation that sound change can start.

As a corollary of this it must be emphasized that pronunciation change itself is not included in the cognitive processes of either speaker or hearer. That is, although distinctive vowels arose in French from sequences of VN, the rule in (14) was not part of either the speaker's or listener's grammar.

(14) VN > Ṽ

Using the terms of the communication engineer, change occurs not in the message source (the speaker's brain) nor the message destination (the listener's brain) but in the transmission channel between them. This includes the speech production system and the listener's decoding system.

Sound change is thus like change in other domains: many of the errors made by scribes copying manuscripts, errors students make when taking notes on professors' lectures, errors in the transcription of codons in DNA – all are unintended and the change itself is not part of any rule set guiding or characterizing the behaviour of the thing doing the copying.

In disposing of teleology from the domain of sound change we free the study of diachronic phonology from many logical and strategic impediments. For example, Lightner (1970) proposed, in essence, that vowels become nasalized after loss of a post-vocalic nasal in order to preserve lexical distinctions. This is not an atypical claim. But there are profound logical inconsistencies in such a claim. First, lexical distinctions *are* commonly lost even though vowel nasalization occurs, for example French [blɑ̃] for both *blond* and *blanc*. Languages teem with homonymy and polysemy without suffering any apparent distress. Preservation of contrast, then, might as effectively be attributed to chance than to the speaker. Secondly, if speakers have such control over their pronunciation as to worry about maintaining the phonetic distinctions between words, then why did they allow the final nasal in such words to disappear in the first place? Why are they helpless in the face of one phonological change but masters of the situation in another?

Furthermore, reliance on teleological accounts of sound change is poor scientific strategy. For the same reason that the mature sciences such as physics and chemistry do not explain their phenomena (any more) by saying 'the gods willed it', linguists would be advised not to have the 'speaker's will' as the first explanation for language change. Not that the will of the gods and the will of speakers may not be the ultimate answer in both cases, but one should explore the less extravagant hypotheses before the more extravagant ones. This strategy has had a splendid payoff in every science that has embraced it. Explanation is, after all, reducing the unknown to the known, not to further unknown, uncertain, or unprovable entities.

None of this is meant to deny the role of teleology in other aspects of language change, especially its spread. I just deny the necessity of teleology in accounting for the pre-conditions or initiation of sound change.

6.2 Sound change is phonetic

This account of sound change also locates the mechanism centrally in the phonetic domain and primarily within the listener. The speaker is responsible for much variation in speech but normally most of this is discounted by the listener and so does not lead to sound change. It is only when the listener fails to normalize the variations in the speech signal or 'corrects' details that did not require correcting that sound change may take place. Thus the inescapable parallelism between diachronic and synchronic variation is accounted for.

Much of sound change can be viewed as a kind of parsing error by the listener (for the use of 'parsing' as it applies to speech perception, see Fowler 1986). In hypo-correction the listener fails to parse or associate a given perturbation or variation in the speech signal with the conditioning environment. In hyper-correction the listener erroneously parses a given event with another event.

Related to this is the fact that in the stream of speech the domain of change is overwhelmingly the word or possibly phrases which occur so often that they could also be said to be lexicalized. Why should this be so? Consider that if sound change were simply a matter of the speaker trying to make pronunciation easier there is no reason to limit change to the domain of the word. It could have been the case, as alluded to above, that in some language no more than one aspirated consonant could be produced per breath group such that the second and subsequent aspirated stops in an utterance would be de-aspirated. But generally this is not what is found. To be sure there are phonological changes that occur across word boundaries, for example French liaison, the palatalization of alveolars in English before a following palatal segment: *gas shortage* > [gæʃ·ɔɹɾədʒ], but even in the latter case there is evidence that such changes occur only with certain lexical items in certain collocations, not across-the-board (Solé and Ohala 1990). The dominance here of lexical units or lexicalized phrases is thus not contradicted. The principal role of the word in sound change follows from sound change being essentially a parsing error on the part of the listener. Accurate parsing of the different phonetic events in a word requires that the parts be separately identifiable. The optimal conditions for this occur when the parts are freely combinable and permutable, that is, when they appear independently of each other. A word is in essence a string of phonetic events frozen in a fixed order. It thus presents the maximum ambiguity to the listener of the separate parts (i.e. as intended by the speaker). This then is the domain where the listener is likely to make the most parsing errors.

Structuralist accounts of sound change (Jakobson 1978 [1931]) provide teleological scenarios that emphasize how a given change was motivated or shaped by its function within the whole system or structure of the language. I have given my views on teleology in sound change above. What about the influence of language structure on sound change? At best, I think this has been seriously exaggerated. While not denying the role of a given language's structure on sound change, especially in the spread of a given type of change to all structurally similar sounds, for example Grimm's Law where all the Indo-European voiceless stops became homorganic voiceless fricatives (but cf. Japanese where just the earlier /p/ changed to a fricative), I think a more readily apparent influence on sound change is the physical phonetic character of the sound involved. In Ohala (1979) I demonstrated how labial velar sounds, e.g. [k͡p, g͡b, w], show substantially similar phonological behaviour even though the languages involved have widely different phonological structure. Labial velars are by no means unique in this respect: similar demonstrations could be made for the behaviour of nasal consonants, nasal vowels, the interaction between voice and obstruents, segments designated [+flat] (Ohala 1975, 1983b, 1985a). A fundamental problem with structuralist accounts of sound change is that they are largely unconstrained: given a certain change, there are a great variety of structural 'pressures' that can be invoked to explain it, after the fact. One can appeal to (a) the language's segment inventory which itself has several degrees of freedom, place and manner contrasts, symmetry or asymmetry in certain contrasts, its size and relative 'density', (b) phonotactics, (c) the lexical or grammatical function of the contrasts, (d) the frequency of occurrence or functional load of the contrast, and so on. There are few rigorous attempts to show via broad cross-language surveys that in a significant number of cases languages showing a given structural trait undergo a sound change claimed to be a response to that trait. The accounts of sound change based on phonetics do not share this weakness.

