

73200255

LINGUISTIC EVOLUTION

WITH SPECIAL REFERENCE
TO ENGLISH

M. L. SAMUELS

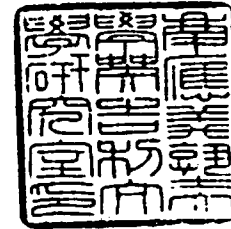
*Professor of English Language
University of Glasgow*

CAMBRIDGE STUDIES IN LINGUISTICS

General Editors · W. SIDNEY ALLEN · EUGENIE J. A. HENDERSON ·
FRED W. HOUSEHOLDER · JOHN LYONS · R. B. LE PAGE · F. R. PALMER ·
J. L. M. TRIM

- 1 DAVID CRYSTAL: *Prosodic systems and intonation in English*
- 2 PIETER A. M. SEUREN: *Operators and nucleus*
- 3 RODNEY D. HUDDLESTON: *The sentence in written English*
- 4 JOHN M. ANDERSON: *The grammar of case*
- 5 M. L. SAMUELS: *Linguistic evolution*
- 6 P. H. MATTHEWS: *Inflectional morphology*

EL 英
8B
321



CAMBRIDGE
at the University Press · 1972

Published by the Syndics of the Cambridge University Press
Bentley House, 200 Euston Road, London NW1 2DB
American Branch: 32 East 57th Street, New York, N.Y. 10022

© Cambridge University Press 1972

Library of Congress Catalogue Card Number: 72-176255

ISBN: 0 521 08385 0

Printed in Great Britain by
Alden & Mowbray Ltd
at the Alden Press, Oxford

Contents

<i>Preface</i>	page vii
<i>Abbreviations</i>	ix
1 Introduction	i
2 Variation in the spoken chain	9
3 System: the phoneme	28
4 Grammar and lexis (i): variation	49
5 Grammar and lexis (ii): systemic regulation	64
6 Diversity and contact	88
7 An assessment of the problem	135
8 Dimensions of study	154
9 Conclusions	177
<i>Bibliography</i>	182
<i>Index of persons</i>	187
<i>Index of subjects</i>	189

Preface

It is often said nowadays that we have no theory of historical linguistics, or that the existing theories are unworkable (though it should be added that some of such statements are prompted by genuine modesty and by a desire to disclaim the ability to provide a fresh theory). But to the present writer it has often seemed that the objections raised to previous theories are pseudo-objections, and that it is unduly pessimistic to say we have no theory at all. Admittedly, a theory that could predict details would be impossible, since the interpenetration of extra- and intralinguistic factors is largely a matter of chance. But it is possible to form a theory that is somewhat less ambitious – one that will, for example, make general predictions, indicate what kinds of evidence to look for, and suggest how it is to be interpreted. The present book results from an attempt to provide a theory of this latter kind. Since it uses previous theories as its starting point, it makes no claim to provide either a revolutionary clean sweep or a panacea; but the resulting principles will, it is hoped, give a fair reflection of the 'why' as well as of the 'how' of linguistic change.

In chapter 2, the examples used are for the most part familiar ones from the better-known European languages. It would have been possible to add others from African and American-Indian languages, the study of which has advanced so greatly since 1945; but this might have run the risk of obscurity for all but specialists in those particular languages. Again, while a book of this scope cannot include examples of all the situations and processes that could arise, it is my hope that the principles outlined can be either applied or adapted to special cases in lesser-known languages.

It may be felt by some readers that this book could have been better expressed in generative terms. My reasons for not doing so refer to generative phonology rather than to transformational grammar proper, and they are twofold. Firstly, to have used it at the stage to which it had developed from 1962 to 1968 would have obscured, rather than clarified, my argument. Secondly, as a procedure it seems to be less concerned with the *causes* of linguistic changes than with their description. Indeed, in 1968 P. M. Postal could still write 'there is no more reason for languages to change than there is for automobiles to add fins one year and remove them the next, for jackets to have three buttons one year and two the next'.¹ Only recently have there been

¹ *Aspects of Phonological Theory* (New York, 1968) 283.

signs that generative phonology is being adapted to fit the facts of heterogeneity in both time and space.¹

It is perhaps natural that, in attempting to bring order to a difficult subject, an author may give the impression of oversimplifying and minimising outstanding problems. I have tried in this book to place into perspective those aspects of the subject that I regard as most relevant; but I am only too aware that some problems are unanswerable, and that our tentative answers to others may be suddenly reversed by the accession of new evidence. The very nature of the subject demands continual caveats and qualifications; if I have occasionally omitted them, I can only plead the exigencies of attempting to present an ordered and coherent argument.

The text of this book was largely completed by 1970, and I was therefore unable to take account of B. M. H. Strang's important *History of English* which appeared in that year.

Finally, I should like to record my indebtedness to W. S. Allen, L. W. Collier, and R. B. Le Page for helpful comments and corrections. For the shortcomings that remain, the responsibility is wholly my own.

University of Glasgow
June 1972

M. L. S.

¹ E.g. in B. E. Newton's recent 'Ordering Paradoxes in Phonology', *Journal of Linguistics* 7 (1971) 31-53; and in publications for the Project on Linguistic Analysis led by W. S-Y. Wang of the University of California at Berkeley.

Abbreviations

(a) Books and periodicals

Books referred to in footnotes by name of author only are listed in the Bibliography on pp. 182-86. If more than one work for a single author is listed, the reference includes the date of publication. The following other abbreviations are used:

<i>AL</i>	<i>Archivum Linguisticum</i>
<i>EAGS</i>	<i>English and Germanic Studies</i>
EETS OS	Early English Text Society, Original Series
<i>ES</i>	<i>English Studies</i>
<i>JEGPh</i>	<i>Journal of English and Germanic Philology</i>
OED	Oxford English Dictionary
<i>PMLA</i>	<i>Publications of the Modern Language Association of America</i>
<i>RES</i>	<i>Review of English Studies</i>
<i>TPS</i>	<i>Transactions of the Philological Society</i>

(b) Languages and dialects

E.	English	MHG	Middle High German
EMnE	Early Modern English	Mn	Modern
Fr.	French	OE	Old English
G.	German	OFr.	Old French
Gk.	Greek	OHG	Old High German
Gmc.	Germanic	ON	Old Norse
Goth.	Gothic	ONhb.	Old Northumbrian
IE	Indo-European	RP	Received Pronunciation
It.	Italian	Sc.	Scots, Scottish
L.	Latin	Skt.	Sanskrit
ME	Middle English	Sp.	Spanish

I Introduction

1.1 Change and evolution

The title of this book ('Linguistic Evolution') was chosen in preference to 'Linguistic Change' although it is about linguistic change. That is because its purpose is to attempt an examination of the large complex of different factors involved, and the title 'Linguistic Change' might have entailed an oversimplification. Nevertheless 'evolution' is itself open to the misunderstanding that some sort of progress is implied, that a clearer or more effective means of communication has been achieved as a result of it. That meaning of 'evolution' is not intended here. We are not concerned here with the prehistoric origins of human language, and, as has often been pointed out, there is today no such thing as a 'primitive' language; every language is of approximately equal value *for the purposes for which it has evolved*, whether it belongs to an advanced or a primitive culture. For English, this can be shown, among other ways, by the fact that in earlier stages certain grammatical distinctions could be made that can no longer be made today (see 4.1 below).

In what sense, then, can a language be said to 'evolve'? We must recognise, firstly, that there is no inevitable one-for-one correspondence between a language and the culture it serves, and that the language need not be a precise mirror of the culture as it exists at any given time. Indeed, the so-called Whorfian hypothesis leaves open the possibility that an individual language, so far from merely accommodating or mirroring a culture's channels of thought, may to some extent even guide and control them, and so, no doubt, in many minor ways, it does. Viewed historically, however, it seems likely that those features in which language today apparently influences culture are relics in the language from a previous stage of the culture. For example, much is made of the Scandinavian terms of relationship, which distinguish automatically, at word-level, between grandparents, according to whether they are on the father's or the mother's side, and so also with uncles and aunts. Yet sociologists are not aware of any striking differences in the *actual* relationships in Scandinavia compared with those of speakers of other Germanic languages. Certainly, they were of special importance in all Germanic-speaking societies 1500 years ago, and their importance may have survived longer in Scandinavia than elsewhere; but today, the survival of the relevant linguistic terms does not seem to have ensured the survival of the corresponding modes of thought.

On the one hand, therefore, language must keep pace with the culture,

progress in knowledge, and changing fashions in the society it serves; but on the other, it may retain many outdated features that are relics of past stages of the culture. There is thus a considerable margin of tolerance to both its contents and its limits.

