Part A

The Speech Community

1

The Darwinian Paradox

It is only reasonable that a book about the principles of linguistic change should begin by asking how important these principles might prove to be – should we succeed in defining them – for the general understanding of language. For a sizeable part of the current linguistic enterprise, there can be very little interest in principles of this kind. The search for a universal, unchanging, and indeed unchangeable, grammar is oriented in an entirely different direction. As such, it lies outside the scope of this work, which is concerned with everything in language that changes or has changed. This seems to include much the larger part of linguistic categories, structures, and substance. It therefore seems natural to ask whether we understand the forces that are responsible for the extraordinary transformations that affect all but a bare skeleton of abstract relations.

The first volume of this work began by admitting that we do not. A long series of inquiries has left us with much detailed knowledge of constraints on and patterns of change, but no general explanation for the scope and persistence of this phenomenon. The continued renewal and far-ranging character of linguistic change is not consistent with our fundamental conception of language as an instrument of social communication. The situation is compounded by the unexpected finding that has emerged from recent sociolinguistic research: that linguistic change is continuing at a rapid rate in every city of North America that has been studied with any care.[1] This result clashes sharply with the common-sense expectation that constant exposure to the network standard on radio and television would lead to convergence and the gradual elimination of local dialects. Language change governs not only our history, but also our immediate present. The immanence of language change makes it easier for us to study, but it also heightens the urgency of the search for explanations. Much of the present volume is devoted to that search; but before we begin it may be helpful to look briefly at the effects of change.

[1] I limit this observation to North America, since it seems to hold for all of the English speaking cities of the USA and for both English and French speaking cities in Canada. Active changes from below have been traced in studies of the cities of South America, the United Kingdom, New Zealand, Egypt, Japan, and Korea, but in Europe and many other areas, community studies have placed more emphasis on changes from above, dialect leveling, and koine formation.

1.1 The social effects of language change

It may be useful to remind every reader of this work that all of us have suffered from the effects of language change in one way or another. These effects range from petty inconveniences to crushing disabilities that can consume years of our lives with unrewarding struggle against hopeless odds. First, we might notice the small family arguments that ebb and flow over the proper use of words. My generation called an ice box an *ice box*, since it used to hold a block of ice before it was electrified, but my children's generation insisted on calling it a *refrigerator*, and confined the use of "ice box" to the freezing compartment that makes ice cubes. My age group will also be ridiculed, perhaps more slyly, for calling something "swell" or "nifty," terms that are now marked as hopelessly old-fashioned unless they are used ironically.

On the other hand, many older citizens find themselves keenly irritated by new forms that have crept into the language, and expend a great deal of effort in demonstrating to their children the illogical character of *hopefully*, *aren't I*, or *like* as a conjunction. But even the most eloquent journalists and educators find that their rhetorical tools are not keen enough to cut the link that ties these forms to the younger speakers of the language. These defective forms return again and again until they are firmly fixed in the fabric of the language – when suddenly they appear as very natural and not at all defective, to all except a small group of traditionalists in professorial and editorial chairs.

The fact that traditionalists are usually fighting a losing battle is not enough to deter them from penalizing the students who take the winning side. Most of us have suffered by having our school papers downgraded by quixotic supporters of a dying tradition, who insist that infinitives cannot be split from their *to*s, that *data* require a plural verb, and that under certain subtle conditions we must spell and even pronounce the word *who* as *whom*.

The emotions aroused by grammatical forms are if anything more temperate than the feelings called up by sound changes, when they finally come to our attention. Middle class American parents in particular feel continually called on to correct the aberrant vowels used by their children, which seem to symbolize their association with the most vulgar elements of local society. These family disagreements over sounds rarely rise to the level of public disputes, because there is no vocabulary available to institutionalize them, but the dispute goes on at the local level with unremitting intensity.

When we observe such controversies in a foreign language, it is easy enough to see them as tempests in linguistic teapots. But in our own language, it is difficult to avoid being caught up in the storm of emotions generated by the contrast between newer and older ways of saying the same thing. It is not easy to step back far enough to ask the fundamental question: why does language change arouse such violent feelings?

Even as we are irritated and confused by linguistic change in progress around us, we suffer more serious disabilities from the results of changes that took place centuries ago. An enormous amount of time and effort is devoted to mastering English spellings such as *bight*, *drought*, *about*, *draft*, *draught*, *cough*, *trough*, and *enough*. The distinctions between *whale* and *wail*, *mourning* and *morning*, *colonel* and *kernel*, must now be mastered by brute memory, though they were quite transparent to earlier generations who had not suffered the effects of the sound changes that collapsed these categories. These spelling-demons are typical of a great many forms that were once a rational representation of the spoken language, but are now the fossilized evidence of language changes that are no longer part of the knowledge of the native speaker.[2] But even greater amounts of time are devoted to learning German, French, Spanish, or Russian, languages that were once mutually intelligible dialects of Proto-Indo-European. And even when a great investment in language learning is made, it may not be enough to overcome the elusive gap between the two language structures that has emerged as the result of language changes over centuries. We may find that no amount of practice allows us to master the native Russian production of palatalized and non-palatalized consonants, the native French rules for schwa deletion, or the use of aspect particular to either of these languages. Worse yet, most of us find that we are not very intelligent in a language other than our mother tongue. These are some of the disadvantages that follow from the fact of language change. What can we point to on the positive side? Some people say that they like to study foreign languages, and some invent secret languages to make it harder for others to understand them.[3] Linguists and language teachers get some employment from the results of linguistic divergence. But that seems to be about all the benefit there is to language change. It is hard to avoid the conclusion that language, as an instrument of communication, would work best if it did not change at all. Though we get some satisfaction from playing with language, and often find it useful to hide behind it, we do not profit in any obvious way from the results of systematic language change.[4]

[2] One might note here the argument advanced by Chomsky and Halle that English orthography can hardly be improved (Chomsky 1964, Chomsky and Halle 1968), based on the demonstration of cases where derivational alternations will support current spelling as a good representation of the underlying form. But the vast number of irregularities in English spelling are not supported by any alternations, but are the result of mergers that eliminate any basis for reconstructing the original forms. In general, morphophonemic alternations are irrelevant to the effects of mergers (see volume 1, chapter 13).

[3] One might point to the widespread use of playful secret languages and other word games as evidence for positive values associated with language learning.

[4] These informal remarks on the consequences of linguistic change will be amplified considerably by the observations and experiments reported in volume 3, which deals with the cognitive consequences of change.

It is not then so hard to understand why the general view of language change is a dismal one. We all seem to be suffering from a linguistic disease that has no cure, and language, like so much of the world around us, is seen as going from bad to worse. Though this "Golden Age" principle is quite general, it assumes an absolute form only in one area: the domain of language. In the course of studies of the speech community that began in the 1960s, I and my associates have interviewed many thousands of speakers in many English dialects and other languages. Whenever language becomes the overt topic of conversation, we find a uniformly negative reaction toward any changes in the sounds or the grammar that have come to conscious awareness. Communities differ in the extent to which they stigmatize the newer forms of language, but I have never yet met anyone who greeted them with applause. Some older citizens welcome the new music and dances, the new electronic devices and computers. But no one has ever been heard to say, "It's wonderful the way young people talk today. It's so much better than the way we talked when I was a kid."

1.2 The parallels between biological and linguistic evolution

The Golden Age principle does not necessarily apply to additions to vocabulary, or the borrowing of prestige features from another system. In all the discussion of language change to follow, I will be focusing on alterations in the mechanism of the language, its system of sounds and grammatical categories: the fundamental process that has led to the mutual unintelligibility of related dialects and languages over many centuries. Our view of this linguistic evolution is of course limited, and confined to those language families whose development has been reliably traced.[5] Indeed, most of the discussion of the causes of linguistic change is cast within the framework of the development of Indo-European. The reconstruction of this vast family tree offered to the scientific community a remarkable parallel to biological evolution that was widely observed and remarked upon. It is often said that the evolutionary transformation of species was demonstrated in linguistics before it had been clearly spelled out in botany and zoology (Lyell 1873:406; Christy 1983, ch. 1). Max Müller wrote: "in language, I was a Darwinian before Darwin" (1861). Though we might trace the origin of the evolutionary viewpoint in the works of many scholars, the problem set for this volume is best seen in Darwin's summary of the situation. *The Descent of Man* (1871) contains a very specific treatment of the parallels between linguistic and biological evolution as he saw them.

[5] For an evaluation of efforts to trace language families back further than this, see Ringe 1992.