6.3 Natural and 'unnatural' sound changes

The account given here provides a consistent, integrated account of sound changes generally regarded as opposite in character, 'natural' vs. 'unnatural', assimilative vs. dissimilative.[8] Central to this is the distinction between hypo-correction where the listener copies at face value the naturally occurring perturbations in the speech signal, thus producing natural, assimilative changes, vs. hyper-correction where the listener unnecessarily corrects the speech signal, thus giving rise to unnatural, dissimilative changes. (The terms 'natural' vs. 'unnatural' are perhaps unfortunate because, as I have tried to argue, dissimilation is natural in the sense that it can be understood by reference to universal perceptual strategies.)

6.4 Sound change in the laboratory

Studying sound change in the laboratory, as in work cited here, is not a novel undertaking. It has at least a century-old tradition (Rousselot 1891; Verner 1913; Stetson 1928; Grammont 1933; Haden 1938; Janson 1986). However, it has never been a 'mainstream' concern in historical phonology for a variety of reasons, not the least being the structuralist de-emphasis of the relevance of phonetic substance. The growing body of experimental literature addressing issues of phonetic and phonological universals and sound change should attract the attention it deserves based on its own merits, that is, what it can do for the historical linguist.

Language history, like history of any sort, and like astrophysics and much of geology deals with a subject matter that is inherently inaccessible to direct study. But conceptual breakthroughs for these latter two disciplines were achieved by adopting the so-called uniformitarian assumption, that is, that the composition of the stars and the earth and the forces and processes which shaped them throughout their remote history are in essence the same elements and forces that are present and detectable now. Studying sound change in the laboratory is also based on an assumption of uniformitarianism: variation in speech studied today parallels variation in centuries past. *In vitro* study of this variation substitutes for the physically impossible *in situ* diachronic study. However, laboratory study does not replace the detective work of the comparative method, it supplements and strengthens it by allowing a more rational choice of sound changes to be posited to convert reconstructed into attested forms.

6.5 Sound change is phonetically abrupt

Almost everyone speculating about sound change allows that it must be going on now as it always has and yet finds it difficult to detect. There is no report in history 'Today, everyone started to pronounce "meat" with an [i] vowel.' From this it follows that sound change must be gradual and taking place at a rate too slow to detect. This led to the supposition that the phonetic shift from one sound to another was progressing in steps too small for the ear to detect. An inherent aspect of the account presented here is that sound change is phonetically abrupt: the shift from one pronunciation to another is large enough to be detected. In most cases the 'before' and 'after' states could be contrasting sounds or sound sequences in some human language (though not necessarily in the language in which the change occurred). Gradualness, then, must lie in other domains: its spread from one speaker to another, from one speaking style to another, from one word to another, etc. There is substantial evidence that such gradual spread does occur (Wang 1977).

6.6 There will never be a 'complete' explanation of sound change

Lass (1980) finds that no account of sound change, including those based on phonetics, permits one to predict (or post-dict) its course by appeal to laws derived from first principles. Historical linguistics thus fails to be a *deductive nomological* discipline, which, according to him, is the goal of all mature scientific disciplines such as physics and chemistry. (See similar remarks by Dinnsen 1980.) The account of sound change given here also is not capable of predicting why a particular sound change takes place in a given language at a particular time and so, by Lass's criteria, this is not a scientific work. But, as outlined in detail in Ohala (1987), there is a fundamental error in Lass's reasoning.

Deductive nomological disciplines do not exist (with the possible exception of mathematics which deals with an artificial universe, not the real one). The problem is with the term 'nomological' or law-based. It implies perfect knowledge of the universe – something which is unattainable. However absolutely scientists may state their beliefs, the actual data on which they are based invariably shows some quantifiable discrepancies which are regarded as negligible. But the history of science suggests that what one age neglects the next uses to overthrow the 'laws' of the previous one. What we are left with in science is *deductive probabilistic* explanations of phenomena and this applies as well to physics and the account of sound change given here. Physics, the more mature science and with more control over the factors influencing events in its domain, makes more accurate probabilistic accounts than anything linguistics is capable of, but there is no qualitative difference between them.

The study of the phonetic bases of sound change is at the very threshold of being able to make deductive probabilistic predictions. Although there will always be more to learn, there is a good understanding in phonetics of the mechanisms for turning gestures and postures of the vocal organs into sound (Fant 1960; Lindblom and Sundberg 1971; Maeda 1990; Carré and Mrayati 1990). There is also some understanding of how physical constraints of this mechanism can give rise to variation (Ohala 1976, 1983b; Goldstein 1983; Westbury and Keating 1985). We have the beginnings of models which can predict auditory response, including rate of confusions, to stimuli based on their acoustic properties (Klein et al. 1970; Wright 1986; Bladon and Lindblom 1981; Lindblom 1986; Stevens 1989) – indeed, this latter measure is inherent in all template-based automatic speech recognition systems (Waibel and Lee 1990).