But there are many changes in language that are not obviously the result of such extralinguistic factors. They may have been triggered by these factors, but to what extent and in what ways is not always clear, and the possibility remains that they arise internally for no other reason than that language, as a tool in constant use is (a) liable to constant fluctuation, and/or (b) in need of constant 'repair' or 'renewal'. The result of changes due to such (apparently) intralinguistic factors parallels those that result from extralinguistic: a series of outdated forms that, from the modern point of view, are 'anomalous'. For example, the alternation corresponding to *stand* – *stood* was regular in the Indo-European system, and so with that corresponding to *seek* – *sought* in Primitive Germanic, and with that corresponding to *feel* – *felt* in Old English. Today, when the only productive alternation is that seen in *move* – *moved*, all these relics of previous systems are more or less anomalous and therefore lessen the efficiency of the modern system. According to Professor A. Martinet, they are due to forces of conservatism and 'inertia', to which are continually opposed the pressures that work towards clearer communication. If that is a fair statement of the position, it would appear, on the one hand, that both externally and internally in language development an equilibrium must be maintained between the old and the new, between the inherited forms and the expressive needs of the present; but at the same time, that the margin of tolerance and choice in the maintenance of that equilibrium is probably fairly wide, except, that is, for certain special points of pressure that vary in each era. It is in that sense that 'evolution' is intended here.

1.2 'Causation'

How far is it either permissible or useful to speak of 'causation' in linguistic change? Some scholars have objected that changes are inevitable and wholly to be expected, and that we should therefore not seek causes or ask 'why', but only investigate the processes and ask 'how'.¹ If the suggestion made above is correct, that the changes are (a) partly to be expected because language is a tool used by a continually changing society, but also (b) partly unpredictable because there are wide limits of tolerance for the direction of change, it then follows that we are attempting both: we are enquiring, firstly, under what conditions the expected changes take place, and secondly, what factors condition the unpredictable changes.

¹ Spence p. 27; cf. Jakobson p. 651 on 'mutability as a constant essential component of any phonetic system'.

There is no logical objection to the term 'cause', but it has the disadvantage that it still carries with it the connotation of a single, indivisible cause. If, therefore, 'cause' is to be preferred stylistically to the more clumsy 'conditioning factor', that connotation must be discounted: it will be one of the main purposes of this book to show that, although one factor may often outweigh others in importance, an *insistence* on a single cause – what Jespersen called 'the all-or-nothing fallacy'¹ – has been, in the past, one of the main drawbacks to progress in diachronic linguistics. In both pure and applied science, the principle of multiple conditioning is a commonplace. Yet in the study of linguistic change it has been treated with suspicion; for long, the linguist who suggested a combination of causes operating in the past was accused of ambivalence. It is only more recently that the controlled study of modern usage has shown that such explanations are nearer the truth than those propounding a single, indivisible cause. In a pioneering article of 1957,² Professor Randolph Quirk demonstrated, from a frequency survey, the factors that influence a speaker in his choice between the relative pronouns *which* and *that*, e.g. position in the clause, length of clause, subject or object function, nature of antecedent; and a remarkable finding, supported by his figures, was that factors operating in combination are more potent than those operating singly. It would be unwise, therefore, to assume that a principle that applies to linguistic change today could not have applied to that of the past; and since, as will be shown later, every linguistic change involves at least some degree of choice, the probability of multiple conditioning must constantly be borne in mind.

By the same token, we must not start with preconceptions that parts of the linguistic process are more important than others, and that these must be given priority in the study of change. For some, at present, phonology is less important than grammar or lexis, being merely 'an input-output device that accepts a terminal string with a labelled bracketing and codes it as a phonetic representation'.³ For others it is grammar that is simpler and less important, being merely a programming device for handling individual items in phonemic shape selected from a lexically based memory store.⁴ But for diachronic purposes it would be wholly misleading to assume any such primacies. Though we know little enough about the brain, we do know that it has capacity to handle complexities at *every* level of the linguistic process. There is, therefore, no reason why changes associated with a particular level should be more or less complex than others: we must start by expecting them to have

¹ Jespersen (1922) 262.

² Quirk (1957), and (1968) 94–108.

³ N. Chomsky, quoted by E. C. Trager in *Papers in Honor of Leon Dostert*, ed. W. M. Austin (The Hague, 1967) 159.

⁴ R. C. Oldfield in *Psycholinguistics Papers*, ed. J. Lyons and R. J. Wales (Edinburgh, 1966) 23.

equal importance and complexity. If the evidence of linguistic evolution suggests otherwise, then that evidence, too, may be used for elucidating the basic processes of language and thought.

1.3 The evidence

The substance of language, from a descriptive point of view, exists in two equally valid and autonomous shapes – spoken and written. As a code, each exists in its own right, and there has been justifiable insistence, in recent decades, that graphetics and graphemics must in the first instance be studied separately, before their relationship with phonetics and phonemics can be considered. This according of equal status to the written language is not to deny that it is ultimately derived from, and dependent on, the spoken language; but it is necessary, if only to disprove the naive speaker's notion that the written language is a mirror-image of what he speaks (or vice versa), or that it is based on contemporary spoken language, and not the spoken language of various periods in the past.

Having granted that for synchronic purposes the two must be carefully distinguished, we find that for diachronic purposes our choice of evidence is far more limited: we must for the most part reconstruct from written records, bearing in mind continually the differences that exist today between spoken and written media. Least difficulty arises in the study of the spoken language of the immediate past, which can be based on comparisons of the speech of older and younger generations; and even for past periods, a certain minimum can be reconstructed, by comparative method, from the present spoken forms of cognate dialects and languages. But for the rest, our only source of detailed knowledge is from written records, and the disadvantages of these are well known. A very small proportion of them – for some periods only – contain useful observations on the spoken language by grammarians, orthoepists, phoneticians and certain others, e.g. a literary writer like Swift with an interest in language; but, apart from the sometimes questionable evidence of rhymes, the vast majority of written records from the past provide no evidence for the spoken language beyond what can be deduced from the establishment of the usual correlations of graphemes with segmental phonemes. To hypothesise further, various procedures are adopted. For example, it is necessary to compare, if possible, specimens from the same dialect and period that differ in style, in order to establish which features are more likely to occur in informal style and may therefore be assumed to have been commoner still in the spoken language. Deviations from the normal written language that seem to arise from carelessness or illiteracy may provide similar hints. Throughout, features of the writings that are obviously due to changing literary fashions and

affectations of style must in the first stages be discounted, and the probable degree of archaism in the written language must be assessed.

Even so, we shall know nothing of the suprasegmental features of stress and intonation. A few hints can be gathered from verse structure, but the main evidence is indirect, consisting of those variations and changes of segmental phonemes that seem best explained as due to a weakening of stress in some positions. Here, even more than elsewhere, we are extrapolating from known distributions in languages spoken today.

A further complication is that, for preliterate periods, we must reckon without any written language at all, even though the language survives in written form from an early date. In other words, we have to envisage that a language that had till then, or till an indeterminate time before then, developed as a spoken language only, without conservative influences from the written language, was then first committed to written form; we must estimate the degree of stylisation that was first adopted, and which may, for all we know, have created an immediate gulf between the two media.

In practice, therefore, there is little chance of keeping the two media distinct. But even if that were possible, it is by no means certain that it would be methodologically desirable. Though there are many differences of convention, the two media both expound what is essentially 'the same language'; and although, for the purpose of describing them at a given point of time they must be kept separate, it may be that for studying their development over a period of time, the interactions between the two media are just as important as their developments as entities. Certainly there is a separate continuity of change and development within each medium; but, unless the written language is of the ideographic or logographic type, the two must keep in step, for, if the level of correspondences between them drops, the written language becomes an uneconomic medium, used only by those few in the community who can afford the time to master it. Yet the artificial regulation of orthographies takes place comparatively rarely; the similarity of the two media is preserved by normal interactions, of which the more typical are shown in fig. 1. The broken lines A, A represent continual mutual interactions, consisting mainly of (i) conservative influence of the written language, inhibiting change in the spoken, (ii) exact written representation of new lexical intake into the spoken language, and its converse (iii) exact spoken representation of new lexical intake into the written language.

The full line B represents adjustments in the written language made later than the corresponding changes (e.g. phonemic) in the spoken. The full line C represents the influence of conservative spellings on the spoken language, resulting in 'spelling pronunciations' which may or may not reproduce an earlier spoken form: e.g. /fɔ:rhɛd/ is produced afresh from the spelling *forehead*, and replaces /fɔ:rd/, the reflex of an earlier pronunciation which had

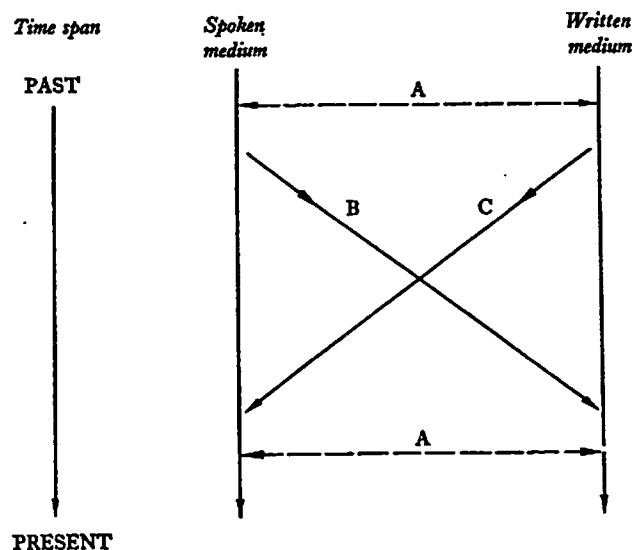


Fig. 1. Interactions between written and spoken media

been closer to the spelling, though not necessarily the same as the new pronunciation.