(1) The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously parallel.

Darwin finds fifteen similarities between the two processes, which he does not elaborate. The list below gives his phraseology in italics, and supplies current examples from the changes in progress that will be traced below.

1 *We find in distinct languages striking homologies due to community of descent*
The Northern Cities Shift (volume 1: pp. 177–201) emerges in the same form in the cities of Rochester, Buffalo, Syracuse, Cleveland, Gary, Detroit, and Chicago. All these cities were settled by the same westward movement from New York State.
2 *and analogies due to a similar process of formation*
The Pattern 3 chain shift developed independently in roughly the same form in Hauteville French, Sweden, Greece, and São Miguel in the Cape Verde Islands, without direct contact among them (Martinet 1955).
3 *The manner in which certain letters or sounds change when others change is very like correlated growth*
Parallel movements of front and back vowels appear in the Great Vowel Shift and in the raising of (æh) and (oh) in New York City, while parallel fronting of back vowels appears in the Southern Shift.
4 *We have in both cases the reduplication of parts*
The bilateral symmetry of phonological and grammatical systems is as marked as the bilateral symmetry of most organisms.
5 *The effects of long continued use*
The extreme reduction of high frequency function words often reaches the point where a morpheme is represented by a single phonetic feature, as in the reduction of the past tense morpheme *wen* in Hawaiian Creole English to a feature of length in [he wːɔk bai ðe we] 'He walked by the way' (Labov 1992).
6 *The frequent presence of rudiments, both in languages and in species, is still more remarkable*
Darwin gives the example of the elimination of schwa in the contraction of *I am* to *I'm*, removing "a superfluous and useless rudiment," as well as the retention of letters in spelling (where we might give the example of the *k* in *knee* and the *g* in *gnome*, which support no alternations).
7 *Languages, like organic beings, can be classified in groups under groups*
This is as true of the English dialects as of the Indo-European languages themselves. Thus the Boston dialect is plainly a member of the Eastern New England sub-group of the Northern dialect region.

8 *They can be classified either naturally, according to descent, or artificially by other characters*
While the Northern dialect area unites dialects with a common history of descent, the classification of the Southern Shift unites Southern England and the Southern United States, with no clear motivation in the history of settlement.

9 *Dominant languages and dialects spread widely*
The influence of dominant cultural centers on American dialects can be seen for many phonological features as a large (roughly) circular region surrounding Boston, Philadelphia, Richmond, and Savannah.

10 *and lead to the gradual extinction of other tongues*
Though many rural American dialects are stable or expanding, it is widely and reliably reported that others are in danger of disappearing (Schilling-Estes and Wolfram 1999) in the same manner as local French and German dialects (Hinskens 1992).

11 *A language, like a species, when once extinct, never . . . reappears*
Though this statement has been questioned in the case of Israeli Hebrew, it is widely accepted by linguists.

12 *The same language never has two birthplaces*
This certainly seems to be true for all of the English dialects we are studying, since a dialect is too specific and complex a configuration to arise independently in several places.

13 *Distinct languages may be crossed or blended together*
This is clearly the case for creole languages, which arise in one place from a mixture of language contributions. Most American English dialects are the result of regional koine formation from many intersecting English dialects.

14 *We see variability in every tongue, and new words are continually cropping up*
Variability is of course the main topic of our investigation. The renewal of regional vocabulary in the United States can be clearly documented in many semantic domains.

15 *Single words, like whole languages, gradually become extinct*
The rural vocabulary that was the principal defining characteristic of the regional dialects of the Atlantic states has to a large extent disappeared, as shown by the obsolescence of *singletree*, *stone boat*, and *darning needle*. This parallel is seen most clearly in the domains of slang and colloquial vocabulary.

Darwin uses the two last parallels to introduce the argument necessary to establish the similarity of biological and linguistic evolution: that linguistic evolution shows the same kind of natural selection that biological evolution does.

16 *The survival or preservation of certain favoured words in the struggle for existence is natural selection.*

Darwin then supports this view with a quotation from Max Müller:

(2) A struggle for life is constantly going on amongst the words and grammatical forms in each language. The better, the shorter, the easier forms are constantly gaining the upper hand, and they owe their success to their own inherent virtue.

The general consensus of 20th-century linguists gives no support to this contention, and finds no evidence for natural selection or progress in linguistic evolution. It is generally agreed that languages that have evolved in societies with subsistence economies based on hunting, gathering, or small-scale agriculture show a structural complexity that is equal to or greater than those spoken in technologically developed societies.[6] Darwin himself quotes Schlegel to this effect:[7]

(3) In those languages which appear to be at the lowest grade of intellectual culture, we frequently observe a very high and elaborate degree of art in their grammatical structure. This is especially the case with the Basque and the Lapponian, and many of the American languages. (cited by Darwin 1871:67)

[6] Hymes (1961) argues that we can recognize evolutionary advance in particular languages if we broaden our view to include writing as well as speech, formal and scientific discussion as well as conversation, international affairs as well as local ones. This "increase in range and variety of adjustments to environments" would include the development of scientific vocabulary, of a meta-language to discuss linguistic structure, and freedom in borrowing forms from other language systems. But Hymes finds it necessary to "bypass the question of evolutionary advance in grammatical features . . . and the question of increased efficiency and economy in language evolution."

[7] Though this quotation is in the paragraph immediately following the one comparing biological and linguistic evolution, it is not cited in reference to the question of natural selection in language but in connection with the argument against evolution from the perfection of language. Darwin's contention that the "perfection" of a language, like that of an organism, is often overestimated on the basis of superficial characteristics, is as sophisticated as one might expect from any 20th-century linguist. "A Crinoid sometimes consists of no less than 150,000 pieces of shell, all arranged with perfect symmetry in radiating lines; but a naturalist does not consider an animal of this kind as more perfect than a bilateral one with comparatively few parts, and with none of these parts alike, excepting on the opposite sides of the body. He justly considers the differentiation and specialisation of organs as the test of perfection. So with languages: the most symmetrical and complex ought not to be ranked above irregular, abbreviated, and bastardised languages, which have borrowed expressive words and useful forms of construction from various conquering, conquered, or immigrant races" (1871:71). Here Darwin's own argument might have been used against the proposition that language shows progressive evolutionary adaptation to its environment.

There is general agreement among 20th-century linguists that language does not show an evolutionary pattern in the sense of progressive adaptation to communicative needs.

(4) Taking linguistic change as a whole, there seems to be no discernible movement toward greater efficiency such as might be expected if in fact there were a continuous struggle in which superior linguistic innovations won out as a general rule. (Greenberg 1959:69)

But it is not merely the absence of evidence for evolutionary adaptation that runs counter to Darwin's argument for natural selection. The almost universal view of linguists is the reverse: that the major agent of linguistic change – sound change – is actually maladaptive, in that it leads to the loss of the information that the original forms were designed to carry. Though there is a wide range of divergent opinions on the nature of sound change, as we saw in Part D of volume 1, there is general agreement on the negative character of this fundamental process. Throughout the 19th century, the basic mechanism of change was seen as dysfunctional, and historical linguists aligned themselves firmly with the enemies of sound change.

Language change as a destructive force

In 1816, Franz Bopp outlined principles for examining the history of language, including a recognition of the "gradual and graded destruction of the simple speech organism . . . and the striving to replace it by mechanical combinations . . ." (Lehmann 1967:43). Rasmus Rask recognized that the recovery from the effects of sound change is not a simple and immediate process: "grammatical inflections and endings are constantly lost with the formation of a new language . . . and it requires a very long time and intercourse with other people to develop and rearrange itself anew" (Lehmann 1967:32). In his first treatment of the Germanic sound shift, Jakob Grimm made it plain that such changes in the sounds of a language were destructive and unfavorable, and referred to them as "barbarous aberrations from which other quieter nations refrained" (Waterman 1963:20). Alexander von Humboldt treated phonetic constraints as a whole as unnatural, which must be subordinated to an intellectual factor, analogical reformation, which corrects them. "What has already been established to a certain extent in the phonetic pattern, violently seizes the new formation and does not permit it to pursue an essentially different path" (Von Humboldt 1836:56).

