Just as change sometimes begins with a small number of words, effects of a change often appear first in the speech of a small number of people. Social pressures play an important role in determining whether a particular innovation will spread through the entire linguistic community. Since speakers can consciously or unconsciously alter the way they speak to approximate to what they perceive to be a more prestigious or socially acceptable variety of speech, once a change has taken hold in the speech of a high prestige group it may gradually spread to other speakers and ultimately affect the entire linguistic community.

There have been numerous examples of this in the history of English, notably the loss of postvocalic [r] in a syllable coda in most parts of England and along the east coast of the United States. This change, which resulted in an '*r*-less' pronunciation of words such as *far* as [fɑ:], originated in the south-east of England in the seventeenth and eighteenth centuries. At that time, postvocalic [r] was still pronounced throughout English-speaking settlements in North America. Two factors accounted for its loss in parts of that continent. First, the children of the New England gentry picked up the new pronunciation in British schools and subsequently brought it back to the colony. Second, the speech of newly arrived immigrants, including colonial administrators and church officials who enjoyed high social status in the colony, typically lacked the postvocalic [r]. As a result, the innovation was widely imitated and ultimately spread along much of the east coast and into the south. Dialects where postvocalic *r* is retained are referred to as *rhotic*, and those where it was lost as *non-rhotic*.

Social pressures were also involved in limiting the spread of this innovation. It did not penetrate Pennsylvania or the other Midland states since the most prestigious group of settlers there were Quakers from northern England, an area which at that time retained the postvocalic [r]. Similarly, in Canada, the influence of Scottish and Irish settlers, whose dialects did not undergo the change in question, helped ensure the survival of postvocalic [r] in all but a few areas where contact with New England was strongest, most notably in the Canadian regions of Lunenberg County in Nova Scotia and Grand Manan Island in New Brunswick. More recently (probably in part because of the decline of British prestige) the '*r*-less' pronunciation has become stigmatized in North America, even in areas where it was previously firmly entrenched, and we now see a trend to restoration of [r] in environments where it had been deleted. However, in England no such reversal in the fortunes of '*r-less*' pronunciation has occurred.

7 LANGUAGE RECONSTRUCTION

When we compare the vocabulary items of various languages, we cannot help but notice the strong resemblance certain words bear to each other. By systematically comparing languages, we can establish whether two or more languages descended from a common parent and are therefore genetically related (see Chapter 9). The **comparative method** refers to the procedure

used to reconstruct earlier forms on the basis of a comparison of later forms. In this way, we can reconstruct properties of the parent language with a great degree of certainty.

7.1 Comparative reconstruction

The most reliable sign of family relationships is the existence of **systematic phonetic correspondences** in the vocabulary items of different languages. Many such correspondences can be found, and Table 8.57 shows a sample of vocabulary items from English, Dutch, German, Danish, and Swedish, all of which are members of the Germanic family of languages.

Table 8.57 Some Germanic cognates

English	*Dutch*	*German*	*Danish*	*Swedish*
man	man	Mann	mand	man
hand	hand	Hand	hånd	hand
foot	voet	Fuß (ß = [s])	fod	fot
bring	brengen	bringen	bringe	bringa
summer	zomer	Sommer	sommer	sommar

Since the relationship between the phonological form and meaning of a word is mostly arbitrary, the existence of systematic phonetic correspondences in the forms of two or more languages must point towards a common source. Conversely, where languages are not related, their vocabulary items fail to show systematic similarities. This can be seen by comparing words from Turkish (Table 8.58), which is not related to the Germanic languages, with their counterparts in the languages cited in Table 8.57.

Table 8.58 Some words in Turkish, a non-Germanic language

adam	'man'
el	'hand'
ajak	'foot'
getir	'bring'
jaz	'summer'

Words that have descended from a common source (as shown by systematic phonetic correspondences and, usually, semantic similarities) are called **cognates**. Cognates are not always as obvious as the Germanic examples in Table 8.57. Where languages from the same family are only distantly related, the systematic correspondences may be considerably less striking. This is exemplified in Table 8.59, showing data from English, Russian, and Hindi, all of which are distantly related to each other. Forms from the unrelated Turkish are included to emphasize the similarities among the first three languages.

Table 8.59 Some distinctly related cognates compared to non-related Turkish

English	*Russian*	*Hindi*	*Turkish*
two	dva	dō	iki
three	tri	tīn	yt∫
brother	brat	bhāī	karde∫
nose	nos	nahi̇̃	burun

Once the existence of a relationship between two or more languages has been established, an attempt can be made to reconstruct the common source. This reconstructed language, or **proto-language**, is made up of **proto-forms**, which are written with a preceding * (for example, **hand*) to indicate their hypothetical character as reconstructions of earlier forms that have not been recorded or are not directly observable.