The scope and types of interaction that have to be taken into account are certain to be much wider and more complex than can be expressed diagrammatically. For example, prescriptivity in grammar acts in part directly from schoolteacher to children, but also indirectly through the written language (cf. 6.6, p. 110). A different kind of complexity occurs where there is a wide gulf between the classical and modern forms of a language (as in the case of Sinhalese): the literary form may influence only the formal register of the spoken language, not the colloquial. For our present purpose, it must suffice to assume that (a) a majority of linguistic changes arise in the spoken language, and may or may not ultimately spread to the written medium; (b) certain (though fewer) changes originate in the written language, and may or may not spread to the spoken medium; and (c) the main influence of the written language is a conservative one – it acts as a brake, inhibiting the general acceptance of many changes that arise in the spoken language.

In subsequent chapters it will be mainly the spoken language that is under discussion; and if the evidence for it consists of written records, they are used subject to the provisos stated above.

1.4 Basic distinctions and approach

The main distinction between internal (intralinguistic) and external (extra-

linguistic) factors has been referred to above. But there is another distinction that is liable to be confused with it if it is not carefully defined – that between *intrasystemic* and *extrasystemic*. The term system is discussed further in 3.1 and 6.7 (p. 129), but we may start by defining intrasystemic factors as those that operate within a single linguistic system, whether it is regarded as belonging to a small group (subdialect) or to a larger area that is less than that covered by a language (dialect), or to the total of these (the language itself); and extrasystemic factors as those that influence a system from outside it as a result of contact with another system, which may be that of another dialect or subdialect of the same language, or belong to an entirely different language.

The confusion between the two sets of terms arises mainly because, even though systems refer only to linguistic systems, everything that is extralinguistic must also, in a strictly literal sense, be extrasystemic (this difficulty does not arise with the term intrasystemic). But this overlap is of no use to our investigation, and will be excluded by restricting the term extrasystemic to refer to intralinguistic only, as follows:

Intralinguistic	{ Intrasystemic Extrasystemic ¹ (linguistic sense only)
Extralinguistic	[Extrasystemic ² (non-linguistic sense, here not used)]

This restriction enables us to separate the two sets, but, even so, we must expect them to co-occur in varying proportions according to situation, as in the following examples:

Example 1. An assimilatory change (e.g. /lɛŋθ/ > /lɛnθ/) arises independently in one community, and spreads to a neighbouring community: the origin of the change is both intralinguistic and intrasystemic, but its spread – viewed from within the second (receiving) system – is extrasystemic.

Example 2. A name is coined for a new invention (e.g. *typewriter*, *ballpoint*). The origin is extralinguistic, but the coinage is made from intrasystemic sources. If the invention is then introduced to a neighbouring community under the same name an extralinguistic factor (introduction of the invention) is still present, but the introduction of its name constitutes an extrasystemic but intralinguistic influence on the receiving system.

Example 3. The rise of a new standard language is extralinguistic, since a certain degree of movement or interchange in the population is a prerequisite. But the contact of different dialects that produces the standard is itself intra-

linguistic, and the process can be studied either intrasystemically, with attention focussed on the single – though as yet not homogeneous – system of the rising standard, or extrasystemically, with emphasis on the various influences of the contributing dialects.

In what follows, the distinction between extra- and intralinguistic will usually be assumed as axiomatic, but that between extra- and intrasystemic will constitute one of the main divisions of approach. In chapters 2–5 the causes of change are viewed as operating within a single system, while the extrasystemic aspect is reserved for chapter 6. The other main division of approach is between (a) *parole*-based variation, i.e. sources of variation at the level of idiolect and (b) functional pressure operating at the level of system. *Parole*-based variation will often be referred to as ‘mechanical’, but will include all variants, whether arbitrary or motivated, and irrespective of whether they are eventually adopted into the system or not; they will be treated as ‘intrasystemic’, but only in the broad sense that they occur in the speech of the individual users of a system. The sources of variation will be discussed in chapters 2 and 4, and functional pressure in chapters 3 and 5; but a number of problems that remain unresolved in these chapters will be subsumed in the latter half of chapter 6. In chapters 7 and 8, an attempt is made to formulate a theory of linguistic change and a method for studying all aspects of it simultaneously.

2 Variation in the spoken chain

2.1 Introductory

Human language is extremely highly developed compared with that of animals, and indeed it is probably the main reason for the development of the human brain;¹ but that does not mean that it is mechanically the best and most efficient means of communication that could be devised. The arbitrariness of its chosen combinations of sounds, the necessity of producing them consecutively with organs that were not originally evolved for the purpose, the fluidity of the word-concept spectrum – these are all typical disadvantages that might be held to support the old popular idea that languages naturally ‘degenerate’. But that is now outdated, and it is recognised that all types of language are equally valid as communication, provided that they are used in their appropriate place or social milieu (*dialects*), or for the appropriate purpose (*registers*: cf. 6.2). The modern view carries with it the assumption that, in any of these types of a language, potential breakdown in communication is constantly and automatically remedied. It follows that, in any investigation of linguistic change, the first question to be asked is: are linguistic changes purely mechanical in origin, arising from the very nature of the substance of the spoken chain? As has often been noted,² success in performing acts of communication varies greatly, and, overall, comparatively few ‘bull’s-eyes’ are scored. This is possible because (a) context and situation, and (b) the redundancy inherent in all linguistic systems, together provide a wide margin of tolerance for the understanding of actual, as distinct from ideal, utterances or writings. For example, the realisations of a single phoneme /t/ in speech might include not only [t^h, t^ʰ, t^ʰ] but also [p] and [k] without breakdown in communication; or the realisations of written *t* might be much nearer certain forms of *l* or *r* and still be understood as the writer intended. Such a margin of tolerance is by no means limited to sounds (phones) or letters (graphs); it applies equally well to grammar and lexis, where the choice of a word or construction may not be sufficiently accurate for the context, or may include, e.g., solecisms or inappropriate metaphors.

The number of variants (phonetic, grammatical or lexical) that are continually produced thus is vast, and it contains the ‘raw material’ of all linguistic

¹ At least, this seems to be the more likely of the two possible deductions that one can make from the fact that both language and a developed brain coincide in man alone.

² Paul p. 43; Hockett p. 440.

change. As regards the step from variant to change, we may adopt, for the present, a simple formula: the variant is misunderstood as an acceptable form (or reinterpreted) by the hearer, and, when this has happened often enough and been subsequently imitated, it ceases to be merely part of *parole* (sporadic usage in the speech of individuals) and is accepted into *langue* (the language, either in abstract or as codified).

But are the variants any more than the raw material of change? Can any universals be detected that would account for linguistic change as due solely to the tolerated margin of aberration? The remainder of this chapter will deal with the phonetic aspect of this problem, the grammatical and lexical aspects being postponed to chapter 4.

2.2 Conditioned phonetic change: evidence

The phonetic study of modern speech shows widespread use of variants according to context, e.g. at word-boundaries, as in /gʊg ɡəl/ 'good girl', /tɛm paʊndz/ 'ten pounds', /hæz ʃi/ 'has she'; within the word, as in /əʊpɪm/ 'open', /lɛnθ/ 'length'; or in unstressed position, as in /ʃəd/ 'should', /hi:d/ 'he would' or 'he had', as well as in the orthographically established *he'd*, *shan't*, *won't*, *can't*, etc. In all such cases, it is evident that short cuts in articulation have been taken: those variants that show phonetically irreversible changes towards simplification are those used in colloquial (i.e. 'less careful') register, while variants not showing such changes are those used in formal register (besides the general distinction of formal-colloquial, other oppositions have been suggested which may or may not co-occur but can produce the same result: forceful-relaxed, careful-slurred, decreased tempo-increased tempo, lento-style-allegro-style). The existence of the written language is a strong influence towards the preservation of formal variants, but that does not alter the basic probability that the synchronic distinction 'careful-careless' reflects a diachronic process that stems from ease of articulation ('the principle of least effort').

Ever since Saussure's insistence that the diachronic and synchronic planes be kept separate, many linguists have been reluctant to extrapolate from synchronic evidence in this way, but it is not easy to see why. The principle of least effort has been amply demonstrated; its most obvious manifestation is that shorter linguistic forms are preferred to longer ones, and that those words which, by formative accident, are inconveniently long for frequent use (*telephone*, *omnibus*, *underground*) are subsequently replaced by shorter forms (*phone*, *bus*, *tube*).