It is curious to find that Max Müller, Darwin's chief supporter for the idea that natural selection governed the evolution of words, himself characterized the major agent of language change as a process which destroyed the nature of language, and so caused "the life of language to become benumbed and extinct" (Müller 1861:54). August Schleicher was perhaps

even more negative, and saw changes in both the sound and form of language as a decay or decline, with a consequent loss of meaning (Lehmann 1967:90). W. D. Whitney was an even stronger exponent of the destructive force of sound change:

(5) A language may become greatly altered by the excessive prevalence of the wearing out processes, abandoning much which in other languages is retained and valued. It is necessary that we take notice of the disorganizing and destructive workings of this tendency inasmuch as our English speech is . . . the one in which they have brought about the most radical and sweeping changes. (Whitney 1904:75)

In declaring that sound change was regular and exceptionless, and distinguishing it clearly from analogy, the Neogrammarians did not depart from the view that its effects were harmful to the major functioning of language. Hermann Paul did not see sound change as unnatural in itself; he attributed it to physiological factors that followed the laws of physics. At the same time, he expressed most eloquently the general consensus on the destructive character of the process:

(6) Thus the symmetry of any system of forms meets in sound change an incessant and aggressive foe. It is hard to realize how disconnected, confused, and unintelligible language would gradually become if it had patiently to endure all the devastations of sound change. (1891:202)

The negative evaluation of sound change continued in the 20th century, though it was not to be expressed so violently. Saussure summed up the situation in this way:

(7) That phonetic evolution is a disturbing force is now obvious. Wherever it does not create alternations, it helps to loosen the grammatical bonds between words; the total number of forms is uselessly increased, the linguistic mechanism is obscured and complicated to the extent that the irregularities born of phonetic changes win out over the forms grouped under general patterns . . . (1949:161)

The linguistic consequences of sound change

Readings in the work of 19th- and early 20th-century linguists make it abundantly clear that they saw sound change as the primary, most systematic and omnipresent mechanism of linguistic change. It was evident to them that sound change interacted with morphological systems, disrupting paradigms, inserting asymmetries, and collapsing the fundamental distinctions that the system maintained; indeed, Saussure devoted an entire chapter to this topic. It was in terms of morphological systems that analogical change

was seen most easily as the restorer of paradigmatic symmetry and efficiency. Since analogy is notoriously sporadic and difficult to systematize, sound change became almost by default synonymous with the notion of "linguistic change." Now that the study of syntax has emerged as a major part, perhaps the major part of linguistic structure, and studies of syntactic change have begun in earnest, it may seem that the study of sound change would become an ever smaller part of the study of linguistic change. Since the present work focuses on change in progress, and comparatively few examples of syntactic change in progress have been located, it might also seem that the study of present trends can give us very little access to linguistic history.[8] Nevertheless, it can be argued that change in the surface phonetics remains the driving force behind a very large number of linguistic changes, perhaps the majority. This includes the processes of cliticization, which triggers any number of syntactic consequences, vowel contraction and consonantal assimilations, shifts of syllabicity and reassignment of syllable boundaries, along with the vast body of segmental changes – lenition and fortition, deletion and epenthesis, monophthongization and diphthongization, change of place and fusion of features, and the development of tone and its intersection with intonation patterns. We receive with increasing frequency the suggestion of wholesale reorganizations of prosodic systems as a causal factor in linguistic change, but what phonetic changes trigger such prosodic revolutions are only dimly perceived. Some of the issues involved here will be developed more fully in volume 3. Here it will be sufficient to state the proposition that a study of the causes and effects of changes in the sound system remains the primary prerequisite for explanation and evaluation of linguistic change in general. Changes in the sound system here refer not only to low level phonetic change but to morphophonemic condensations specific to particular grammatical locations.

Given that understanding, we have little basis for quarreling with the 19th-century understanding of the effects of sound change upon language as a whole. The view of language change as pathological is not mere rhetoric. Volume 3 will report extensive studies of the cognitive consequences of sound change which document by observation and experiment that sound change has led to a considerable degree of mutual unintelligibility of the phonologies of North American dialects. Volume 1 brought forward quantitative linguistic evidence that the reduction of functional elements of high frequency, noted in parallel 5 in Darwin's list above, can hardly be seen as an improvement. As chapter 20 of volume 1 showed, the loss of final sounds in Spanish and Portuguese leads to a measurable loss of information. When tautosyllabic final /s/ disappeared in French, a number

[8] The most important being studies of the elaboration of syntactic structures in developing pidgins and creoles (Sankoff and Laberge 1973), which will play a prominent role in this volume and the following one.

of compensating processes preserved the plural meaning. But there remain many cases where the plural meaning can no longer be signaled in standard French by grammatical means. Thus De Gaulle once declared in a public speech, *Je m'addresse aux peuples* . . . As a result of French sound changes completed many centuries ago, the *x* of *aux* in pre-consonantal position and the *s* of *peuples* in final position exist only in writing: singular *au peuple* and plural *aux peuples* are homonymous. De Gaulle was forced to recognize the inability of spoken French to distinguish singular and plural at this point by adding the meta-comment *au pluriel*.

One of the most widely studied processes of consonant reduction in English is the simplification of clusters ending in /t/ or /d/. In my own Northern New Jersey speech, the high frequency of simplification of the consonant cluster *nt* in *can't* has made it difficult to distinguish positive *can* from negative *can't*.[9] It is not uncommon for a speaker of this dialect to ask, "Did you say C-A-N or C-A-N-T?"

Perhaps the most dramatic shortening of words has occurred in the history of northern Mandarin, with a consequent augmentation in the number of homonyms. In compensation for this, most Mandarin words are now two characters or morphemes instead of one. It would be difficult for Darwin to argue that the shorter form had triumphed due to its own inherent virtue, when in compensation it developed a form that is roughly twice as long.

What then are we to make of Darwin's final statement?

(8) The survival or preservation of certain favoured words in the struggle for existence is natural selection.

One can hardly argue with this conclusion: in this form, it is nothing but a restatement of the fact that some words survive and others do not. But its significance depends on the answers to two questions: (a) Are the factors that lead to the survival or preservation of individual words the same as those that operate to form the abstract sets of relations between sound and meaning? (b) Can the survival of particular forms or relations be shown to be the result of adaptation of language to its environment? So far, the answers to both questions are "probably not."

As far as words are concerned, the replacement of vocabulary seems to have many of the characteristics of random variability. It is not simply the existence of statistical regularity[10] which leads to this conclusion; it is the apparent impossibility of saying which words have a better chance of surviving and which do not, whether abbreviations will persist, and whether

[9] The New York City and Philadelphia rule that laxes the vowel of auxiliary *can* and distinguishes it from the tense vowel in *can't* does not operate here.

[10] That is, the lexicostatistic finding that roughly 19% of the basic vocabulary is replaced every 1000 years (Swadesh 1971).

at a given time the vocabulary will expand or contract.[11] The same situation does not prevail for linguistic structure, where a number of directional principles have emerged. The study of sound shifting has shown that vowels have a high probability of moving in a particular direction (volume 1: chs 5–9), that mergers expand at the expense of distinctions (volume 1: chs 11–14), that vowels before liquids are far more likely to merge than vowels before obstruents, that inflections in some positions in the paradigm are much more likely to disappear than others (Greenberg 1969), and that in general, heavily marked structures are less stable than unmarked ones.

As far as structural alterations are concerned, the consensus is reflected in quotations (5, 6, 7). Sound change, the most general and pervasive source of such changes, is not the result of any adaptation of language to its environment. Though analogy and dialect borrowing may compensate for some of the damage to linguistic structures caused by sound change, their operation is far too episodic and unpredictable to be compared to the systematic operation of natural selection.

Thus we cannot support Darwin's hope to complete the fifteen parallels between biological and linguistic evolution by including a sixteenth parallel: natural selection. We might sum up the situation as **Darwin's paradox**:

(9) **The evolution of species and the evolution of language are identical in form, although their fundamental causes are completely different.**

Throughout this volume, we will be alert to the possibility of responding to this paradox. It would be strange indeed if the detailed resemblances between linguistic and biological evolution were in no way dependent on the fundamental mechanism of change. It would be too ambitious to say that this paradox can be resolved; it can be interpreted and dealt with in ways that will suggest the shape of a resolution.[12]

One immediate way of reducing the force of this paradox was suggested by Darwin himself in the sentence immediately preceding his final conclusion (8).

(10) To these more important causes of the survival of certain words, mere novelty and fashion may be added; for there is in the mind of man a strong love for slight changes in all things.

[11] This is seen most clearly in the rapid replacement of the slang vocabulary, which affects many words but not others. While *super*, *swell*, *nifty*, and *keen* have shown signs of obsolescence in American English, *fantastic*, *great*, *terrific* have not, over a comparable period of time.