We are thus almost in a position to give probability estimates of which sounds will be confused based on deductions from first principles. We can already do this inductively (e.g. based on confusion matrices such as those published by Winitz et al. 1972). Thus we can predict that confusions of the sort [gi] > [di] will be more common than [gu] > [bu]. In the laboratory these predictions apply to the population of listeners hearing speech under specified conditions (usually with no higher-order linguistic redundancies present to influence their judgements). The predictions do not apply to individual listeners. No one can say why listener A identified the stimulus as X, not Y, but listener B identified Y correctly. Likewise, when extrapolating these predictions to languages, it will not be possible to say anything about individual languages or specific time periods in their history, simply that considering a large and representative sample of human languages, such-and-such confusion or change is more likely than another. But equally, it may be no more interesting or fruitful to try to identify the factors which initiated a change in a specific language at a specific time in history than it is to enquire further into why listener A responded differently from listener B in laboratory-based speech perception experiments. As in public health and epidemiological science, it should be sufficient to be able to make useful predictions about influence on the large mass of the population without being able to say anything certain about what will happen to individuals.

What should one make of the many attempts in the historical phonology literature to explain sound changes in a specific language at specific times by referring to contemporary cultural and psychological forces? At their best, they are accounts of why these sound changes *spread* – because at any given time all languages are probably flooded by all applicable mini-sound changes. Maxi-sound changes arise due to some of these mini-sound changes spreading selectively whereas others fail to spread; cultural and psychological factors undoubtedly play a role in this. At their worst, they are Kiplingesque 'Just So' stories: *ad hoc* hypotheses never subjected to empirical test and stated in terms of explanatory principles that themselves require explaining. Historical phonology would do well to wean itself away from this latter genre of explanation; they are neither helpful nor necessary to a larger understanding of sound change.

6.7 Relevance of historical phonology to practical domains

The relevance of phonetic research in practical domains such as communication disorders and speech technology has never been questioned (for the most part). I have attempted here to show

the relevance of phonetic research to an understanding of the mechanism of sound change initiation. If this is accepted then it implies that basic research on speech production and perception as done in phonetics, the study of sound change, and applied research on speech all have a common scientific core. The flow of useful information and hints could be from any of these domains to the others, including from historical phonology to speech technology. Whether in speech synthesis or automatic speech recognition a knowledge of all the multiple cues used to differentiate words is useful; hints on these multiple cues for a given sound can be obtained by seeing what sorts of changes it induces in neighbouring sounds. For example, Lea (1973) proposed that the perturbations of FO on vowels following consonants with distinctive voicing contrasts – which diachronically led to the development of tones – could be used in a speech recognition task. Many other examples could be provided (Ohala 1975, 1985b, 1986). It would benefit all researchers studying the behaviour of speech, including the historical linguist and the communication engineer, to integrate their data, methods, and theories; let us not repeat the folly of the blind men each describing a separate part of an elephant.

Notes

1. But also in other languages, e.g. Punjabi (Gill and Gleason 1969: 33), and Nama (Beach 1938: 247–53).
2. Hombert et al. (1979) cautiously endorsed the hypothesis that the consonantally induced perturbation of FO on vowels was caused by a difference in larynx height which in turn created a difference in vocal cord tension. They reviewed but offered counterarguments to the hypothesis of Halle and Stevens (1971) that differences in vocal cord stiffness were responsible. However, given the study by Löfqvist et al. (1989) and the work by Riordan (1980) and Kingston (1985) casting doubt on the larynx height hypothesis, it seems the Halle and Stevens hypothesis has the most empirical support.
3. The historical processes leading up to alternations such as those in (4) are somewhat complex. From this data it is safe only to conclude that when nasalization affects vowel quality it can cause reduction in height distinctions and tends to lower non-low vowels.
4. See Ohala 1974a,b, 1975, 1978a, 1981, 1983a,b, 1985a, 1986, 1990a; Ohala and Lorentz 1977; Ohala and Ohala in press.
5. The qualification 'appropriate structure' is added here to eliminate trivial exceptions to this generalization, e.g. one could not very well find out how vowels are treated before nasal consonants in a language that either did not have nasals or did not have syllables closed by nasals. Similarly, one could only study the influence of consonant

voicing on following vowels if a language had consonants contrasting in voicing.
6. The voice quality called 'glottalization' is often found on voiced sonorants flanking 'glottalized' consonants, e.g. ejectives and implosives.
7. Taking vowel duration as a cue to 'voicing' is a prime example of confusing the classificatory (not necessarily phonetic) feature [+VOICE] with the physical entity 'voice'.
8. Also certain cases of syncope vs. epenthesis; see Ohala 1991.

References

ABRAMSON, ARTHUR and LISKER, LEIGH (1985) 'Relative power of cues: FO shift versus voice timing.' In FROMKIN, VICTORIA A. (ed.) *Linguistic Phonetics*. New York: Academic Press.

ALLEN, W. SYDNEY (1957) 'Aspiration in the Hāṛautī nominal.' *Studies in Linguistics Analysis* (Special Volume of the Philological Society of London). Oxford: Blackwell, pp. 68–86.

AMMAN, JOHANN CONRAD (1700) *Dissertation de Loquela*. Amsterdam: Wolters.

BEACH, DOUGLAS M. (1938) *The Phonetics of the Hottentot Language*. Cambridge: W. Heffer.