It may be objected that the existence of two variants today provides no proof of change in the past; that all we know is that the same individual says

/gəʊɪŋ tə ɡəʊ/ in formal register but /gənə ɡəʊ/ in colloquial register, and that the precise diachronic link between the two escapes us. It is true that we are unable to tell whether a given form continually undergoes fresh and repeated reduction by each new user, or whether it is a repetition of the same form learned from others. But this argument is relevant only to the *chronology* of the change. For example, even where there is historical evidence for full and reduced variants in the same text (as in Shakespeare's *Ile/I will*, *Ild/I would*), the reduced forms may have existed much earlier in the spoken language, the only limit here being provided by the Southern ME *ich*, which combined in corresponding but quite different contractions *ichil* and *icholde*. Apart from such uncertainties of chronology, the fact remains that the only possible etymologies for the two variants leave us with no alternative but to assume that one variant has 'come from' the other; and the assumption is confirmed by the overall absence of other variants to disprove it.

Again, if it is objected that the process of change from one form to another in the language as a whole cannot be treated as parallel to a pair of register-variants today, we may answer that the process assumed for the past *is* parallel – that assimilated or contracted forms existed at first as colloquial register-variants beside the full forms, but subsequently spread beyond colloquial register. This problem will be discussed more fully in connection with extrasystemic influences (ch. 6). For the present, we are restricting the argument to those forms or variants in a single system for which reasonably certain etymologies can be deduced from forms in a closely related antecedent system. On that basis alone, we may agree with E. Coseriu that the Saussurean dichotomy has long been interpreted too rigidly: the language itself is continually changing, and it is only synchronic *descriptions* of it that are static and give us, as it were, a still photograph of the moving object.¹

In principle, therefore, and subject to further conditions that will appear in chapter 6, it is reasonable to assume that the same process, prompted by ease of articulation, is to be seen (a) in certain changes evidenced only in successive stages of a language and shown, e.g., in a formula like *eall swa* > *al swa* > *also* > *alse* > *als* > *as*, and (b) in contemporary variants like *cannot*–*can't* that belong respectively to 'careful' and 'careless' varieties of language.

Finally, it should be noted that variants arise *in spite of* the fact that every speaker can check or 'monitor' his output from the feedback through his hearing mechanism, i.e. there is less care in checking as well as in articulation. Whether this fact lends support to the theory that most phonetic variants arise in childhood, when monitoring of feedback has not yet been perfected, must at present remain an open question (cf. 6.7); but it confirms that, for studying the origin of variants, articulatory criteria take precedence over acoustic (cf. 3.2).

¹ Spence p. 24.

2.3 Conditioned phonetic change: extent

The relation of conditioned change to the principle of least effort has been discussed at some length because the so-called 'ease-theory' of linguistic change has so often been questioned in the past. Obviously it cannot provide an explanation for all change; but its part is by no means negligible, as may be seen from the following outline of the main types of conditioned change.¹ (The examples that will be given are of changes that could take place within a single system as a result of inertia, but this is not to deny that the same changes could take place from extrasystemic contact simply because the structure of the receiving language does not contain certain sequences.)

(i) Assimilatory phenomena

The articulation of a segment is brought nearer to that of those adjoining it. The more important types are:

(a) Regressive assimilation of consonants, as in L. *adcedere* > *accedere*, *sec + mentum* > *segmentum*, **scribunt* > *scriptum* > It. *scritto*, OE *heahre* > *hearre*, *cyþde* > *cydde*, *wifman* > *wimman*, *æmete* > *ant*, E. dial. *lopster* 'lobster', /dluv/ 'glove', /tlo:z/ 'clothes', /bretθ/ 'breadth', Sc. dial. /tneif/ 'knife'.

(b) Progressive assimilation of consonants, as in OE **bitþ* > *bitt* 'bites', *pisre* > *pisse*, ON **gulþ* > *gull* 'gold', OE *eln* > *ell*, ME *mylne* > *mill*, ME *wiltu* 'wilt thou'; cf. /tj/ > /tʃ/ in *feature*, *nature*, /dj/ > /dʒ/ in *did you*?

(c) Influence of consonants on vowels, e.g. of /w/ in OE *sweord* > *sword*, *weoruld* > *woruld*, *wille* > *wylle* (ME *wol*, cf. *will-won't*), ME *wimman* > *woman*, MnE *want*, *swan*, *quality*, E. dial. /twɔv/ 'twelve', cf. G. *zwölf* < MHG *zwelf*; of /l/ in /tʃoldrən/ 'children'; of /r/ in ME *kerve* > *carve*, *sterre* > *star*, MnE *burn*, *bird*, *earth*.

(d) Influence of vowels on consonants, e.g. palatalisation of /k, g/ in L. *caelum* > Fr. *ciel*, *centum* > *cent*, *canem* > *chien*, *cantare* > *chanter*, *causa* > *chose*, *argentum* > *argent*; E. *chin*, *cheese* (cf. G. *Kinn*, *Käse*), *yellow*, *yield* (cf. G. *gelb*, *gelten*).

(e) Voicing of intervocalic consonants (i.e. voicing ceases to be interrupted), e.g. American English [d] in *butter*, *sitting*; Latin intervocalic /p, t, k/ as in *ripa*, *vita*, *mutare*, *pacare*, *saponem* voiced to /b, d, g/ in the early western

¹ The terms 'conditioned' and 'isolative' are retained here. The introduction of distinctive feature analysis (and of prosodic analysis in the Firthian sense) has tended to blur the distinction between the two, and it is not used here as one that can be regarded as clear-cut throughout. Nevertheless, the distinction between features that belong to whole clauses and those that are limited to segments or syllables (or occasionally sequences of syllables within a word) remains an important one for the tracing of historical developments, and, as will be suggested below, bears a close distributional relationship to the traditional distinction between 'isolative' and 'conditioned'.

Romance languages, and to some extent in Italian (e.g. *pagare*). Cf. (f) below.

(f) Loss of plosion in intervocalic consonants, so that the flow of breath continues through three segments: Western Romance /b, d, g/ (as in (e) above) > /β, ð, γ/, preserved in Spanish and Portuguese (*riba*, *vida*, *mudar*, *pagar*, Sp. *jabon*, Port. *sabão*) but with loss or further change in French to *rive*, *vie*, *muer*, *payer*, *savon*. Cf. ME *fader*, *hider*, *togedere* > *father*, *hither*, *together*. 'Lenition' of plosives between voiced segments in Celtic may also be classified here, though it is more extensive in that it includes 'initial mutation' conditioned by a preceding word-ending (cf. (k) below).¹

(g) Replacement of non-initial /t/, and to some extent of /p, k/, by [ʔ] (i.e. plosion is preserved, but is performed by the vocal cords themselves, not by tongue or lips): regional and colloquial English, not only intervocalically, as in [bʌʔə] 'butter' [kɛʔəl] 'kettle', but also before another consonant, as in [kʌʔfʊl] 'cupful', [fʊʔbɔ:l] 'football', and finally, as in [ʌʔ] 'up', [bʊʔ] 'book'.

(h) Unvoicing of final consonants (i.e. voicing is not maintained till the end of the word, especially before a pause): ME (W. Midl.) *hont* 'hand', *lomp* 'lamb', E. dial. /nʌθɪŋk/ 'nothing'; /-p, -t, -k/ in G. *Weib*, *Lied*, *Tag*, etc.

(i) Vocalisation of consonants, especially of [ʃ] to [o] or [u] as in [fi:ʊd] 'field', [mɪrɒk] 'milk'. Cf. L. *alterum* > Fr. *autre*, *parabola* > **parawla* > *parole*, OE *dæg* > *day*, *dragan* > *draw*, *hafoc* > *hawk*, Sc. dial. /bɔ/ 'ball', /fʊ/ 'full'.

(j) New articulation resulting from fusion, as in OE **bidþ* > *bitt* 'awaits'. Cf. /sʃ/ > /ʃ/ as in ME *nacioun*, MnE /neɪʃən/, and /zj/ > /ʒ/, as in ME *visioun*, MnE /vɪʒən/.

(k) Alternation of forms resulting from different assimilations at word-boundaries ('sandhi'), e.g. ME (*Ancrene Riwe*, Nero MS) *de ualse uikelare* but *wurst fikelare*, *one ureond* but *his freond*. Initial consonant alternations of this type are a prominent feature of the Celtic languages, and include nasalisation and lengthening (gemination) that result from assimilation with the consonants that originally ended the preceding word, as well as 'lenition' (cf. (f) above).

(l) Alternation of forms according to well-defined variations of clause-stress, e.g. MnE /haz, hæz, (ə)z/ 'has', /and, (ə)nd, (ə)n/ 'and' (cf. (ii) below, and also 5.5 for discussion of reduction in weakly stressed positions); and alternation that results from regular variation of word-stress, e.g. /rɛb(ə)l/ noun ~ /rɪ'bɛl/ verb 'rebel' (cf. 'Verner's Law', 3.5 below).