[12] Among recent writers on linguistic evolution from a sociolinguistic point of view, Chambers (1995) argues the most vigorously for the adaptive value of linguistic variation, and this seems consistent with Weinreich, Labov, and Herzog's (1968) argument that a thoroughly homogeneous language would be dysfunctional. Chambers does not, however, examine new linguistic changes in progress and their disruptive effect upon communication.

The desire for novelty is introduced by Darwin as a less important, minor factor. But if the major factor of natural selection is discounted, then we might conclude that the driving force behind linguistic evolution is random variation. Indeed, genetic variability is an important component of the evolutionary mechanism in biology. There are sufficient random variables in linguistic structure to account for the gradual isolation of languages separated by geographic barriers to communication, so that the diversification of Oceanic languages in Micronesia and Polynesia would be comparable to the development of distinct species on the Galapagos Islands. Yet variability is only a necessary condition for biological evolution: without natural selection, variability is not sufficient to account for the rapid evolution of distinct species and radiation of organisms with different adaptive structures into distinct evolutionary niches. Elevation of the novelty principle (10) to the major factor in linguistic evolution may reduce the force of the Darwinian paradox to (9′):

(9′) **The evolution of species and the evolution of language are identical in form, although the fundamental mechanism of the former is absent in the latter.**

Such a re-formulation would only sharpen the problem of understanding the causes of sound change. No amount of nondirectional variability or drive to exaggerate that variability can account for the directional chain shifts, mergers, and splits that were presented in volume 1.[13] If there is no adaptive radiation in language, and no natural selection, what then are the fundamental causes of sound change? There is no shortage of answers to this question. Before we attempt to apply current findings to the problem, it may be helpful to review the answers that have already been given.

1.3 Earlier proposals for the causes of sound change

From the beginning of the 19th century, linguists have made many efforts to identify the causes of sound change. Those who have considered the matter most deeply give a uniform report on the difficulty of the problem. In 1856, von Raumer summed up the state of current knowledge in this way:

(11) . . . we ascertain that the sounds of words have changed when we compare the older state of languages with the more recent. The process of the change itself however has not yet been investigated enough. If we penetrate deeper into the darkness which in many ways veils these questions, we find a huge multitude of highly different processes at work. (1856:72)

[13] For an effort to account for sound change as simply the random drift of the mean value around which tokens of a phoneme are dispersed, see Hockett 1958:441.

Some 60 years later, Saussure reviewed the situation in similar terms:

(12) The search for the causes of phonetic changes is one of the most difficult problems of linguistics. Many explanations have been proposed, but none of them thoroughly illuminates the problem. (1949 [1916]:147)

Finally, we may quote Bloomfield, writing in 1933:

(13) Although many sound-changes shorten linguistic forms, simplify the phonetic system, or in some other way lessen the labor of utterance, yet no student has succeeded in establishing a correlation between sound-change and any antecedent phenomenon: the causes of sound-change are unknown.

In spite of these cautions, many linguists have argued strongly for a particular explanation of sound change, and in the course of time a great many approaches to the topic have been exposed to argument and debate. Some will be of more value than others in our exploration here.

Beginning with the less valuable, we find that many explanations of linguistic change put forward in the 19th and early 20th centuries were materialist in spirit, but were supported by only the weakest kind of empirical evidence. As more data accumulated, explanations based on climate or topography are easily set aside as the counter-examples come to outnumber the examples. Furthermore, the mechanisms that were proposed for the link between the cause and the effect usually seem to us today naive in the extreme.[14] Of even less interest are the explanations based on physiological differences between speakers of various languages, which seem to be motivated more by convictions of racial superiority than scientific evidence.[15] The traditional arguments advanced for the causes of sound change that will most concern us here are basically three: the principle of least effort, the principle of density, and the principle of imitation.

The principle of least effort

This principle seems to have been a part of linguistic thinking about change from the very beginning. It is cited today most often in the formulations of Saussure (1949:148–9), of Jespersen (1921), and of Bloomfield. Bloomfield's seems the most precise:

(14) It is safe to say that we speak as rapidly and with as little effort as possible, approaching always the limit where our interlocutors ask us to repeat our utterance, and that a great deal of sound-change is in some way connected with this factor. (1933:386)

[14] As for example that speakers in cold climates had to keep their mouths closed to prevent the cold air from entering their vocal cavity, and so had fewer open vowels.

[15] See Saussure's critical review of the arguments advanced for the causes of phonetic changes (1959:147).

In this formulation, the principle of least effort is a precise structural principle, bound between two limiting factors that determine exactly the extent of the reduction concerned. If the principle were made one step more explicit, it might state that the reduction of phonetic form stops at exactly before the point where information would be lost. Let us try to restate this accordingly:

(14′) Principle of least effort I. We speak with the least effort that is required to be understood by our addressees, but with sufficient effort to ensure that we are understood.

However, (14′) is not consistent with the view that sound change is destructive of meaning, as cited in (5, 6, 7). The very term *least effort* implies a limiting, asymptotic factor which can only be the preservation of meaning. Bloomfield did not disagree in the least with the traditional view that sound change destroys meaning.[16] If not, then the principle of least effort would require an alternative formulation of a very different character:

(14″) Principle of least effort II. We speak with less effort than is required to convey all of our meaning to our addressees.

But (14″) loses the characteristic Bloomfieldian precision. It says nothing of interest on how much or by what cause we fall short. (14′) defines how much reduction is possible, and implicitly attributes the reduction to a rational principle of efficiency, while (14″) says very little at all. To have any interest for a theory of language change, it would have to be reinforced as

(14‴) Principle of least effort III. Under the influence of factors $\mathbf{a}_1$, $\mathbf{a}_2 \ldots \mathbf{a}_n$, we reduce the phonetic information that we convey to our addressees, sometimes to the point that they do not understand us.

At this point, the principle of least effort would no longer lie at the focus of efforts to explain change. Rather, the task would be to identify the factors that lead to this behavior. *Laziness*, *carelessness*, and *ignorance* are perhaps the most frequent candidates for underlying causes, not only in popular treatments but in scholarly works of the 19th century. In his general treatment of the causes of linguistic change, Whitney refers to "linguistic degeneration," which is caused by:

(15) the wholly regrettable inaccuracies of heedless speakers, their confusion of things which ought to be carefully held apart, their obliteration of valuable distinctions. (1904:84–5)

[16] "In fact, sound-changes often obliterate features whose meaning is highly important . . . Homonymy and syncretism, the merging of inflectional categories, are normal results of sound-change" (1933:388).

Here Whitney focuses on the carelessness of these speakers, though elsewhere he deals equally with laziness and ignorance. At first glance, the three terms *laziness*, *carelessness*, and *ignorance* seem simply to express of the same moral disapprobation. Yet they can be differentiated in their implications for the mechanism of sound change when we introduce the dimension of rapidity of speech. In dealing with the principle of least effort, Whitney points out that

(16) we may call it laziness, or we may call it economy . . . it is laziness when it gives up more than it gains; economy, when it gains more than it abandons. (1904:70)

Syllable length is regularly associated with degree of approximation to target articulations, not only for the syllable nucleus, but for syllable margins as well.[17] The low level of effort associated with laziness would generally be correlated with slow speech, while the low level of effort attributed to carelessness would be associated with rapid speech. Thus for the careless speaker, a low level of attention or effort directed to the norms of correct speech would combine with the mechanical temporal effect of shorter time to reduce the phonetic information produced, while for the lazy speaker, the temporal effect would operate in the opposite direction.

On the other hand, ignorance has no direct relationship either to tempo or to the principle of least effort. If the speakers described by Whitney are indeed ignorant of classical and time-honored usages and of valuable distinctions, the changes in their language cannot be ascribed to the principle of least effort. Thus in the course of a merger, speakers who are aware of the distinction between *whale* and *wail* might neglect it through carelessness or laziness; it is their children who would then complete the sound change through ignorance of the distinction.[18]

Speech tempo

The tempo of speech may be considered a distinct factor in sound change, since tempo may vary independently of effort. Bloomfield notes that Wundt attributed sound change to an increase in the rapidity of speech, and this in turn to the community's advance in culture and general intelligence. In contrast to the other treatments of the period, Wundt did not view such changes as a consequence of human failings, but rather as a product of

[17] This applies to changes in progress as well as stable variation. Volume 1, table 18.1 illustrates how following syllables, which shorten syllable duration, restrict the raising of (æh) for Carol Meyers in Philadelphia.