BEDDOR, PATRICE SPEETER (1983) *Phonological and phonetic effects of nasalization on vowel height*. Bloomington, IN: Indiana University Linguistics Club.

BEDDOR, PATRICE S., KRAKOW, RENA A. and GOLDSTEIN, LOUIS M. (1986) 'Perceptual constraints and phonological change: A study of nasal vowel height', *Phonology Yearbook* **3**: 197–217.

BENNETT, PATRICK R. (1967) 'Dahl's Law and Thagicu.' *African Language Studies* **8**: 127–59.

BLADON, R., ANTHONY, W. and LINDBLOM, BJÖRN (1981) 'Modeling the judgment of vowel quality differences.' *Journal of the Acoustical Society of America* **69**: 1414–422.

BLOCH, JULES (1920) *La Formation de la Langue Marathe*. Paris, E. Champion.

BLOCH, JULES (1965) *Indo-Aryan: From the Vedas to Modern Times*. Paris: Librairie d'Amerique et d'Orient Adrien-Maisonneuve.

BLOOMFIELD, LEONARD (1914) *An Introduction to the Study of Language*. New York: Henry Holt.

BROAD, DAVID J. and FERTIG, RALPH H. (1970) 'Formant-frequency trajectories in selected CVC-syllable nuclei', *Journal of the Acoustical Society of America*, **47**: 1572–82.

DE BROSSES, CHARLES (1765) *Traité de la Formation Méchanique des langues, et de Principes Physiques de l'Étymologie*. Paris: Chez Saillant, Vincent, Desaint.

BROWMAN, CATHERINE P. and GOLDSTEIN, LOUIS (1990) 'Tiers in articulatory phonology, with some implications for casual speech.' In KINGSTON, J. and BECKMAN, M. (eds) *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, pp. 341–76.

CAMPBELL, LYLE (1977) *Quichean Linguistic Prehistory*. University of California Publications in Linguistics, No. 81. Berkeley: University of California Press.

CARRÉ, RENÉ and MRAYATI, MOHAMMED (1990) 'Articulatory-acoustic-phonetic relations and modelling, regions and modes.' In HARDCASTLE, WILLIAM J. and MARCHAL, ALAIN (eds) *Speech Production and Speech Modelling*. Dordrecht: Kluwer Academic, pp. 211–40.

CUTLER, ANNE (ed.) (1982) *Slips of the Tongue and Language Production*. The Hague: Mouton.

DARDEN, BILL J. (1970) 'The fronting of vowels after palatals in Slavic.' *Proceedings of the Regional Meeting, Chicago Linguistic Society*, Vol. 6, pp. 459–70.

DELATTRE, PIERRE (1962) 'Some factors of vowel duration and their cross-linguistic validity', *Journal of Acoustical Society of America* **34**: 1141–2.

DEVINE, ANDREW M. and STEPHENS, LAWRENCE D. (1977) *Two Studies in Latin Phonology*. Saratoga: Anma Libri.

DINNSEN, DANIEL A. (1980) 'Phonological rules and phonetic explanation', *Journal of Linguistics*, **16**: 171–91.

DURAND, MARGUERITE (1955) 'Du rôle de l'auditeur dans la formation des sons du langage', *Journal de Psychologie Normale et Pathologique* **52**: 347–55.

EDKINS, JOSEPH (1864) *A Grammar of the Chinese Colloquial Language Commonly Called the Mandarin Dialect*. Shanghai: Presbyterian Mission Press.

ELLIS, ALEXANDER J. (1889) *On Early English Pronunciation*, Part V. London: Trübner.

VON ESSEN, OTTO (1964) 'An acoustic explanation of the sound shift [ɨ] > [u] and [l] > [i].' In ABERCROMBIE, DAVID, FRY, D. B., MACCARTHY, P. A. D., SCOTT, N. C. and TRIM, J. L. M. (eds) *In Honour of Daniel Jones*. London; Longmans, pp. 53–8.

FANT, GUNNAR (1960) *Acoustic Theory of Speech Production*. The Hague: Mouton.

FLICKINGER, DAN P. (1981) 'Dissimilation in Gothic without Thurneysen's Law', *Proceedings of the Annual Meeting, Chicago Linguistic Society*. Vol. 17, pp. 67–75.

FOWLER, CAROL A. (1986) 'An event approach to the study of speech perception from a direct realist perspective', *Journal of Phonetics* **14**: 3–28.

FUJIMURA, OSAMU and LINDQVIST, JAN (1971) 'Sweep-tone measurements of vocal-tract characteristics', *Journal of the Acoustical Society of America* **49**: 541–58.

GAMKRELIDZE, TAMAS V. and IVANOV, VIACHESLAV V. (1973) 'Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse', *Phonetica* **27**: 150–6.