7.2

Techniques of reconstruction

Reconstruction can be undertaken with some confidence because (as discussed in the previous sections) the processes underlying language change are systematic. Once the processes are uncovered by linguists, they can be reversed allowing us to infer earlier forms of the language. Although it is possible to reconstruct all components of a proto-language (its phonology, morphology, syntax, lexicon, and semantics), we will focus in the following on phonological reconstruction, the area in which linguists have made the most progress.

Reconstruction strategies

Reconstruction of a proto-form makes use of two general strategies. The most important one is the **phonetic plausibility strategy**, which requires that any changes posited to account for differences between the proto-forms and later forms must be phonetically plausible. Secondarily, the **majority rules strategy** stipulates that if no phonetically plausible change can account for the observed differences, then the segment found in the majority of cognates should be assumed. It is important to note that the first strategy always takes precedence over the second; the second strategy is a last resort.

Consider the following cognates (somewhat simplified) from members of the Romance family.

Table 8.60 Romance cognates

French	*Italian*	*Rumanian*	*Spanish*	
si	si	∫i	si	'yes'

The data exemplify a correspondence between [s] and [∫] before the vowel [i]. To account for this, we could assume either that Rumanian underwent a

7.3

Internal reconstruction

Sometimes it is possible to reconstruct the earlier form of a language even without reference to comparative data. This technique, known as **internal reconstruction**, relies on the analysis of morphophonemic variation within a single language. The key point is that the sound changes that create allomorphic and allophonic variation can be identified and then used to infer an earlier form of the morpheme. The data in Table 8.72 are from French; because of borrowing, English exhibits a parallel set of contrasts involving [k] and [ʃ].

Table 8.72 [k]–[s] correspondence in French

maʒik	'magic'	maʒis-jɛ̃	'magician'
lɔʒik	'logic'	lɔʒis-jɛ̃	'logician'
myzik	'music'	myzis-jyɛ̃	'musician'

The root morpheme in each row exhibits two forms, one ending in [k], the other ending in [s]. The same methods and principles used in comparative reconstruction can be applied here to reconstruct the historically earlier form of the root morpheme. If a root ending in **s* is posited, no phonetically plausible change can account for the [k] in the left-hand column. By contrast, if a root-final **k* is posited, the [s] can be accounted for by assuming that the **k* was fronted under the influence of the high front vowel of the suffix (palatalization) and became an affricate [tˢ] (affrication), which was later simplified to a fricative [s] (deaffrication). All of these changes are phonetically plausible and listed in the catalogue in Table 8.3. Accordingly, internal reconstruction indicates that at an earlier point in the development of French, the root morphemes in Table 8.72 ended in the consonant **k*.

7.4

The discovery of Indo-European

The late eighteenth-century discovery that Sanskrit (the ancient language of India) was related to Latin, Greek, Germanic, and Celtic revolutionized European linguistic studies. Sir William Jones, a British judge and scholar working in India, summed up the nature and implications of the findings in his 1786 address to the Royal Asiatic Society in Calcutta, a part of which follows:

> The Sanskrit language, whatever be its antiquity, is of a wonderful structure; more perfect than the Greek, more copious [having more cases] than the Latin, and more exquisitely refined than either, yet bearing to both of them a stronger affinity, both in the roots of the verbs and in the forms of the grammar, than could possibly have been produced by accident; so strong indeed, that no philologer could examine them all three, without believing them to have sprung from some common source, which, perhaps, no longer exists; there is a similar reason ... for supposing that both the Gothic and the Celtic ... had the same origin with the Sanskrit; and the old Persian might be added to the same family.

This discovery led to several decades of intensive historical-comparative work and to important advances in historical linguistics during the nineteenth

century. By studying phonetic correspondences from an ever-increasing number of languages, linguists eventually ascertained that most of the languages of Europe, Persia (Iran), and the northern part of India belong to a single family, now called Indo-European. By applying the techniques of the comparative method, they began reconstructing the grammar of the proto-language from which these languages evolved, **Proto-Indo-European (PIE)**.

A number of individuals advanced this research. In 1814, the Danish linguist Rasmus Rask carefully documented the relationships among cognates in a number of Indo-European languages, and at the same time established the methods that would govern the emerging science of historical-comparative linguistics. He wrote:

> When agreement is found in [the most essential] words in two languages, and so frequently that rules may be drawn up for the shift in letters [sounds] from one to the other, then there is a fundamental relationship between the two languages; especially when similarities in the inflectional system and in the general make-up of the languages correspond with them.

Rask worked without access to Sanskrit. The first comparative linguistic analysis of Sanskrit, Greek, Persian, and the Germanic languages was done by the German scholar Franz Bopp in 1816. In 1822, another German, Jakob Grimm, extended Rask's observations and became the first person to explain the relationships among the cognates noted by Rask in terms of a **sound shift**, the systematic modification of a series of phonemes. Some of the correspondences on which he based his work are given in Table 8.73.