[18] In this case, since the distinction between voiced /w/ and voiceless /ʍ/ is still registered in spelling, and some schools continue to teach this usage, the younger generation could be said to exhibit carelessness, laziness, and ignorance in failing to acquire it in their formal speech.

intelligent behavior, associating greater speed of speech with higher intelligence. It is generally recognized that morphophonemic condensations that involve cliticization, syncope, degemination, consonant cluster simplification, and assimilation are associated with fast speech rules (Dressler and Grosu 1972, Gay 1977, Kaisse 1977, Beckman et al. 1992), and that over the course of time some of these contractions become institutionalized in the more formal structure of the language or as underlying forms.[19] Sankoff and Laberge 1973 report that one of the characteristic differences between native speakers of Tok Pisin and second language speakers is that native speakers talk much faster. Later, we will examine the linguistic changes that accompany this increase in tempo.

Discontinuities in communication

A more general explanation of linguistic change was advanced by Bloomfield in his treatment of dialect geography, cited extensively in Part C of volume 1. In dealing with the high degree of differentiation in local European dialects, he wrote:

(17) The reason for this intense local differentiation is evidently to be sought in the principle of density. Every speaker is constantly adapting his speech-habits to those of his interlocutors; he gives up forms he has been using, adopts new ones, and perhaps oftenest of all, changes the frequency of speech-forms without entirely abandoning any old ones or accepting any that are really new to him. The inhabitants of a settlement, village, or town, however, talk much more to each other than to persons who live elsewhere. When any innovation in the way of speaking spreads over a district, the limit of this spread is sure to be along some lines of weakness in the network of oral communication, and these lines of weakness, in so far as they are topographical lines, are the boundaries between towns, villages, and settlements. (1933:476)

To the extent that this is true, a large part of the problem of explaining the diffusion of linguistic change is reduced to a simple calculation. Given the degree of variability indicated above, discontinuities in the networks of communication would inevitably lead to a random drift of neighboring dialects in different directions. Though Bloomfield thought that his own hypothesis was not within reach of empirical confirmation, it can be tested with figures on vehicular traffic and telephone communication. An examination of the dialect boundaries of the Eastern United States on the basis of average daily traffic flow show that Bloomfield's hypothesis holds for all

[19] Hock 1986:352–4 gives a characteristic example of the reduction of the relative marker *yo* in Old Irish, leading to a reinterpretation of the lenition of the following consonant as the relativizing signal.

boundaries but one (Labov 1974). Furthermore, the principle of density implicitly asserts that we do not have to search for a motivating force behind the diffusion of linguistic change. The effect is a mechanical and inevitable one; the implicit assumption is that social evaluation and attitudes play a minor role.

Language and dialect contact

No treatment of the causes of linguistic change could be complete without a consideration of the effect of one system on another. Extensive treatments of the effects of dialect contact on language change are available in Trudgill 1986, Kerswill 1993, Chambers 1995, and Williams and Kerswill 1999. The present work is primarily concerned, however, with those changes that emerge from within a linguistic system, in which the problem of causation arises in its sharpest form.

Optimization of communicative function

In more recent times, a number of theories of linguistic change have portrayed the process as part of a smoothly functioning mechanism that serves to maximize the communication of information. Far from interfering with communication, change is seen as maximizing the flow of information and the ease of obtaining it. The most prominent of these accounts is the functional approach of Martinet (1955). Martinet sees most sound changes as governed by the need to maximize the distinctiveness of phonemes. Phonemes shift their target positions and their fields of dispersion in order to preserve their margin of security. The instability of phonetic systems is due to the presence of two conflicting pressures: the psychological preference for symmetry, and the asymmetrical construction of the organs of articulation. Thus there is a tendency to preserve symmetry with the same number of distinctions of height in the front and the back, and a tendency to have fewer distinctions in the back since there is a smaller physiological space to differentiate back vowels.

Among the many empirical demonstrations of Martinet's position, two rank among the most substantial. Moulton (1962) demonstrated that the position of the allophones of /a:/ in Northern Switzerland was highly determined by the configuration of other low and mid long vowels in the front and the back. Haudricourt and Juilland (1949) showed that in a large range of languages, the fronting of the nucleus of /u/ and /o/ was associated with a reduction of the number of degrees of height among the back vowels from four to three. In volume 1, it was seen that the force of their argument is somewhat diminished by the extensive fronting of the long back vowels in American English dialects with a prior merger of long and short open /o/ and only three degrees of height in that region. Liljencrants and Lindblom

(1972) support the tendency to maximal dispersion of vowels with a review of reported vowel systems and a numerical simulation.

Chapter 20 of volume 1 argued that the chain shifting of vowels could be seen as conforming to Martinet's principles of maintaining margins of security in a way that maximized the efficiency of communication. The proposed mechanism of shift depended upon the consequences of misunderstanding: outlying vowel productions in the direction of smaller margins of security would have a greater tendency to be misunderstood than those in the direction of larger margins of security. As a result, the mean number of tokens in the data base available to a language learner would be shifted in the direction of the greater margin of security, and the field of dispersion of the phoneme would expand in that direction. In this mechanism, the teleological aspect of functional explanations disappears. However, two major problems remain unresolved under the functional explanation of chain shifting. First, it does not account for the massive mergers that are as common as or more common than chain shifts (volume 1, chs 10–12). Secondly, it does not include any account of the driving force that moves the vowel system in the first place.

More recent efforts to explain linguistic change depend upon more abstract characteristics of rule systems. King (1969) proposed to account for all linguistic changes as forms of rule simplification, though King (1975) retracted this argument in favor of a multivariate approach that takes social factors into account. Kiparsky (1971, 1982) argued that linguistic change tends to favor feeding relations of rules, maximizing their application, and that change also tends to minimize opacity and maximize transparency. Volume 1 argued that the most characteristic sound change is a change in the phonetic realization of a phoneme at a low level of abstraction, a postlexical output rule. The symmetrical generalization of such rules would represent rule simplification and maximization of application.[20] It was also argued in volume 1, chapter 8 that chain shifting could be treated as a unitary process no different in character from parallel shifting. Explanations from rule systematization could then compete with the functional explanations of Martinet.

To what extent can the various causes of sound change advanced be seen as adaptations of language to its environment and environmental needs? Here we must draw a fine line between the facilitation of communication and communication itself. Many factors involve shortening of the effort, mental or physical, required for the act of communication. The principle of least effort is such a form of facilitation, as is rule simplification (which may facilitate acquisition as well as production) and the maximization of

[20] Parallel arguments arise in the constraint-based mechanism of optimality theory; the generalization of a rule corresponds to an elevation in the ranking of a more general constraint.

transparency (which facilitates interpretation and acquisition). On the other hand, functional explanations proper are usually based on maximization of the information conveyed – either by increasing the total amount of information in the signal, or by calling the receiver's attention to a particular piece of information.

The pessimistic views (11, 12, 13) expressed by the major thinkers about linguistic change can now be understood more clearly. For each explanation brought forward, there is a competing explanation which can account for the same change. The Great Vowel Shift can be seen as the preservation of distinctions in the face of some unknown force that raised the long vowels, or as a generalization of that raising that simplifies the phonological system. Not only are there competing explanations for each phenomenon, but most of these explanations can predict the opposite of what occurred. Thus the maintenance of any given distinction, like that between /w/ and /ʍ/, conveys more information, but makes it harder to learn the language – not only because there is one more distinction to be maintained, but because one term of the opposition is a marked articulation. The generalization of a change from front to back vowels simplifies the structure of the system, but frequently leads to mergers in the back vowels.

There is a deeper problem that makes all of these explanations less than satisfactory. They all depend upon some permanent properties of the organism of the language structure; yet sound change is characteristically sporadic, accelerating at unpredictable rates and terminating at unpredictable times. Bloomfield was well aware of this aspect of the problem:

(18) Every conceivable cause has been alleged . . . No permanent factor, however, can account for specific changes which occur at one time and place and not at another.