GHAZELI, SALEM (1977) 'Back consonants and backing coarticulation in

Arabic'. Unpub. Ph.D. Dissertation, University of Texas, Austin.
GILL, HARJEET SINGH and GLEASON, HENRY A. JR (1969) *A Reference Grammar of Punjabi*. Patiala: Punjabi University.
GOLDSTEIN, LOUIS (1983) 'Vowel shifts and articulatory-acoustic relations.' In COHEN, A. and BROECKE, M. P. R. (eds) *Abstracts of the Tenth International Congress of Phonetic Sciences*. Dordrecht: Foris, pp. 267–73.
GRAMMONT, MAURICE (1933) *Traité de Phonétique*. Paris: Librairie Delagrave.
GRIERSON, GEORGE (1922) 'Spontaneous Nasalization in the Indo-Aryan languages', *Journal of the Royal Asiatic Society of Great Britain*, July: 381–8.
HADEN, ERNEST F. (1938) 'The physiology of French consonant changes.' *Language Dissertations* No. 26.
HALLE, MORRIS and STEVENS, KENNETH N. (1971) 'A note on laryngeal features.' *Quarterly Progress Report* (Research Laboratory of Electronics, MIT). Vol. 101, pp. 198–213.
HARSHMAN, RICHARD, LADEFOGED, PETER and GOLDSTEIN, LOUIS (1977) 'Factor analysis of tongue shapes', *Journal of the Acoustical Society of America* **62**: 693–707.
HOMBERT, JEAN-MARIE (1978) 'Consonant types, vowel quality, and tone.' In FROMKIN, V. A. (ed.) *Tone: a Linguistic Survey*. New York: Academic Press, pp. 77–111.
HOMBERT, JEAN-MARIE, OHALA, JOHN J. and EWAN, WILLIAM G. (1979) 'Phonetic explanations for the development of tones', *Language* **55**: 37–58.
HOPPER, PAUL J. (1973) 'Glottalized and murmured occlusives in Indo-European'. *Glossa* **7**: 141–66.
HOUSE, ARTHUR S. (1957) 'Analog studies of nasal consonants.' *Journal Speech and Hearing Disorders* **22**: 190–204.
HOUSE, ARTHUR S. and STEVENS, KENNETH N. (1956) 'Analog studies of the nasalization of vowels', *Journal Speech and Hearing Disorders* **21**: 218–32.
JAKOBSON, ROMAN (1978) 'Principles of historical phonology.' In BALDI, PHILIP and WERTH, RONALD N. (eds) *Readings in Historical Phonology*. University Park, PA: Pennsylvania State University Press, pp. 253–60.
JANSON, TORE (1986) 'Sound change in perception: an experiment.' OHALA, JOHN J. and JAEGER, JERI J. (eds) *Experimental Phonology*. Orlando, FL: Academic Press, pp. 253–60.
JAVKIN, HECTOR R. (1979) *Phonetic Universals and Phonological Change. Report of the Phonology Laboratory*. No. 4. University of California, Berkeley.
JHĀ, SUBHADRA (1958) *The Formation of the Maithilī Language*. London: Luzac.
KAWASAKI, HARUKO (1986) 'Phonetic explanation for phonological universals: the case of distinctive vowel nasalization.' In OHALA, JOHN J. and JAEGER, JERI J. (eds) *Experimental Phonology*. Orlando, FL: Academic Press, pp. 81–103.

KENT, ROLAND (1936) 'Assimilation and dissimilation', *Language* **12**: 245–58.
KEWLEY-PORT, DIANE (1982) 'Measurement of formant transitions in naturally produced stop consonant-vowel syllables', *Journal of the Acoustical Society of America* **72**: 379–89.
KEY, T. HEWITT (1855) 'On vowel-assimilations, especially in relation to Professor Willis's experiment on vowel-sounds.' *Transactions of the Philological Society*, Vol. 5, pp. 191–204.
KING, ROBERT D. (1969) *Historical Linguistics and Generative Grammar*, Englewood Cliffs, NJ: Prentice-Hall.
KINGSTON, JOHN C. (1985) 'The phonetics and phonology of the timing of oral and glottal events.' Unpublished PhD Dissertation. University of California, Berkeley.
KLATT, DENNIS H. and STEVENS, KENNETH N. (1969) 'Pharyngeal Consonants', *Quarterly Progress Report, Research Laboratory of Electronics* (MIT) **93**: 207–16.
KLEIN, W., PLOMP, REINIER and POLS, LOUIS C. W. (1970) 'Vowel spectra, vowel spaces, and vowel identification', *Journal of Acoustical Society of America (JASA)* **48**: 999–1009.
LADEFOGED, PETER (1984) '"Out of chaos comes order"; Physical, biological, and structural patterns in phonetics.' In VAN DEN BROECKE, MARCEL P. R. and COHEN, ANTONIE (eds) *Proceedings of the Tenth International Congress of Phonetic Sciences*. Dordrecht: Foris, pp. 83–95.
LASS, ROGER (1980) *On Explaining Language Change*. Cambridge: Cambridge Univeristy Press.
LEA, WAYNE A. (1973) 'Segmental and suprasegmental influences on fundamental frequency contours.' In HYMAN, L. M. (ed.) *Consonant Types and Tone*, Southern California Occasional Papers in Linguistics, No. 1. Los Angeles: University of Southern California, pp. 15–70.
LEHISTE, ILSE (1964) *Acoustical Characteristics of Selected English Consonants*. International Journal of American Linguistics (IJAL) Publication **34**.
LEHISTE, ILSE and PETERSON, GORDON E. (1961) 'Transitions, glides, and diphthongs', *Journal of the Acoustical Society of America* **33**: 268–77.
LIGHTNER, THEODORE M. (1970) 'Why and how does vowel nasalization take place?' *Papers in Linguistics* **2**: 179–226.
LINDBLOM, BJÖRN (1963) 'Spectrographic study of vowel reduction', *Journal of the Acoustical Society of America* **35**: 1773–81.
LINDBLOM, BJÖRN (1986) 'Phonetic universals in vowel systems.' In OHALA, JOHN J. and JAEGER, JERI J. (eds) *Experimental Phonology*. Orlando, FL: Academic Press, pp. 13–44.
LINDBLOM, BJÖRN and SUNDBERG, JOHAN (1971) 'Acoustical consequences of lip, tongue, jaw, and larnyx movement', *Journal of the Acoustical Society of America* **50**: 1166–79.
LÖFQVIST, ANDERS, BAER, THOMAS, MCGARR, NANCY S., and SEIDER STORY, ROBIN (1989) 'The cricothyroid muscle in voicing control', *Journal of the Acoustical Society of America* **85**: 1314–21.
MAEDA, SHINJI (1990) 'Compensatory articulation during speech: evidence