(ii) Simplification by loss of segments

Decrease in the length of words due to the loss of one or more segments is

¹ The term 'lenition' is used here to refer to the changes that took place in prehistoric Celtic before the loss of endings, and not to the surviving morphophonemic alternations.

often closely related to assimilation, e.g. the omission of /k/ in L. *pasc + tor* > *pastor*, *quinctus* > *quintus* leaves the purely dental-alveolar sequences /st, nt/; or loss may provide the conditions for a subsequent assimilation, as in OE *æmete* > ME *amte* > *ant*. The loss of vowels is from syllables with least stress in the word or clause. It is a succession of such losses of both vowels and consonants that is mainly to account for changes like L. *testimonium* > Fr. *témoïn*, L. *consobrinum* > Fr. *cousin*, IE */bherǝnǝm/ > E. *bear* (cf. Skt. *bharanam*). The principal types are:

(a) Loss of intervocalic consonant with contraction of vowels, as in Gmc. */sexwan/ > OE *sēon* 'see' (cf. Goth. *saihw-an*, OHG *sehan*), L. *vita* > OFr. *vidhe* > Fr. *vie*, *pavorem* > *peur*, *securum* > *sûr*, Mn Sc. /dil/ 'devil'.

(b) Loss of medial vowel (syncope), as in L. *homines* > Fr. *hommes*, *bonitatem* > *bonté*, *anima* > *âme*, *tabula* > *table*, *manducare* > *manger*, *calidum* > It. *caldo*, OE *hatte* (cf. Goth. *hâtada*), *heht* (cf. Goth. *hathait*), *hierde* (cf. Goth. *hausida*), *ænlipig* > *ælpig* > ME *alpi* 'single'; cf. MnE *business*, *evening*, *every*, *ordinary*, *Wednesday*, *Salisbury*.

(c) Simplification of consonant clusters, as in L. *gnatus* > *natus*, */dɪdksko/ > *disco*, *testa* > Fr. *tête*, *patrem* > *père*, *rubeum* > */robju/ > *rouge*, OE *cwipst* > *cwist* 'thou sayest', *betst* > *best*, *godspell* > *gospel*, *hlæfdige* > *lady*, *hwylc* > *which*, ME *swuch* > *such*, *laund* > *lawn*, MnE *know*, *gnaw*, *write*, *right*, *bought*, *castle*, *climb*, *lamb*, E. dial. /o:l/ 'old', /tʃaɪl/ 'child', Sc. dial. /fak/ 'fact'; and by combination with (b) above, L. *dubitare* > Fr. *douter*, *civitatem* > *cité*, OE *heafod* > ME *heued* (*hefd*) > *head*, *hlaford* > *lord*, *macode* > *made*, *nawiht* > ME *nazt*, *nozt*, *not*.

(d) Loss of initial and final unstressed vowels (apocope): Gmc. nom. sg. masc. */-az/ (IE */-ǝs/), preserved in Finnish loanword *kuningas* and Runic *stainar*, but reduced in Goth. *dags*, ON *dagr* and lost in OE *dæg*; L. *murus*, -um > It. Sp. *muro*, but OFr. *murs*, *mur*; ON *gnógr*, *gllkr* (cf. OE *genoh*, *gelic*), MHG *gelouben* > *glauben*, *beliben* > *bleiben*, E. dial. /gri:/ 'agree', /bli:v/ 'believe', /naf/ 'enough', /levn/ 'eleven'.

(iii) Addition of segments or glides

'Addition' of substance would not at first sight appear to involve economy of effort. Yet the occurrences (although of different types and origins) have one feature in common – that they assist in smoothness of transition between segments or words:

(a) The insertion of a glide-vowel between the two consonants of a cluster is parallel and alternative to assimilation or reduction of the consonants, as may be seen by comparing /kənɛɪ/ 'knee' in S. Scottish dialect with /tni/ (obsolete) and /tneɪv/ in E. Scottish. The insertion of the glide enables the inherited features of the main allophones of the consonants to be preserved

instead of being adapted or lost. Cf. Sc. dial. /wərəɪt/ 'write', E. dial. /fɪləm, vɪləm/ 'film', /kərəɪn/ 'corn'. The quality of the glide itself may, however, be determined by features elsewhere in the word, as in OE nom. sg. *burg* > *burug* but dat. sg. *byrg* > *byrig*, ONhb. *wyrihta* but *worohton*, *perh* > *perih* but ME *purh* > *puruh* (cf. (vi) (a) below).

(b) A similar process is to be seen in the conditioned diphthongs that arise in the sequence VC, e.g. the essential features of front [æ] and back [ɪ] in early OE */æld/ 'old' were preserved in Southern OE *eald*, but in Northern OE *ald* */æ/ was retracted to [ɑ] by assimilation to [ɪ]; cf. OE */xertæ/ > *heorte* (conditioned diphthong), and ME *hert(e)* > *heart* (assimilation of vowel). Other examples are OE */fex/ > *feoh*, ME *sah* > *sauʒ* 'saw', *ehte* > *eight*, E. dial. /fleiʃ/ 'flesh', /aɪʃ/ 'ash'.

(c) A plosive is often inserted into consonant clusters that demand complex transitions, notably those containing /m, n, l, r/: Gk. *ἀνῆρ* nom. but *ἀνδρός* gen., L. *em(p)tum*, *camera* > Fr. *chambre*, *simulare* > *sembler*, *pulverem* > *pol(d)re* > *poudre*, *tenerum* > *tendre*, OE *æmtig* > *empty*, *þunor* > *thunder*, ME *nymel* > *nimble*. In some cases this change is due to faulty coordination of the articulators: in clusters starting with [m, n], the plosives [b, d] are automatically produced if, in anticipation of the next consonant, the nasal passage is closed too soon. In other cases it may be due to misinterpretation by the hearer of the particular allophones used, e.g. a strongly trilled [r] after [l, n] may be interpreted as [dr].

(d) Linking in hiatus is in many cases automatic, e.g. rounded or high-front vowels and diphthongs in English are followed by the corresponding consonantal glides [j, w] before another vowel, as in [gəʊʊən] 'go on!', [hiʊɪz] 'he is', [krɑɪʊt] 'cry out'; and this explains the retention of the fuller forms for *to* and *the* in hiatus, e.g. [tʊʊəs] 'to us', [ðɪʊɑ:t] 'the art', compared with [tə, ðə] before consonants.

Typical hiatus-links in English, French and German are not, strictly speaking, 'added', since they survive from older forms, e.g. *a(n)*, /r/ in *sore arm* but not in *sore foot*, Fr. /n, t, z/ in *un homme*, *a-t-il*, *mes enfants*. But they may be regarded as additions when extended to contexts where they are not historical, as with 'intrusive' /r/ in /lɔ:rənɔ:də/ 'law and order', Bavarian dial. /vɪɪrɪ/ = *wie ich*, /n/ in ME *he hafden al* sg. 'he had all' (Lazamon's *Brut A*), Swiss dial. /vo:nɪ/ = *wo ich*.

(iv) Distance assimilation (vowel harmony)

This change comprises the extension of a distinctive feature (front or back quality of a vowel) to the vowels of other syllables in a word. Some doubt attaches to whether the feature spreads through intervening consonants (in which case the change is a combination of the assimilations in 1(c) and (d)

above) or whether the change is psychological, resulting from a simplification habit ('rule') adopted in the coordinating mechanism; probably either or both factors may apply in different languages.¹ In some languages the change is merely occasional, e.g. L. *ne hilum* > *nini*, OE *finger* but *hunger*, *dages* but *dages*; but in others, especially Turkish, it is a prominent structural feature. A related change in Germanic, umlaut (mutation), evidently has word-stress as a conditioning factor, since it takes place only in stressed vowels when followed by unstressed /r, j, v/ in the next syllable. This to some extent differentiates the Germanic mutations as a special type, since it must be partly because of the weakening of the unstressed syllables that their most prominent mark is anticipated and added to the preceding syllable, e.g. Gmc. *fo:it~fo:tiz/ eventually yields *foot-feet*, cf. *land~landu/>ON *land-land*.

(v) Dissimilation

Dissimilation of adjoining consonants is parallel to (ia, ib) and (iia) above in that it provides a simpler articulation, but nearer to (iia) in that it preserves, to a greater extent than assimilation, the distinctive features of the cluster as a whole. Thus in OE *ciet(i)þ > *ciet* '(he) chooses', one of the two fricatives is dissimilated to a plosive, and likewise in OE *þeoff* > *þeft*, *heþu* > *height* *geþp* > *stgh*.