Saussure is more elaborate on this point:

(19) . . . why did the phenomenon break through at one time rather than another? The same question applies to all the preceding causes of phonetic changes if they are accepted as real. Climatic influence, racial predisposition, and the tendency toward least effort are all permanent or lasting: why do they act sporadically, sometimes on one point of the phonological system and sometimes on another? (1959:15)

Meillet gave a precise answer to these questions:

(20) From the fact that language is a social institution, it follows that linguistics is a social science, and the only variable element that we can resort to in accounting for linguistic change is social change, of which linguistic variations are only consequences, sometimes immediate and direct, more often mediated

> and indirect. . . . We must determine which social structure corresponds to a given linguistic structure, and how in general changes in social structure are translated into changes in linguistic structure. (Meillet 1926:17–18; my translation)

This quotation is from Meillet's inaugural lecture of February 1906 as he assumed the professorial chair of the College de France formerly held by Bréal. The lectures that followed were explicitly devoted to this program. Meillet's social arguments draw upon well-established facts about the social relations of speech communities, and only occasionally refer to dialect differences or variation within the community. Nevertheless, his insight remains fundamental to the sociolinguistic approach to linguistic change developed in this volume. Curiously enough, the main proposal for the social correlates of language that was advanced in Meillet's time was that of Gabriel Tarde, who considered his own theory of society to be a major competitor to that of Durkheim.

Imitation

In his *Laws of Imitation* (1873), Tarde developed a general theory of language based on an "inter-psychology" of individuals, diametrically opposed to the Saussurian concept of *langue* as a social fact. His argument included a theory of language change:

(21) It appears to me almost beyond dispute that language is a phenomenon of imitation: its propagation from high to low, from superior to inferior, whether it be without or within the nation, the acquisition of foreign words by fashion and their assimilation by custom, the contagion of accent, the tyranny of usage in itself, suffices to show at one glance its imitative character. (Tarde 1873: ch. 5)

Tarde discussed at length the nature of the creative act performed by these innovators. These are portrayed as superior individuals who are imitated by the "public" that admires them. Tarde also recognized the "law of least effort," but did not see it as the product of careless and heedless speech. Instead, he saw its operation as "inevitable" and "teleological," that is, tending toward an efficient form of simplification. At the same time, he insisted on a complementary force of "phonetic reinforcement," which "serves to introduce a new sense or emphasize the expression of an accepted sense." Far from regarding sound change as a blind, mechanical force, Tarde regarded it as a positive, creative process; he saw no clear separation between semantic and phonetic change.

Though Tarde's view of imitation is unidirectional in the social hierarchy, it is not necessary to limit the process to a transfer of features from higher to lower social groups. To explain the fact that speech forms stigmatized

by the dominant social classes are maintained over long periods of time, and even expand in the face of that stigmatization, one is forced to consider the existence of an opposing set of values that do not readily emerge in formal situations (Labov 1972b:313), and some firm evidence has been produced for the existence of such covert prestige (Trudgill 1972, Labov et al. 1968). However, the force of Tarde's explanation may be considerably weakened if the term "prestige" is allowed to apply to any property of a linguistic trait that would lead people to imitate it. Thus the fact that a linguistic form has prestige would be shown by the fact that it was adopted by others.

Differentiation and alignment of social groups

Bloomfield's principle of density, given as (17) above, dealt with geographic differentiation. He later generalized this principle to apply to social differentiation in a single community, in a description that applies closely to the results of recent sociolinguistic studies of urban communities:

(22) We believe that the differences in density of communication within a speech community are not only personal and individual, but that the community is divided into various systems of sub-groups such that the persons within a sub-group speak much more to each other than to persons outside their sub-group. The lines of weakness and, accordingly, the differences of speech within a speech community are local – due to mere geographic separation – and non-local, or as we usually say, social.

This account plainly gives a picture of the growth of social differentiation within the community, and in particular of the divergence of the dialects spoken by highly segregated racial groups in North American cities (Labov and Harris 1986, Bailey 1993). However, it does not account for the progressive diffusion of linguistic change across social groups, which is one of the main phenomena that we have to deal with in this volume, or the way in which the entire speech community advances in the course of linguistic change.

The orientation to the relations of language and society that is closest to my own point of view is that of Sturtevant (1947). He viewed the process of linguistic change as the association of particular forms of speaking with the social traits of opposing social groups. Those who adopt a particular group as a reference group,[21] and wish to acquire the social attributes of that group, adopt the form of speaking characteristic of that group. The opposition between the two forms of speaking continues as long as the social opposition endures, and terminates in one way or another when the social distinction is no longer relevant.

[21] In the technical sense developed by Merton 1957.

1.4 Different kinds of sound change

The various proposals for the causes of sound change cover physical, psychological, and social parameters of the speaker's situation. But very few of these discussions discriminate among the different kinds of sound change involved, in spite of the fact that many of the causes proposed apply to only a limited range of types. In the effort to bring empirical evidence to bear upon these proposals, it will be essential to make that discrimination.

Sound shifts

For all of the discussions involving the principle of least effort, it is obvious that the writer was focusing, consciously or unconsciously, upon changes that reduce the amount of phonetic information provided by the articulation of speech.[22] Yet the principle of least effort applies primarily to changes of manner: consonant lenition, vowel reduction, and the deletion of segments. It does not apply at all to those sound changes that alter place of articulation, like the shift of Austronesian /t/ to /k/ in Hawaiian, the shift of apical obstruents to velars in Skikun discussed by Li (volume 1: 16–17), or any of the vowel shifts presented in volume 1:5–9. In both the Northern Cities Shift and the Southern Shift, a majority of the vowel changes involve an increase in the complexity and energy of articulation – lengthening, diphthongization, and movement to more extreme positions in phonetic space. Very few of these vowel shifts involve shortening, and even the laxing of diphthongal nuclei to the nonperipheral track frequently involve an increase in nucleus–glide differentiation within the syllable. Typically, the shift of New York /ay/ from [aɪ] to [ɒːi] and Philadelphia /aw/ from [æʊ] to [eːɔ] comprise a considerable increase in the length and complexity of the trajectory of the vowel. In general, changes in the place of articulation of segments cannot be explained by the principle of least effort, or by any of the factors that are used to motivate that principle: laziness, careless, or ignorance. Nor would rapidity of speech apply to such sound changes. Many efforts have been made to show that these changes represent an optimization of rule systems, though an overall assessment seems to show as much complication as simplification. The imitation of dominant social groups seems equally unlikely, since when such sound changes come to public attention, they are almost always stigmatized by the dominant social groups. Certainly sound shifts can carry social evaluation, as demonstrated by the subjective reaction tests carried out in New York City (Labov 1966a: ch. 12) and Philadelphia (chapter 6,

[22] Whitney is perhaps the most explicit on this point; lenition and deletion are the only types of phonetic change that are considered in his entire volume (1904).

this volume). It is an empirical question as to which sound shifts are the vehicles of which social values associated with which other groups in the social spectrum.

As opposed to sound shifts, there are many other types of changes that reduce phonetic information in the speech chain over time. Though lenition, merger, and deletion are all similar in this respect, they have different relations to the proposed causes of sound change presented above.

Lenition

There is no shortage of sociolinguistic variables that are characterized by the lenition of phonetic forms. Studies of variation have focused on the aspiration of Spanish and Portuguese postvocalic (s), the vocalization of postvocalic (n) in Portuguese, Chinese, and African-American English, the loss of initial aspiration and vocalization of liquids in English, and so on. We can observe in Liverpool a modern counterpart of the lenition of voiceless stops to fricatives that marked Grimm's law, this time in post-vocalic position. These are paradigmatic candidates to register the influence of the principle of least effort. In so far as these changes can be represented as a reassignment or spreading of features, rather than merger or loss of features, they are not easily interpreted as responding to the need for optimization of rule systems. On the other hand, they are usually not as heavily marked for social evaluation as other sociolinguistic variables and show more moderate stylistic shift with increase of audio-monitoring. Thus the role of imitation and reference group association may be not as prominent for changes involving lenition.

The data available for the empirical investigation of such changes in progress is limited. Most of the variables that have been studied are now quite stable, though they are undoubtedly the product of active sound changes at some time in the past. Chapter 3 of volume 1 gave real-time and apparent-time evidence for change in progress in Cedergren's study of the lenition of (ch) in the Spanish of Panama City (1973, 1984). The vocalization of (l) in American English appears to be a recent and vigorous change in progress from below (Ash 1982a,b), and we will draw heavily upon this phenomenon for an understanding of the social trajectory of a lenition rule later in this volume.

Mergers and splits

In historical comparative linguistics, "sound change" is almost equivalent to merger, since mergers are the changes that are preserved most clearly in the historical record and in the comparison of languages. The innovations that identify nodes in family trees are therefore most heavily concentrated in this type of sound change. Whether the principle of least effort applies

to such mergers is an interesting question; I do not know of any discussions of the topic. Clearly a merger represents a reduction in the amount of information provided by the speaker, though the mechanism of merger proposed by Herold represents merger as a gain of information (1990; see volume 1, ch. 12). One might argue that a merger is a conceptual type of least effort, just as the perseverance of variables or concord of number or gender may be argued to facilitate speech production.