from the analysis and synthesis of vocal-tract shapes using an articulatory model.' In HARDCASTLE, WILLIAM J. and MARCHAL, ALAIN (eds) *Speech production and speech modelling*. Dordrecht: Kluwer, pp. 131–49.

MANN, VIRGINIA A. and REPP, BRUNO H. (1980) 'Influence of vocalic context on perception of the [š] vs [s] distinction', *Perception and Psychophysics* **28**: 213–28.

MASPÉRO, HENRI (1912) 'Études sur la phonétique historique de la langue annamite: les initiales', *Bulletin de l'École Française d'Extrême Orient* **12**: 114–16.

MASSARO, DOMINIC W. and ODEN, GEORGE C. (1980) 'Evaluation and integration of acoustic features in speech perception', *Journal of the Acoustical Society of America* **67**: 996–1013.

MATISOFF, JAMES A. (1975) 'Rhinoglottophilia: the mysterious connection between nasality and glottality.' In FERGUSON, CHARLES A., HYMAN, LARRY M. and OHALA, JOHN J. (eds) *Nasálfest: Papers from a Symposium on Nasals and Nasalization*. Stanford: Language Universals Project, pp. 265–87.

MEILLET, ANTOINE (1964) 'Introduction à l'étude comparative des langues indo-européennes.' (Reprint of 8th edn, 1937). University, Alabama: Univeristy of Alabama Press.

MERINGER, RUDOLF and MAYER, KARL (1895) *Versprechen und Verlesen: Eine Psychologisch-Linguistische Studie*. Stuttgart: Göschensche Verlagsbuchhandlung.

MILLER, GEORGE A. and NICELY, PATRICIA E. (1955) 'Analysis of perceptual confusions among some English consonants', *Journal of the Acoustical Society of America* **27**: 338–53.

MULLER, ERIC M. and BROWN, WILLIAM S., JR (1980) 'Variations in the supraglottal air pressure waveform and their articulatory interpretations.' In LASS, N. J. (ed.) *Speech and Language: Advances in Basic Research and Practice*. Vol. 4. New York: Academic Press, pp. 317–89.

MÜLLER, F. MAX. (1864) *Lectures on the Science of Language* (second series). London: Longman, Green, Longman, Roberts and Green.

O'CONNOR, J. D., GERSTMAN, LOUIS, LIBERMAN, ALVIN M., DELATTRE, PIERRE and COOPER, FRANK S. (1957) 'Acoustic cues for the perception of initial /w, j, r, l/ in English', *Word* **13**: 24–43.

OHALA, JOHN J. (1974a) 'Experimental historical phonology.' In ANDERSON, JOHN M. and JONES, CHARLES (eds) *Historical linguistics*, Vol. II, *Theory and Description in Phonology* (Proceedings of the 1st International Conference on Historical Linguistics. Edinburgh, 2–7 September 1973.) Amsterdam: North Holland, 353–89.

OHALA, JOHN J. (1974b) 'Phonetic explanation in phonology.' In BRUCK, ANTHONY, FOX, ROBERT A. and LAGALY, MICHAEL W. (eds) *Papers from the Parasession on Natural Phonology*. Chicago: Chicago Linguistic Society, pp. 251–74.

OHALA, JOHN J. (1975) 'Phonetic explanations for nasal sound patterns.' In FERGUSON, CHARLES A., HYMAN, LARRY M. and OHALA, JOHN J. (eds) *Nasálfest: Papers from a Symposium on Nasals and Nasalization*.

Stanford: Language Universals Project, 289–316.

OHALA, JOHN J. (1976) 'A model of speech aerodynamics.' *Report of the Phonology Laboratory*, University of California, Berkeley. Vol. 1, pp. 93–107.

OHALA, JOHN J. (1978a) 'Southern Bantu vs. the world: the case of palatalization of labials.' *Berkeley Linguistic Society, Proceedings, Annual Meeting*, University of California, Berkeley. Vol. 4, pp. 370–86.

OHALA, JOHN J. (1978b) 'Phonological notations as models.' In DRESSLER, WOLFGANG U. and MEID, W. (eds), *Proceedings of the 12th International Congress of Linguists, Vienna, 28 August–2 September 1977*. Innsbruck: Innsbrucker Beitrage zur Sprachwissenschaft, pp. 811–16.

OHALA, JOHN J. (1979) 'Universals of labial velars and de Saussure's chess analogy.' *Proceedings of the 9th International Congress of Phonetic Sciences*, Vol. 2. Copenhagen: Institute of Phonetics, pp. 41–7.

OHALA, JOHN J. (1981) 'The listener as a source of sound change.' In MASEK, CARRIES, HENDRICK, ROBERTA A. and MILLER, MARY FRANCES (eds) *Papers from the Parasession on Language and Behavior*. Chicago: Chicago Linguistic Society, pp. 178–203.