Dissimilation of non-contiguous consonants results less from the articulation itself than from difficulty in coordinating it. In most languages simple reduplications constituting the whole word seem to cause no difficulty, and indeed are sometimes created by the minor change known as distance assimilation, e.g. OFr. *cercher* > *chercher*. But when the repeated consonants occur at arbitrary, irregular intervals among other consonants in the same word (as in tongue-twisters over the space of several words), one instance of the repeated consonant is often replaced by another that is similar in articulation, notably /l/ by /r/, as in L. *Marcus dies* > L. *marcolidi*, *pergrinus* > *pellagrinu*, *arborum* > Sp. *arbol*, Fr. dial. *albre*. Cf. L. **caululeus* > *caeruleus*, **stingulatis* > *stingularis*, **medie* > *meridie*, OE reduplicated pret. **lealt* > *leort* (Goth. *laltot*), Gk. *θθημι > τθημι, *φεφρυα > τφεφρυα, E. dial. *frail* = Std. E. *frail*.

(vi) Minor confusions

(a) Confusion in the ordering of segments (metathesis) is of two types.

¹ In descriptive analysis, such situations are simply treated as the extension, over a number of segments, of a distinctive feature (or prosody in the Firthian sense) like palatalisation or nasalisation. But since the ultimate origins of these features are single segments, e.g. /j, n/ (whether lost or preserved), it is preferable for historical purposes to retain a segmental treatment throughout.

Firstly, the order of /r/ and a vowel may be reversed, as in L. **turbulare* > Fr. *troubler*, OE *pridda* > *third*, OE *þurh* > *through*, Sc. dial. /modrən/ 'modern'; this may amount to no more than insertion of a glide-vowel and misinterpretation of the stress (OE *worhte* > *worhite* 'wrought', *þurh* > *þurh* > ME *þorou3* > *through*). Secondly, the order of two adjoining consonants may be reversed. In some cases this may be due to preference for a sequence that is commoner in the language, e.g. Gk. *ττρκω > ττρκω, L. **vespa* > *vespa* (and conversely, English dialects with /wɒps/ as their original form for 'wasp' may also show /krɪps/ 'crisp', /klaps/ 'clasp'); cf. also OE *ascian* > *asian*, *fixas* > *fixas*, at a period when /sk/ was losing in distribution to /ʃ/ (3.6). But forms like L. *ascia* 'axe', *viscum* 'mistletoe' (cf. Gk. ἰσός) cannot be explained thus; it is possible that, at some time in the history of Latin, /sk, ks/ were especially liable to confusion since, as a result of normal phonetic processes, they were coming to be combined in the same paradigm (*misceo*, *mixtum*).

(b) Conflation of repeated sequences (haplogy, as in L. **nutri-trix* > *nutrix*) is sometimes described as a short cut in which the first /t/ is mistaken for the second in articulation, but the process may be more complex. To judge from British habits of articulating words like *temporary*, *veterinary*, there exist both hesitation and slurring in attempts to pronounce both syllables; then, in order to avoid this, speakers settle for /tempəri, vetnəri/, and a convention gradually arises that the shortened forms will serve. The change is to some extent an alternative to dissimilation, since it eliminates repetition of a segment.

The above short survey under headings (i) to (vi) could not include all the subtypes and variations of conditioned change to be found in different languages (some further examples will occur in later chapters); but it is intended as a representative selection from the main types. In a majority of these, ease of articulation is prominent as a prime source of variation, while the remainder appear to have a close connection with attempts to ease difficulties in coordinating the articulation of complex forms. These factors, combined with those of faulty transmission and/or interpretation, are sufficient to account for the origin of conditioned phonetic variants, and it can hardly be claimed that this type of change constitutes a problem in the history of language. With the obvious exception of straightforward addition to the lexis for the naming of new objects, it is the most easily explained type of linguistic change; in view of the nature of language and its transmission, it is wholly to be expected, and, indeed, its absence would be more difficult to explain than its presence.

An objection frequently raised¹ to the above explanation is that, if such changes are so simple, they should appear in all languages and dialects to a full and equal extent, and not in varying proportions in each. The answer to

¹ Most recently by King p. 189.

this will be given more fully in the remainder of this chapter and in chapters 5 and 6, but a few general points may be anticipated here. Firstly, there must be a limit to the rate and quantity of assimilations, contractions and other rearrangements that is tolerable in a single dialect if intelligibility is to be maintained, so that, even if a large proportion of them are continually occurring in certain colloquial styles, only a small proportion could be in process of acceptance into other styles at any given time. Secondly, the fact that their incidence differs according to language and dialect follows from the very nature of dialects, which owe their existence to development in part-isolation (6.2): in each dialect, there are differences in the basis of articulation (2.4), in the effects of suprasegmental features on the spoken chain (2.5), in the forms that are socially accepted or rejected as status-markers (6.5), or selected for functional utility (5.2), or merely redistributed at random (2.7, 6.7). Since all these factors have relevance for the subject as a whole and not merely for conditioned phonetic change, they are dealt with in turn in later chapters. For the present, we may conclude that conditioned phonetic change, though it covers a wide sphere and has many repercussions on other spheres, is not difficult to explain either as regards basic causes or varying incidence. It has been dealt with first so that it can, as it were, be taken for granted and not allowed to complicate discussion of the more difficult problems that remain.

2.4 Isolative phonetic change: introductory

The main problem of phonetic change arises from 'spontaneous' (isolative) changes in which a phoneme changes in all, or nearly all, of its occurrences irrespective of environment, as in OE *hām, stān, hlāf* > *home, stone, loaf*. The neogrammarians of the last century tabulated such changes in great detail, but they were mainly concerned to observe correspondences between different stages of a language, and to formulate 'laws' for the changes. The phonetic grounds for them were never specified beyond what H. Paul summarised as 'displacement of the motory sensation'¹. Individual changes might be attributed to 'raising of the tongue', 'lip-rounding', etc., and a battery of metaphorical terms like *Aufhellung* and *Brechung* was evolved which helped to hypostatise them, for a number of generations, as so well based as to need no further explanation. Later, we find terms like 'trend' or 'drift' of a language, which no more explain the changes than did 'the genius of the language' in the eighteenth century. It is small wonder, then, that modern scholars have turned from a purely mechanical approach and explored other possibilities – the requirements of the language system, the influence of substrata or of neighbouring languages, or the social origins of change of fashion. Before

¹ Paul ch. II.

proceeding to these, we must assess how far purely phonetic factors can explain the origins of 'spontaneous' sound-change.

Firstly, the boundary between those failures in performance that are purely accidental and those that are not is by no means clear-cut. The number of speakers with genuine speech defects like cleft palate or hare lip is comparatively small, but they engage in as much speech activity as others, and their speech reinforces tendencies already existing in other speakers, such as the substitution of a glottal plosive for other plosives, or of a voiceless lateral fricative for [s]. If such influences cannot be regarded as more than marginal, others are stronger. For example, the extraction of all teeth and the wearing of full sets of dentures (often of poor fit) from an early age were at one time common practice in certain sections of the Glasgow population: in these sections, even in speakers still possessing all their own teeth, the distinction between /s/ and /ʃ/ is less clear than elsewhere, and it is probably only preserved at all under influence of the standard language.¹ This is a special case because there is direct interference with the speech mechanism, but the influence of paralinguistic features, such as the voice qualities typical of different regions, is much wider. It has been suggested that the prevalence of adenoids in Liverpool speakers has been responsible not only for the local voice quality but for changes in the phonetic system of the dialect.² It is probably reasonable to assume that inherited (i.e. environmentally transmitted or 'learnt') features of voice quality are the indirect causes of linguistic change in many regions; and although it has yet to be shown that they could affect whole languages, they cannot be ignored in dialect research.

With voice quality must be considered another phonetic feature that is environmentally inherited and is still more likely to affect the development of a language. It is what has been variously termed its 'basis of articulation' or 'articulatory setting', and refers to the configuration of the movable speech organs typical of a given language. It was once thought³ that this was purely a 'rest-position' (*G. Ruhelage*) which was connected in some vague way with national or regional habits of both physique and temperament; but it has now been shown⁴ to be a special setting that is adopted whenever speech is in progress, differing from a rest position and including numerous adjustments of muscular tensions throughout the speech organs. According to this latter view, the setting is that which is ideal for the pronunciation of all the sounds in a given language, especially its commonest sounds; but this assumption would be extremely difficult to prove, for, when one considers how great is

¹ Confusion of /s/ and /ʃ/ without any special phonetic abnormality might be attributed to contact with Gaelic, but in the Glasgow area the sound in question is much nearer to the voiceless lateral fricative typical of a speech disorder.

² Abercrombie (1967) 94–5.

³ Sievers (1894) 55–7, 272; Heffner pp. 98–9.

⁴ B. Honikman, 'Articulatory Settings', in Abercrombie (1964) 73–84.

the range of allophonic choice available in any language, it is hard to see why the setting should be determined by the particular allophones in use. One might equally well assume that it was determined by the sounds used at various times in the past history of the language, and that it now determines, and is not determined by, the allophones actually selected.