When changes in place of articulation are accompanied with conditioned mergers, they may have strong effects upon the morpheme structure rules of a language and so are subject to arguments of rule simplification. Thus one stage of the reduction of final consonants in unstressed syllables of Greek, Italic and Romance, and Germanic led to a severe limitation on the features found in final position and in affixes – essentially to apicals. In many formulations, this would lead to a great simplification of the phonological representation of grammatical formatives.

In the area of social evaluation, mergers are distinctly opposed to sound shifts. It will become evident that mergers are almost invisible to social evaluation, and it is difficult to think of them as diffusing under the social pressures of social imitation and association. There are many mergers in progress in American English, and we have ample data on their social distribution. The discussion of the merger or near-merger of *ferry* and *furry* in Philadelphia in volume 1, chapter 14 provides a fine-grained view of variation at different levels of attention to linguistic categorization, and this will be related to a closer examination of stylistic effects on merger.

Deletions

The lenition rules discussed above are often closely coupled with deletions, which may also be thought of as the last stage of lenition, but also as alternations with zero or merger with zero. Quantitative studies have been more closely involved with this type of variation than any other. The deletion of final (s) and (n) in Spanish and Portuguese, the deletion of final /t/ or /d/ in English and Dutch consonant clusters, and the alternation of /s/ morphemes with zero in African American English, are located squarely on the intersection of phonology and morphology, where we can study most closely the relation between sound change and the information available in the speech signal (volume 1, ch. 20). They are certainly open to interpretation as consequences of the principle of least effort, particularly when it is formulated as constrained by the need to convey information. All of these deletions also have profound effects on the distribution of syllable types, and ultimately on the canonical form of the syllable, as does the deletion of initial /l/ and medial schwa in French. Though they may represent complications of the grammar in their initial form, they are open to interpretation as simplifications in their final stages.

Unfortunately, none of these deletion variables have been shown to be involved with change in progress, so that our view of the initial and final stages is limited to what we can glean from the historical and comparative record.

Deletions resemble lenitions in their availability for social evaluation: a limited amount of social affect is displayed, considerably less than for sound shifts. We find, for example, that the social stratification of *-t/d* deletion is not by any means as sharp as that of (ing), and that stylistic differentiation is even smaller (Labov et al. 1968).

1.5 The narrow interface between language and society

At one point in the development of sociolinguistics, it was not uncommon for scholars to suggest that the social and linguistic aspects of language were coextensive in the sense that each linguistic element had a social aspect or evaluation. Yet the actual situation seems to be quite the reverse. For the most part, linguistic structure and social structure are isolated domains, which do not bear upon each other. As indicated above, those sound changes with clear structural consequences – mergers – are almost entirely without social evaluation. The force of social evaluation, positive or negative, is generally brought to bear only upon superficial aspects of language: the lexicon and phonetics. However, social affect is not in fact assigned to the very surface level: it is not the sounds of language which receive stigma or prestige, but rather the use of a particular allophone for a given phoneme. Thus the sound [iːə] is not stigmatized in general, since it is the prestige norm in *idea*, but it is stigmatized as an allophone of /æ/ in *man*. Similarly, social criticism is not directed at the word *finalize*, but rather at the stem /faynəlayz/, since it is equally shared by *finalizing*, *finalized*, and *finalizes*.

The evidence for the isolation of abstract linguistic structures from social evaluation and differentiation comes from many sources. In the quantitative analyses of variation, it is found that changes made by the addition or subtraction of internal, linguistic factors are reflected in changes in the values of other internal, linguistic factors, while values of the external, social factors remain identical; the same situation applies inversely when external, social factors are added to or subtracted from the analysis (Weiner and Labov 1983, Sankoff and Labov 1979). When analyses of linguistic factors are carried out independently for different social classes or for men and women, very few significant differences are found in the values for the two social groups (Braga 1982). Though the overall level for socially marked variables may vary widely across age groups or social classes, the internal constraints show remarkable constancy (Kroch 1989). In those parts

of sociolinguistic interviews that deal overtly with language and its social evaluation, it is almost unknown for subjects to speak spontaneously of the existence or nonexistence of a contrast, or differences in conditions on rules. On the other hand, we will present in this volume evidence for strong social reactions to the phonetic realizations of particular phonemes, and for social evaluation of those realizations in particular words. Under some conditions, the presence or absence of particular grammatical formatives is remarked on, but primarily by those who have taken on the responsibility of enforcing a literary tradition.[23]

The relative segregation of social and structural elements in language is a major factor in distinguishing the possible causes of different types of sound change. Since this volume deals with the role of social factors in change, the primary focus will be on those elements of language that are most likely to be highly stratified in use and strongly evaluated in social perception.

1.6 The social location of the innovators

Sturtevant's views on the diffusion of sound change were based upon his own informal but penetrating observations of sociolinguistic patterns in the first half of the 20th century. His explanation of the path of linguistic change is essentially that it is a reflection of social change, responding most directly to the appeal of Meillet (1921). Yet it barely touches the question of the underlying causes of the continued renewal of change, and Sturtevant's brief comments do nothing to implement his views by showing how a particular change followed the course outlined. This volume will undertake that task, tracing the diffusion of linguistic change through the various layers of social structure. The goal is not only to describe the path of the change, but also to advance our understanding of its fundamental causes. The strategy to be followed here is to transform the traditional question "Why does language change?" into a different form: "Who are the leaders of linguistic change?"

Many of the earlier writers on sound change cited indicated that it would indeed be helpful to know which speakers were responsible for its initiation. If social factors are in fact connected with the onset and continuation of this process, it would be essential to know something about the social class, sex, ethnicity, or occupations of the innovators. For those

[23] There are exceptions to this generalization: negative concord in modern English is one. In one way or another the entire speech community shows sensitivity to this abstract structural pattern, which does not depend on the presence or absence of any one surface form. There is no change evident in this sociolinguistic pattern today, but chapter 3 will present some information on how this change in English structure came about.

who looked on sound change as an unmixed evil, the search was more or less a criminal investigation. Thus Whitney:

(23) Such phonetic changes . . . are inevitable and creep in of themselves; but that is only another way of saying that we do not know who in particular is to blame for them. Offenses needs must come, but there is always that man by whom they come, could we but find him out. (1904:43)

Whitney's description of the innovators cited above in (15) as "uncultivated and careless speakers . . . to whom the preferences of the moment are of more account than anything in the past or the future" is a classic description of the lowest social class in the eyes of the upper class. Such a prediction fits in with the negative character of the explanations that are usually advanced for sound change. But theoretical notions do not account altogether for Whitney's pursuit of these corrupters of the language. He plainly saw the opposition as one of social loyalties:

(24) New dialects are wont to grow up among the common people, while the speech of the educated and lettered class continues to be what it has been. (1904:44)

In condemning the effects of linguistic change, Whitney calls for social action to oppose them. The full quotation from which (15) was drawn is:

(25) The wholly regrettable inaccuracies of heedless speakers, their confusion of things which ought to be carefully held apart, their obliteration of valuable distinctions – all these are part and parcel of the ceaseless changes of language . . . they are only that part against which the best public sentiment, a healthy feeling for the conservation of linguistic integrity, arrays itself most strongly. (1868:84–5)[24]

In general, those linguists who pointed to the principle of least effort as the major factor in linguistic change would look for the most extreme examples of the change in progress among the lowest social classes. To the extent that discontinuities of communication within a speech community are the causes of change, with resulting ignorance of the normative standard, we would also expect to find the leaders of change in the lowest social class. However, the opposite prediction would be made by Tarde, who believed that linguistic change was always initiated by the highest group in the social

[24] The strongly moral overtones from these quotations, characteristic of 19th-century reflections on this subject, may be misleading. Whitney took a much more objective view of the effects of sound change than the moral overtones of these quotations convey. In fact, he saw "phonetic corruption" as the chief creative force in "the life and growth of language," as Lecture III from Whitney 1904 demonstrates throughout.

scale. Aligned with Tarde would be Wundt. Both believed that the condensations of rapid speech represented an increase in efficiency of speech, and were characteristic of the most intelligent and educated speakers.

