American Anthropologist 68/4 (1966)

Dialect, Language, Nation[1]

EINAR HAUGEN
Harvard University

The impossibility of stating precisely how many "languages" or "dialects" are spoken in the world is due to the ambiguities of meaning present in these terms, which is shown to stem from the original use of "dialect" to refer to the literary dialects of ancient Greece. In most usages the term "language" is superordinate to "dialect," but the nature of this relationship may be either linguistic or social, the latter problem falling in the province of sociolinguistics. It is shown how the development of a vernacular, popularly called a dialect, into a language is intimately related to the development of writing and the growth of nationalism. This process is shown to involve the selection, codification, acceptance, and elaboration of a linguistic norm.

THE taxonomy of linguistic description—that is, the identification and enumeration of languages—is greatly hampered by the ambiguities and obscurities attaching to the terms "language" and "dialect." Laymen naturally assume that these terms, which are both popular and scientific in their use, refer to actual entities that are clearly distinguishable and therefore enumerable. A typical question asked of the linguist is: "How many languages are there in the world?" Or: "How many dialects are there in this country?"

The simple truth is that there is no answer to these questions, or at least none that will stand up to closer scrutiny. Aside from the fact that a great many, perhaps most, languages and dialects have not yet been adequately studied and described, it is inherent in the very terms themselves that no answer can be given. They represent a simple dichotomy in a situation that is almost infinitely complex. Hence they have come to be used to distinguish phenomena in several different dimensions, with resultant confusion and overlapping. The use of these terms has imposed a division in what is often a continuum, giving what appears to be a neat opposition when in fact the edges are extremely ragged and uncertain. Do Americans and Englishmen speak dialects of English, or do only Americans speak dialect, or is American perhaps a separate language? Linguists do not hesitate to refer to the French language as a dialect of Romance. This kind of overlapping is uncomfortable, but most linguists have accepted it as a practical device, while recognizing, with Bloomfield, "the purely relative nature of the distinction" (1933:54).

The two terms are best understood against the perspective of their history. In English both words are borrowed from French. *Language* is the older, having partially displaced such native words as "tongue" and "speech" already in Middle English. The oldest attestation in the OED is from 1290: "With men þat onder-stoden hire langage." The French word is itself late, being a popular derivative of Latin *lingua* with the probable form **linguāticum*, first attested in the 12th century. *Dialect*, on the other hand, first appears in the Renaissance, as a learned loan from Greek. The oldest OED citation is from 1579 in reference to "certain Hebrue dialectes," while the earliest French I have found (in Hatzfeld and Darmesteter's dictionary) is only 16 years earlier and speaks

of Greek as being "abondante en dialectes." A 1614 citation from Sir Walter Raleigh's *The History of the World* refers to the "Aeolic Dialect" and confirms the impression that the linguistic situation in ancient Greece was both the model and the stimulus for the use of the term in modern writing.

There was need for some such term in Greece, since there was in the classical period no unified Greek norm, only a group of closely related norms. While these "dialects" bore the names of various Greek regions, they were not spoken but written varieties of Greek, each one specialized for certain literary uses, e.g., Ionic for history, Doric for the choral lyric, and Attic for tragedy. In this period the language called "Greek" was therefore a group of distinct, but related written norms known as "dialects." It is usually assumed that the written dialects were ultimately based on spoken dialects of the regions whose names they bore. These spoken dialects were in turn descended by normal linguistic divergence from a Common Greek language of an older period, which can be reconstructed by comparison of the dialects with each other and with their Indo-European kinsmen. In the postclassical period, however, the Greek dialects disappeared and were replaced by a rather well-unified Greek norm, the *koinḗ*, essentially the dialect of Athens. So, in the Hellenistic period "Greek" became the name of a norm that resulted from a linguistic convergence. The differences among the dialects were eliminated in favor of a single, triumphant language, based on the dialect of the cultural and administrative center of the Greeks.

The Greek situation has provided the model for all later usage of the two terms "language" and "dialect." Much of the unclarity in their application stems from the ambiguities present in that situation. This has become evident with their extension to other countries and with their adoption into the technical terminology of linguistics. In a descriptive, synchronic sense "language can refer either to a *single* linguistic norm, or to a *group* of related norms. In a historical, diachronic sense "language" can either be a common language on its way to dissolution, or a common language resulting from unification. A "dialect" is then any one of the related norms comprised under the general name "language," historically the result of either divergence or convergence.