Whatever the precise origins of articulatory settings, they vary greatly from language to language. The position of the tongue may vary from high and front to low and retracted, the lips may be rounded, neutral or spread, and the jaws may vary in degree of aperture (the nature of the variation may be seen most clearly by comparing the hesitation-phenomena of various languages, e.g. English /əm/ or /ɜ/, Scots /e:(ə)/, Norwegian /ø/, Russian /ɨ/). It follows that articulatory movements must vary in each language, not only for the obvious fact that the phones themselves differ, but because the starting and finishing points of the articulators are different. It is likely, therefore, that when changes take place, there will be an inbuilt preference for fronting in some languages, for raising in others, and in yet others for rounding or retraction. The subject has not been extensively studied, but the possibility remains that the direction of sound-change may be determined by a feature which, though perhaps partly functional in origin, is environmentally transmitted in each language, i.e. it is automatically learnt by imitation in childhood.

The factors just outlined have relevance for all phonetic change, whether isolative or conditioned. But a certain minority of isolative changes are similar to conditioned change in that they require little explanation beyond that of economy of effort. They are those that involve (a) the loss of a distinctive feature from a phoneme in all, or most, of its contexts, e.g. the loss of fricativity in the change of prehistoric Gk. /s/ to /h/, or of Gmc. (especially initial) /x/ to /h/; or (b) ultimate loss of all distinctive features resulting in complete loss of a phoneme, e.g. of IE /p/ in Celtic, IE /w/ in Gk.; or (c) the simplification of a complex articulation, e.g. /ā/ > /a/, Gmc. /ai/ > OE /a:/, Gmc. /xʷ/ > OE /hw/ > MnE /w/, sometimes by prior assimilation of one element to another, e.g. /ai/ > /e:/, /au/ > /ɔ:/, o:/, both of which are common changes, as may be seen from later Latin, Old Saxon, and Early Modern English. (These latter monophthongisations may have close systemic connection with the opposite process of 'spontaneous' diphthongisation (cf. 2.5, 3.3), but the origin of variation is different in each case.)

The distribution of such simplifications is not always complete, and to that extent they resemble conditioned changes; but since there is no obvious conditioning factor, it is methodologically preferable to class them with the residue (i.e. with all the isolative changes requiring explanation). The history of some of them (e.g. /hw/ > /w/, 6.7; /x/ > /h/, 8.3) shows that the feature was lost first in some contexts only, and its loss was then gradually generalised. It

suggests that isolative change may have the same origins as conditioned change, but that, for reasons that are not immediately obvious, it is extended to include all contexts. We shall return to this possibility at the end of this chapter.

2.5 Isolative change: principal mechanism

The majority of variants of segmental phonemes, in so far as they are not conditioned by contiguous segments (2.3) or by well-defined alternations of suprasegmentals (p. 13), arise under the influence of those suprasegmental features whose incidence is not constant but varies according to both context and situation: stress and intensity, pitch and intonation, tempo, quantity and rhythmic pattern.¹ For our present purpose we need not define whether these belong to one or more of paralinguistic, phonetics or phonology. The point is that, whereas in 2.2 above only *one* extreme in the range of register-variations was adduced to account for 'careless' (e.g. unstressed, relaxed or slurred) variants, here, in looking for the sources of isolative changes, *both* extremes must be considered, i.e. we must also include variants arising in 'careful', 'forceful' or 'committed' styles.

To consider vowels first, there is evidence to suggest that the stressed variants found in more forceful styles may show a *higher* or *more fronted* tongue-position, whereas the less stressed variants of relaxed styles may show *lower*, *more centralised* or *retracted* tongue-positions. The evidence for this claim is circumstantial but can hardly be ignored. Firstly, such a distribution of variants has been often observed, e.g. in the London dialect:

I'm 'going down the 'road: [daʊn].
Get 'down! : [dɔʊn].

A similar distribution of [ɛ] and [iɛ] has been observed in Sicilian dialects.² If it is objected that such variants could have arisen in a quite different way, and have been merely selected to carry the distinction, there is a long-standing corpus of experimental evidence which shows a close correlation of the increased muscular energies required for (a) greater stress, and (b) vowels with higher and more fronted tongue-positions.³

The reverse process – the use of a lowered variant as a result of decreased

¹ Included here are all prosodic and paralinguistic features except the few already specified as 'well-defined alternations' on p. 13 above. Such exceptions do not differ in principle from the rest; they happen merely to yield such pronounced and much-used contrasts as to condition doublets from originally single segmental sequences.

² Schmitt pp. 68ff.

³ Dieth p. 83, with references there quoted; Ladefoged ch. 1. The experimental evidence does not of itself prove the point in question, but it is significant when taken together with the evidence of stress-variants.

muscular energy – may be deduced from a comparison of the diaphones of certain English diphthongs. The vowel of *made* or *say* is today realised variously as [eɪ] (RP), [eː] (Scots), [ɛː; ɛə] (Yorks.), [ɛɪ] (London and some types of RP), [æɪ, aɪ] (Cockney). The evidence of seventeenth- and eighteenth-century writers on pronunciation indicates that the main Southern reflexes descend from a sound no lower than mid-front [ɛː]. As the accompanying sonagrams¹ show, the present diphthong [ɛɪ] does, *at one point of its span*, achieve the same acoustic effect as the monophthong [eː]. The origin of [ɛɪ, æɪ, aɪ] is therefore likely to have been relaxed style, in which the speaker began at a lower tongue-height than that needed for [eː], but then compensated by raising the tongue to a height greater than that required initially. An analogy, which both illustrates the process and suggests that relaxed style is the origin, is to be found in a certain type of languid crooning (or incompetent singing, as professionals might regard it) in which the performer never hits the higher notes at the point of expectation, but starts them lower and climbs up to them *glissando*. Other diphthongs of this origin are to be seen in Cockney [səɪ] ‘see’ and [fəʊd] ‘food’: the original vowels [iː, uː] were begun from lower and more central positions (at first [ɪ, ʊ], then [ə]) and their quality was reached only towards the end of the articulation. The same could be said of Cockney [gəʊ] ‘go’ and probably of RP [gəʊ], which both show lowering and centralisation of the first element of [ou].

Such cases suggest that if the whole range of suprasegmental features is taken into account, the initial occurrence (if no more) of most new vocalic and diphthongal variants could be explained. The combinations in which such features can occur are many and complex; indeed, stress alone implies a combination of muscular energy which can manifest itself at various points on any or all of the scales of quantity, loudness and softness, tenseness and laxness, high and low pitch, and rising and falling patterns of intonation. In the past, certain phoneticians² have suggested that it is *varying* combinations of such features that produce different types of vocalic and diphthongal change. These suggestions have been largely ignored in recent decades, as is shown by an authoritative statement like Bloomfield’s ‘no permanent factor . . . can account for specific changes which occur at one time and place and not at another’.³ But the validity of such a negative approach depends on how much is being asked of the ‘permanent factors’. The influence of suprasegmental features probably does account for a majority of variants that exist in the spoken chain, but not for the selection of one variant or another at a given point in the history of a dialect, and its generalisation to all contexts of the phoneme. It was thus the claim that this factor provided the whole cause for

¹ My thanks are due to Dr A. Classe for his kindness in supplying me with these.

² E.g. Fouché (1927), Schmitt (1931).

³ Bloomfield p. 386.

‘He made haste to say’

(i) Scots

(ii) English



change that gave rise to suspicion; if it is viewed as the mere origin of variation, there are no grounds for rejecting it. Where subsequent developments confirm the possibility, we do not deny it: the distinction between *holy* and *holiday* shows what was at one time a complex distinction between configurations of stress, rhythm, tempo and quantity which happens to have been preserved and emphasised for us by subsequent divergent development. In particular, the effect of suprasegmental features on the phonetic shape of diphthongs was established long ago by Duraffour for the Provençal dialects;¹ it is borne out by a pair like [gúəd] 'good' but [gwódlək] in Frisian dialects;² and it is possible to find similar fluctuations in some types of English, e.g.

(a) *I can't 'pay* (deliberate statement with falling intonation): [pe¹].

(b) *You can't 'pay?* (surprised or outraged question with rising intonation): [pei].

It may be objected that the variety of isolative changes is so great that it is impracticable to attempt an account of their origin from suprasegmental influences; and, in view of our lack of detailed evidence on the subject, the objection has some cogency. However, it is possible, partly by the reverse process of deducing what kinds of suprasegmental a given variant would best fit, to arrive at a tentative, hypothetical scheme, the object of which is to show merely how different combinations of suprasegmentals might be expected to produce the commonest types of variants to be found both in past records and in modern dialects.

A. Long vowels and diphthongs

1. Forceful style, greater stress, with

(a) level or rising intonation: raising, e.g. [e:] > [i:], [au] > [æu]; fronting, e.g. [ɑ:] > [æ:], [u:] > [y:], [ɑu] > [øy].

(b) falling intonation: centring diphthongisation, with the first element higher than the original vowel, e.g. [e:] > [iə], [o:] > [uə].

2. Relaxed style, less stress, with

(a) level or rising intonation: [i:] > [əi], [u:] > [əu].

(b) falling intonation: centring diphthongisation without raising of the first element, e.g. [ɛ:] > [ɛə].