OHALA, JOHN J. (1983a) 'The phonological end justifies any means.' In HATTORI, SHIRÔ and INOUE, KAZUKO (eds) *Proceedings of the XIIIth International Congress of Linguists, Tokyo, 29 August–4 September 1982*. Tokyo: Sanseido Shoten, pp. 232–43.

OHALA, JOHN J. (1983b) 'The origin of sound patterns in vocal tract constraints.' In MACNEILAGE, PETER F. (ed.) *The Production of Speech*. New York: Springer-Verlag, pp. 189–216.

OHALA, JOHN J. (1985a) 'Around *flat*.' In FROMKIN, VICTORIA A. (ed.) *Phonetic Linguistics. Essays in Honor of Peter Ladefoged*. Orlando, FL: Academic Press, pp. 223–41.

OHALA, JOHN J. (1985b) 'Linguistics and automatic speech processing.' In DE MORI, RENATO and SUEN, C.-Y. (eds) *New Systems and Architectures for Automatic Speech Recognition and Synthesis*. NATO ASI Series, Series F: Computer and System Sciences, Vol. 16. Berlin: Springer-Verlag, pp. 447–75.

OHALA, JOHN J. (1986) 'Phonological evidence for top-down processing in speech perception.' In PERKELL, JOSEPH S. and KLATT, DENNIS H. (eds) *Invariance and Variability in Speech Processes*. Hillsdale, NJ: Lawrence Erlbaum, pp. 386–97.

OHALA, JOHN J. (1987) 'Explanation in phonology: Opinions and examples.' In DRESSLER, WOLFGANG U., LUSCHÜTZKY, HANS C., PFEIFFER, OSKAR E. and RENNISON, JOHN R. (eds) *Phonologica 1984*. Cambridge: Cambridge University Press, pp. 215–25.

OHALA, JOHN J. (1990a) 'The phonetics and phonology of aspects of assimilation.' In KINGSTON, JOHN and BECKMAN, MARY (eds) *Papers in Laboratory Phonology*, Vol. I, *Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, pp. 258–75.

OHALA, JOHN J. (1990b) 'Respiratory activity in speech.' In HARDCASTLE, WILLIAM J. and MARCHAL, ALAIN (eds) *Speech Production and Speech*

Modelling. Dordrecht: Kluwer, pp. 23–53.

OHALA, JOHN J. (1991) 'What's cognitive, what's not, in sound change.' Paper A297. Linguistic Agency of the University of Duisburg, (LAUD) Universität Duisburg Gesamthochschule.

OHALA, JOHN J. (1992) 'The segment: Primitive or derived?' In DOCHERTY, G. J. and LADD, D. R (eds), *Papers in Laboratory Phonology: Gesture, segment, prosody*. Cambridge: Cambridge University Press, pp. 166–83.

OHALA, JOHN J. and AMADOR, MARISCELA (1981) 'Spontaneous nasalization', *Journal of the Acoustical Society of America* 68: S54–S55 (Abstract).

OHALA, JOHN J. and RIODAN, CAROL J. (1979) 'Passive vocal tract enlargement during voiced stops.' In WOLF, JARED J. and KLATT, DENNIS H. (eds) *Speech Communication Papers*. New York: Acoustical Society of America pp. 89–92.

OHALA, JOHN J. and LORENTZ, JAMES (1977) 'The story of [w]: an exercise in the phonetic explanation for sound patterns.' *Berkeley Linguistic Society, Proceedings, Annual Meeting*, University of California, Berkeley. Vol. 3, pp. 577–99.

OHALA, JOHN J. and SHRIBERG, ELIZABETH E. (1990) 'Hyper-correction in speech perception.' *Proceedings, ICSLP 90* (International Conference on Spoken Language Processing, Kobe 18–22 November 1990), Vol. 1, pp. 405–8.

OHALA, MANJARI (1983) *Aspects of Hindi Phonology*. Delhi: Motilal Banarsidass.

OHALA, MANJARI and OHALA, JOHN J. (1991) 'Nasal epenthesis in Hindi', *Phonetica* 48: 207–20.

OSTHOFF, HERMAN and BRUGMANN, KARL ([1878] 1967) Preface to *Morphological Investigations in the Sphere of the Indo-European Languages I*. In LEHMANN, WINFRED P. (ed. and trans.) *A reader in Nineteenth-Century Historical Indo-European linguistics*. Bloomington: Indiana University Press, pp. 197–209. (Originally published as Preface. *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen I*. iii–xx, Leipzig: S. Hirzel, 1878.)

PERKELL, JOSEPH S. and COHEN, MICHAEL H. (1989) 'An indirect test of the quantal nature of speech in the production of the vowels [i], [a] and [u]', *Journal of Phonetics* 17: 123–33.

PETERSON, GORDON E. and LEHISTE, ILSE (1960) 'Duration of syllable nuclei in English.' *Journal of the Acoustical Society of America* 32: 693–703.

VON RAUMER, RUDOLF (1863) *Gesammelte Sprachwissenschaftliche Schriften*. Frankfurt: Heyder & Zimmer.

DE REUSE, WILLEM J. (1981) 'Grassmann's Law in Ofo', *International Journal of American Linguistics* 47: 243–4.

RIORDAN, CAROL J. (1980) 'Larynx height during English stop consonants', *Journal of Phonetics* 8: 353–60.