Table 8.73 Some Indo-European phonetic correspondences

Greek	*Latin*	*English*
patḗr	pater	father
treîs	trēs	three
hekatón	centum	hundred

The crucial observation is that where English has [f], [θ], and [h] (here, in word-initial position), Greek and Latin have [p], [t], and [k]. Grimm tabulated a series of consonant shifts for Proto-Germanic which differentiated it from other Indo-European languages. **Grimm's Law** (Table 8.74) is the name given to the consonant shifts which took place between Proto-Indo-European and Proto-Germanic.

Table 8.74 The sound shifts underlying Grimm's Law

Proto-Indo-European	p	t	k	b	d	g	bh	dh	gh
Germanic	f	θ	x	p	t	k	b	d	g

Some additional examples of the relationships captured by these shifts are given in Table 8.75. The Proto-Indo-European consonants were either

maintained in Sanskrit, Greek, and Latin or in some cases underwent changes different from those found in Germanic.

Table 8.75 Some examples of the consonant shifts underlying Grimm's Law

Shift in Germanic	*Sanskrit*	*Greek*	*Latin*	*English*
p > f	pād-	pod-	ped-	foot
t > θ	tanu-	tanaós	tenuis	thin
k > x	çatam	hekatón	centum	hundred
b > p	—	—	lūbricus	slippery
d > t	daça	déka	decem	ten
g > k	ajras	agrós	ager	acre
bh > b	bhrātā	phrā́tēr	frāter	brother
dh > d	vidhavā	ēítheos	vidua	widow
gh > g	hansas	khḗn	(h)ānser	goose

Although there appeared to be exceptions to Grimm's Law (Table 8.76), they turned out to be systematic and could be traced to specific environments. For example, voiceless stops were not subject to Grimm's Law when they were immediately preceded by *s*.

Table 8.76 A systematic exception to Grimm's Law

Original s + voiceless stop				
Latin	spu-ō	English	spit	[sp] (not [sf])
Latin	stāre	English	stand	[st] (not [sθ])
Latin	piscis	Gothic	fisks	[sk] (not [sx])

A particularly important discovery was made by Karl Verner, who traced a group of exceptions to Grimm's Law to the original accentual pattern of Proto-Indo-European. In a generalization that came to be known as Verner's Law, he proposed that a word-internal voiceless fricative resulting from Grimm's Law underwent voicing if the original Proto-Indo-European accent did not immediately precede it. Since stress came to be fixed on the root syllable in Germanic subsequent to the changes covered by Verner's Law, the original environment was obscured. However, Sanskrit provides direct evidence for Verner's claim since Sanskrit was very conservative in its maintenance of the original Proto-Indo-European accent. Although the English forms are complicated by other developments, the effects of Verner's Law are apparent in the Gothic examples shown in Table 8.77 (overleaf). In the Gothic word for *brother*, PIE **t* becomes [θ] according to Grimm's Law, whereas in the word for *father* it becomes [ð] in accordance with both Grimm's and Verner's Laws.

Table 8.77 Verner's Law

PIE	*Sanskrit*	*Grimm's Law*	*Verner's Law*	*Gothic*	
*t	bhrā́tā	*t > θ	—	[bro:θar]	'brother'
*t	pitā́	*t > θ	θ > ð	[faðar]	'father'

It should also be noted here that borrowing is an important factor which must be taken into consideration when comparative reconstruction is being carried out. For example, Table 8.78 shows that English has many words that do not show the effects of Grimm's Law.

Table 8.78 English words not showing the effects of Grimm's Law

Expected by Grimm's Law	*Latin*	*English*
p > f	ped-	pedestrian
t > θ	tenuis	tenuous
k > h	canalis	canal

The apparent failure of Grimm's Law here stems from the fact that the English words were borrowed directly from Latin or French many centuries after the sound shifts described by Grimm's Law had taken place. The task of reconstruction can often be complicated by such borrowings.

Subsequent developments

By the middle of the nineteenth century, the study of language had made great strides, especially in the field of phonetics, which opened the way for the detailed comparison of linguistic forms. One influential hypothesis at that time was that sound laws operated without exception. A group of linguists known as the **Neogrammarians** adopted this idea and made many important contributions to the fledgling science of linguistics by applying it to new and more complicated data. Although such factors as lexical diffusion and social pressures were more or less ignored by the Neogrammarians, their hypothesis represented an important and daring advance in the scientific study of language.

The nineteenth century also saw major advances in the classification of languages. A German scholar, August Schleicher, developed a classification for the Indo-European languages in the form of a genealogical tree. Genetic classification is discussed in much more detail in the chapter on language typology that follows.

Work in comparative reconstruction is far from finished. In particular, linguists are now considering the possibility of superfamilies. One such proposed family is Nostratic, which includes Indo-European, Afro-Asiatic (for example, Arabic, Hebrew), Altaic (for example, Japanese, Korean, Turkish), and Uralic (for example, Finnish, Hungarian). Comparative reconstruction is also playing an important role in determining the genetic relationships of the hundreds of North American indigenous languages, a topic which still remains highly controversial.