B. Short vowels

These are less capable of showing the effects of one of the elements of increased

¹ Duraffour (1932), especially p. 66.

² Frings pp. 25-6 and references there quoted; J. M. N. Kapteyn in W. Steller (ed.), *Festschrift T. Siebs* (Breslau, 1933) 152ff; and for parallels from Sicilian and Russian, Schmitt pp. 68ff.

stress – muscular tension. Increased stress may be manifested by lengthening of the syllable, either by lengthening of the vowel itself,¹ or by doubling of the following consonant² (cf. 3.10). Under conditions of less stress, lowering and centralisation may occur, e.g. [ɪ] > [ə], [ʊ] > [ə, ʌ] (but cf. 3.9).

The above scheme includes only the main suprasegmental features. It does not include a distinction like *clipped–drawled*,³ though these features are probably responsible for further variants. Certain types of RP that are characterised by drawling show diphthongised forms like [bæ^əd] ‘bad’, and it can hardly be a coincidence that diphthongisation of short vowels has gone still further in those United States dialects that are described as having a ‘Southern drawl’. At the other extreme, ‘clipped’ pronunciations may account for shortened variants, and these may arise in great numbers when there is a change in the rhythm of the language, e.g. from stress-timed to syllable-timed. In the latter type of rhythm there is a tendency to level the quantity of each syllable, and this may account for the stress-shift in diphthongs in later Old and Middle French, e.g. /roi/ > /rwe/. A similar change, though less radical, presumably underlies the change of early ON [bæ^ərya] > *bjarga*, [hæ^əlpa] > *hjalpa* (later *hjdþpa*); and a parallel has occasionally occurred in English when a short variant of a diphthongised vowel has come to stand in unstressed position, as in /wan/ ‘one’, Sc. /e:n/ > /jɪn/.

As regards variations of consonants, the directions likely to be taken in isolative change are (apart from the simplifications mentioned in 2.4) more limited than for vowels, since their articulations depend on discrete positions like dental or velar, not on a continuum of tongue-positions as in the case of vowels (though, within groups of closely related consonants, it is not hard to see why there can be interchange between, say, [ɣ] and [R], or between [v], [w] and [β]). It is possible, too, that the effects of suprasegmental variation are less on consonants, which occur at syllable boundaries, than on vowels, which occur at the peak of the syllable. But even so, the effects of the forceful/relaxed distinction must still be reckoned with, especially on the coordination of articulatory movements required for plosives. In relaxed articulation, if the active articulator of the plosion (lip, tongue or velum) is not withdrawn soon enough for a clear-cut plosion, the result is affrication: since the articulator is still in its original position after plosion has started, friction intervenes between plosion and the vowel that follows, so that alveolar [t] is realised as [tʃ], dental [t] as [tʰ], [p] as [pʰ] and [k] as [kʰ]. Conversely, in certain types of forceful articulation, the onset of voice after plosion may be delayed *without* fricative interference, so that [p, t, k] are realised as aspirated [p^h, t^h, k^h].

¹ See especially Laziczius pp. 38–63.

² Cf. B. E. Newton, ‘Spontaneous Gemination in Cypriot Greek’, *Lingua* 20 (1968) 15–57 and references there quoted.

³ For the great variety of contexts in which these and other prosodic features are found in Modern English, see Crystal pp. 300–8.

One of the best known of irreversible consonant-changes is that of voiceless plosives to fricatives or affricates, as in IE [p, t, k] > Gmc. [f, θ, x], Ancient Gk. [ph, th, kh] > Mn Gk. [f, θ, x], OHG [p, t, k] > [pf, ts, kx]. In all these, without postulating affrication or aspiration as the causes, we may at least regard them as the sources of variation.¹

2.6 A problem: preliminary statement

One of the problems that lie at the root of this enquiry may now be stated in preliminary fashion. If an isolative change such as the raising of vowels is due – even only initially – to the influence of suprasegmental features, why is it that this conditioning factor does not apply equally to *all* vowels, and what relation has it to mergers and circular shifts? In the case of the latter, it may be objected that variants of *opposite* origins have to be assumed for different parts of the shift, e.g. for the Great Vowel Shift, forceful-style variants for the changes /e:/ > /i:/, /a:/ > /e:/, /ɛ:/ > /e:/, /o:/ > /u:/ and /ɔ:/ > /o:/, but relaxed-style variants for the diphthongisations /i:/ > /aɪ/, /u:/ > /aʊ/. This apparent contradiction is, however, only superficial. All linguistic change entails a process of selection (7.2), and all the variants are of mechanical origin and are continually and automatically present in the spoken chain. Moreover, the combination of variants that originally belonged to different registers (or dialects) into a single register is, as we shall see, the basis of grammatical restructuring, and may therefore reasonably be assumed for phonology also.

The more basic problem that remains is: why do forceful-style variants preponderate at different times and places in linguistic history, and why are redistributions of the type just mentioned a necessary consequence of such imbalances? Answers to this problem will be attempted in 3.7 and in chapter 6.

2.7 Conditioned and isolative change

We have seen that the variants contributing to both conditioned and isolative change can be termed broadly ‘stylistic’ in origin. In conditioned change, those typical of relaxed style are uppermost, i.e. the effects of ease in articulation are prominent; in isolative change, the variants may be those typical of *either* relaxed *or* forceful style, i.e. both ease and greater intensity of articulation have to be taken into account. Furthermore, there is a close relation in origin between the two types. In conditioned change, the variants arise (i) from

¹ The Celtic lenitions (2.3) are sometimes classed with these changes on the ground that they include the parallel result [p, t, k] > [f, θ, x]; but they always involve a split which was conditioned by the alternation of forceful and relaxed articulations according to phonetic context, and not by stylistically induced suprasegmentals.

assimilation or other adaptation of contiguous segments, (ii) from causing the coordination of a whole succession of segments, and (iii) more rarely, from assimilation of segments to certain well-defined alternations of suprasegmentals, such as those of word-stress in 'Verner's Law' (3.5). In isolative change, too, the segments in question are assimilated or adapted, but in this case to suprasegmentals that are less predictable in distribution: those that may accompany different registers, prosodies, paralinguistic features and the like (2.4-5).

But if the two types are as closely related as this, why is there such a difference in their distribution? Why do conditioned variants remain limited in certain contexts, whereas isolative change, although similarly limited in origin, is extended to all contexts of the phoneme? A clue is provided by the difference in the types and distributions of assimilation that have just been referred to. In conditioned change, various allophones can remain widely dispersed because the conditioning is overt (i.e. the distribution can be initiated and learnt); but in isolative change the conditioning is covert – the distribution in which the whole range of register-variants can occur is not obvious. If the conditioning factor can be perceived and learnt, the variants can survive, irrespective of whether the distributions remain purely allophonic or are in time reallocated to different phonemes (3.5-6). But the harder the conditioning factor is to perceive, the less chance there is of this, and the result will be that, of two covertly conditioned variants, one is generalised at the expense of the other: if the newer form is preferred, isolative change takes place; if the older form is preferred, there is no change at all. (The notion that a child learns the overt distinctions by 'rule', but that there is no rule that could be formulated for covert and more fluid distributions, would no doubt be another way of putting the matter; but it is by no means a *sine qua non* for the description of the overall historical process.)

The above explanation is aimed at the two extremes – change that is obviously conditioned, and change which, to judge from evidence available, affects a majority of the instances of a given phoneme more or less simultaneously. But there are other types that may be regarded as midway between these two extremes. As already suggested above (p. 20f), simplifactory changes are more likely to have taken place gradually, context by context (see further pp. 120-3 and 8.3). A similar gradual spread may apply to any conditioned change if the conditions are successively widened, or simply because a new generation misunderstands the previous distribution and extends it to contexts where no conditioning applies¹ (cf. p. 124f below). Furthermore, the rise of suprasegmentally conditioned doublets is well attested, and preferences for one

¹ Especially in cases of isolative change which show an unusual direction, it is necessary to explore the possibility that they might conceal a long process of spread from an originally limited conditioning to other relevant parts of the lexis. Cf. the term

or other of the two forms may take time to become established, affecting different parts of the lexis in different ways. For example, the present distribution of the types *food*, *mood* and *good* reflects changes of quantity at different periods, and results from selections – whether arbitrary or motivated – that were independent of the original conditioning factors (cf. pp. 35 and 125).

To sum up thus far: variation of mechanical origin in the spoken chain provides a sufficient explanation for conditioned phonetic change; it will also explain the origins of isolative change, but the question of its relative importance for the process and incidence of isolative change as a whole is left open for the present, and will be resumed in chapters 3, 6 and 7.

¹lexical diffusion' proposed by Wang pp. 12-18.

Other unexpected changes may be open to similar explanations. For example, the change of initial *gr-* to *r-* in South Italian dialects (e.g. *ramme* = *grande*) is unlikely to be a straightforward simplification since that would imply an unusually low standard of what is 'hard to pronounce'; more probably, therefore, it is an extension from another change normal in the same dialects, in which *gr-* > [ʎ].