ROCK, IRVIN (1983) *The Logic of Perception*. Cambridge, MA: MIT Press.

ROTHENBERG, MARTIN (1968) 'The breath-stream dynamics of simple-released-plosive production', *Bibliotheca Phonetica* 6.

ROUSSELOT, L'ABBÉ PIERRE-JEAN (1891) *Les Modifications Phonétiques du Langage, Études dans le Patois d'une Famille de Cellefrouin (Charente)*. Paris: H. Welter.

SATO, T. (1950) 'On the differences in time structure of voiced and unvoiced stop consonants', in Japanese, *Journal of the Acoustical Society of Japan* 14: 117–22.

SHRIBERG, ELIZABETH E. and OHALA, JOHN J. (1991) '"Correction" in the perception of filtered vowels.' Paper presented at 121st Meeting of the Acoustical Society of America, Baltimore, 29 April–3 May 1991.

SOLÉ, MARIA JOSEP and OHALA, JOHN J. (1990) 'Cognitive representation and the varying pronunciation of frequently occurring expressions.' Paper presented at International Pragmatics Conference. Universitat de Barcelona, 9–13 July.

SOLÉ, MARIA JOSEP and OHALA, JOHN J. (1991) 'Differentiating between phonetic and phonological processes: the case of nasalization.' In *Proceedings of the 12th International Congress of Phonetic Sciences, Aix-en-Provence, 19–24 August, 1991*. Aix-en-Provence: Institut de Phonétique. Vol. 2, pp. 110–13.

STETSON, RAYMOND H. (1928) *Motor Phonetics*. La Haye: Martinus Nijhoff.

STEVENS, KENNETH N. (1971) 'Airflow and turbulence noise for fricative and stop consonants: static considerations', *Journal of the Acoustical Society of America* 50: 1180–92.

STEVENS, KENNETH N. (1980) 'Discussion.' *Proceedings, Ninth International Congress of Phonetic Sciences*, Vol. 3. Copenhagen: Institute of Phonetics, University of Copenhagen, pp. 185–6.

STEVENS, KENNETH N. (1989) 'On the quantal nature of speech', *Journal of Phonetics* 17: 3–45.

STEVENS, KENNETH N. and KEYSER, SAMUEL JAY (1989) 'Primary features and their enhancement in consonants', *Language* 65: 81–106.

STEVENS, KENNETH N., KEYSER, SAMUEL JAY and KAWASAKI, HARUKO (1986) 'Toward a phonetic and phonological theory of redundant features.' In PERKELL, JOSEPH R. and KLATT, DENNIS H. (eds) *Invariance and variability in speech processes*. Hillsdale, NJ: Erlbaum, pp. 426–49.

SVANTESSON, JAN-OLOF (1983) *Kammu phonology and morphology* (Travaux de l'institut de linguistique de Lund 18). Lund: Gleerup.

SWEET, HENRY (1874) *History of English sounds*. London: Trübner.

SWEET, HENRY (1877) *A Handbook of Phonetics*. Oxford: Clarendon Press.

TRAUNMÜLLER, HARMUT (1981) 'Perceptual dimension of openness in vowels', *Journal of the Acoustical Society of America* 69: 1465–75.

TURNER, SIR RALPH L. (1921) 'Gujarati phonology', *Journal of the Royal Asiatic Society* pts 3–4.

TURNER, SIR RALPH L. (1923–25) 'The Sindhi recursives or voiced stops preceded by glottal stop', *Bulletin of the School of Oriental Studies* 3: 301–15.

VERNER, KARL (1913) (Letters to Hugo Pipping). *Oversigt over det kongelige Danske Videnskabernes Selskabs Forhandlinger*. No. 3. 161–211. Copenhagen.

WAIBEL, ALEX and LEE, KAI-FU (eds) (1990) *Readings in Speech Recognition*. San Mateo, CA: Morgan Kaufmann.

WANG, WILLIAM S.-Y. (ed.) (1977) *The Lexicon in Phonological Change*. The Hague: Mouton.

WEISMER, GARY (1979) 'Sensitivity of voice-onset-time (VOT) measures to certain segmental features in speech production', *Journal of Phonetics* 7: 197–204.

WESTBURY, JOHN R. (1979) 'Aspects of the temporal control of voicing in consonant clusters in English.' PhD Dissertation, University of Texas.

WESTBURY, JOHN R. and KEATING, PATRICIA A. (1985) 'On the naturalness of stop consonant voicing.' *Working Papers in Phonetics* (UCLA) Vol. 60, pp. 1–19.

WHITNEY, WILLIAM DWIGHT (1867) *Language and the Study of Language*. New York: Charles Scribners.

WINITZ, HARRIS, SCHEIB, M. E. and REEDS, J. A. (1972) 'Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech', *Journal of the Acoustical Society of America* 51: 1309–17.

WRIGHT, JAMES T. (1986) 'The behavior of nasalized vowels in the perceptual vowel space.' In OHALA, JOHN J. and JAEGER, JERI J. (eds) *Experimental phonology*. Orlando, Florida: Academic Press, pp. 45–67.

YOUNG, THOMAS (1819) 'Remarks on the probabilities of error in physical observations and on the density of the earth considered especially with regard to the reduction of experiments on the pendulum.' *Transactions of the Philosophical Society*, Cambridge, pp. 70–95.

ZUE, VICTOR W. (1976) *Acoustic Characteristics of Stop Consonants: a Controlled Study*. Bloomington, IN: Indiana University of Linguistics Club.