Given this radical disagreement, it appears that there are significant theoretical consequences to the social location of the innovators of sound change. The Project on Linguistic Change and Variation in Philadelphia [LCV] accepted the challenge laid down by Whitney (23): to identify the social groups responsible for the continued course of sound change. Following Meillet's argument that the course of sound change must be accounted for by its interaction with social forces, we can identify those social forces by charting the position of the leading groups in the multidimensional fabric of the speech community.

The curvilinear pattern

As noted above, early theories of the causes of linguistic change would predict that the innovators would be at either the top or the bottom of the social hierarchy. The first sociolinguistic studies of change in progress, in Martha's Vineyard (Labov 1963) and New York City (Labov 1966a), did not find either of these patterns. The first general sociolinguistic model of the mechanism of linguistic change (Labov 1965) proposed that change within the system could originate in any social group, and following Sturtevant's suggestion, would spread gradually through each neighboring social group until it reached in one form or another all members of the community.

Kroch (1978) pointed out that there were no examples of systematic linguistic change (as opposed to borrowing from outside, or change from above) initiated by an upper class. He proposed a dichotomous model in which natural linguistic change was initiated by working class speakers, while middle and upper class speakers reacted against such changes, correcting their speech in a direction opposed to natural change.

In an early discussion of the social location of the innovators of change from below (Labov 1972b:294–5), it was stipulated that the highest social class is not "as a rule" the innovating group, but it was also pointed out that "it seldom happens" that innovation spreads upward from the lowest social group.[25] Instead, it was observed that the innovating groups were always located in an upper working class, or lower middle class, and that

[25] This formulation presumes an analysis of the social hierarchy into more than two components. The actual basis for the division of social classes seems to be immaterial: some support for the pattern described below comes from studies with education as the class indicator; others with occupation; and still others with combined indices. But unless three, or preferably four, divisions of the social hierarchy are distinguished, the curvilinear pattern will be concealed. Thus from the point of view of the study of linguistic change in progress, descriptions of communities in terms of upper vs. lower class, or middle vs. working class, are not informative and may actually conceal whatever change is taking place.

in many cases, these two groups were almost identical in the advancement of the change in progress in vernacular speech. Thus the crucial division in the society from the point of view of language change was not middle class vs. working class, but rather centrally located groups as against peripherally located groups.

From these observations was formed the **curvilinear hypothesis**: while stable sociolinguistic variables showed a monotonic social class distribution, a monotonic distribution in age groups was associated with a curvilinear pattern in the socioeconomic hierarchy. The major evidence for this hypothesis was drawn from the raising of (oh), (ay), and (aw) in New York City (Labov 1966a), the backing of (el) in Norwich (Trudgill 1974b), and the lenition of (ch) in Panama City (Cedergren 1973). Figure 1.1 is drawn

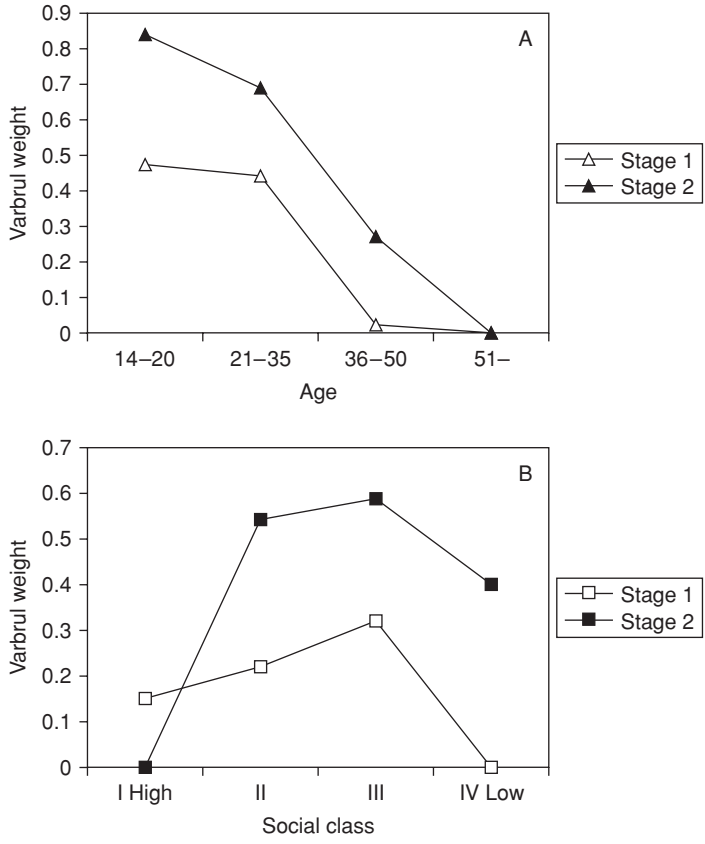

Figure 1.1 Curvilinear pattern shown by Varbrul weights for social constraints on the lenition of (ch) in Panama City. Stage 1: weakened affricate. Stage 2: fricative. (A) Monotonic function of age. (B) Curvilinear function of social class (adapted from Cedergren 1973)

from Cedergren's Panama City data: it shows the characteristic coupling of a steady rise in the (ch) index with progressively younger speakers, and a curvilinear pattern in the social class domain.

The Project on Linguistic Change and Variation in Philadelphia [LCV] was proposed explicitly to test the curvilinear hypothesis in a community where about two-thirds of the vowels appeared to be involved in change in progress. Chapter 3 of volume 1 presented the results of the studies of change in apparent time, reinforced by observations in real time, that led to the establishment of five levels of change within the vowel system. Of the eleven changes in progress described there, the most important for the hypothesis are the new and vigorous changes: the fronting and raising of (aw) in *out*, *down*, etc.; the raising and fronting of (ey) in checked syllables in *made*, *pain*, etc.; and the centralization of (ay) before voiceless consonants in *right*, *fight*, etc. The social distribution of these and other changes will be described in chapter 5, which will provide ample evidence to confirm or disconfirm the curvilinear hypothesis. The social location of the leaders of linguistic change will then be examined further along many other social dimensions, and the results will be applied to illuminate and perhaps reduce the Darwinian paradox.

1.7 Individual, group, community

Many writers on sociolinguistic themes, including those whose work plays a major role in this volume, have argued that the major focus of sociolinguistic analysis should be placed on the individual speaker rather than the group (L. Milroy 1980:133–4, Douglas-Cowie 1978; see also Fillmore, Kempler, and Wang 1979). If the net result of such a policy is to plunge more deeply into the internal composition of the group, it is likely to be productive. This volume will begin with larger components of social structure, and proceed with finer and finer analysis until the leaders of linguistic change are located as specific individuals. The main data base will be the 112 speakers of the Philadelphia Neighborhood Study whose vowel systems were analyzed acoustically. The leaders of linguistic change will be located as outliers within a particular social class, a particular gender, in specific positions within local social networks. To understand the forces operating in linguistic change, we will necessarily be focusing upon a handful of individuals. We will study their personal statements, their social histories, and their philosophies of life. This focus on individuals is not inconsistent with the argument of my 1966 study of New York City that the behavior of the individual speaker cannot be understood until the sociolinguistic pattern of the community as a whole is delineated.

This investigation is not a search for individuals, but rather for social locations and social types. The leaders of linguistic change are not individual

inventors of a certain form, but rather those who, by reason of their social histories and patterns of behavior, will advance the ongoing change most strongly. In tracing the forces that underlie linguistic change, I would follow Meillet in rejecting the reduction of social factors to the social psychology of individuals – the "inter-psychology" invented by Tarde. This approach continues the program advanced by Weinreich, Labov, and Herzog 1968, centered about the concept that the speech community and not the idiolect is the primary object of linguistic investigation. It is true enough that when we examine a community closely enough, it will inevitably appear that each individual's linguistic pattern differs in some respects from that of everyone else. Yet this unique object, the individual speaker, can only be understood as the product of a unique social history, and the intersection of the linguistic patterns of all the social groups and categories that define that individual. Linguistic analysis cannot recognize individual grammars or phonologies. Individual rules or constraints would have no interpretation and contribute nothing to acts of communication. In this sense, the individual does not exist as a linguistic object. However, each individual shows a personal profile of the comparative use of resources made available by the speech community.

Those who work outside of sociolinguistic principles of accountability must hope that the intuitions of several individuals will be sufficiently representative of the speech community to make the description of the language a valid one – a situation that is rarely realized. It is for this reason that all sociolinguists agree that the productions and interpretations of the individual speaker are the primary site for linguistic investigation. The position of this study is that these individuals are not the final units of linguistic analysis, but the components that are used to construct models of our primary object of interest, the speech community.