Since this historical process can be indefinitely repeated, the two terms are cyclically applicable, with "language" always the superordinate and "dialect" the subordinate term. This is also clear from the kind of formal structures into which they can be placed: "X is a dialect of language Y," or "Y has the dialects X and Z" (never, for example, "Y is a language of dialect X"). "Language" as the superordinate term can be used without reference to dialects, but "dialect" is meaningless unless it is implied that there are other dialects and a language to which they can be said to "belong." Hence every dialect is a language, but not every language is a dialect.

In addition to the ambiguities provided by the synchronic and diachronic points of view distinguished above, increasing knowledge concerning linguistic behavior has made the simple application of these two contrasting terms ever more difficult.

In French usage a third term developed, *patois*, which applied primarily to the spoken language. The term *dialecte* is defined in the dictionary of the Académie française and other French dictionaries as "variété régionale d'une langue." Littré (1956) explicitly requires that a dialect "include a complete literary culture" (comportant une complète culture littéraire). As pointed out by André Martinet (1964), this usage reflects the special French situation, in which there were a number of regional written standards, which were then superseded by the written standard of Paris. The French dialects were regional, like the Greek, and literary, but not functionally distinguished like the Greek. When the dialects ceased to be written, they became "patois": "Après le XIV^{e} siècle, il se forma une langue littéraire et écrite, et les dialectes devinrent des patois" (Littré). Even more succinctly, Brun (1946) writes: "Un patois est un dialecte qui s'est degradé." A patois, then, is a language norm not used for literary (and hence official) purposes, chiefly limited to informal situations. Thus Provençal might be considered a French dialect, but its local, spoken varieties are all patois. This distinction introduces a new dimension in our discussion: the social functions of a language. In terms of the language-dialect distinction, we may say that a patois is a dialect that serves a population in its least prestigious functions. The distinction of patois-dialect is therefore not one between two kinds of language, but between two functions of language. The definition in Littré (and others like it) clearly suggests a pejorative attitude toward the patois, since it no longer carries with it "a complete literary culture."

In English the term "patois" has never been seriously adopted in the description of language, and "dialect" has carried the full burden of both scientific and popular usage. Older writers, cited in the OED, often used it for any specialized variety of the language, e.g., "the lawyer's dialect." Samuel Butler (*Hudibras*, 1663) railed against "a Babylonish dialect, which learned pedants much affect." General usage has limited the word largely to the regional or locally based varieties, such as "Lancashire dialect" or "Irish dialect" in reference to varieties of English. It is less customary to speak of "London dialect" or "Boston dialect," except in reference to the lower-class speech of those cities. Nor is it common to speak of "British dialect" in reference to cultivated English speech, and Americans are generally resentful of being told they speak "American dialect" when reference is had to the speech of educated people. Martinet is therefore beside the mark when he writes that in America "the term denotes every local form of English but without any suggestion that a more acceptable form of the language exists distinct from the dialects" (1964:146). It is quite different with the word "accent": an American may inoffensively be described as having a "New England accent" or a "Southern accent," and, of course, all Americans speak of the English as having an "English accent." "Dialect" is here as elsewhere a term that suggests informal or lower-class or rural speech. In general usage it therefore remains quite undefined whether such dialects are part of the "language" or not. In fact, the dialect is often thought of as standing outside the language: "That

isn't English." This results from the *de facto* development of a standard language, with all the segregation of an élite and the pyramidal power structure that it has usually implied.

As a social norm, then, a dialect is a language that is excluded from polite society. It is, as Auguste Brun (1946) has pointed out, a language that "did not succeed." In Italy, Piedmontese is from every linguistic point of view a language, distinct from Italian on the one hand and French on the other, with a long tradition of writing and grammatical study. But because it is not Tuscan, and Tuscan became the standard language of all Italy, Piedmontese is only a "dialect," yielding ground to Italian with every generation and kept alive only by local pride and linguistic inertia (Clivio 1964). Only if a "dialect" is watered down to an "accent"—that is, an intonation and a set of articulations, with an occasional lexical item thrown in for color—does it (say in Germany or Italy or England) become "salonfähig." As a complete structure it is out in the cold limbo of modern society. In America the stigma is placed not so much on local dialects, since these are few and rarely heard, as on "bad" English, which is quite simply lower-class dialect. The language of the upper classes is automatically established as the correct form of expression. They cannot say only, "L'état, c'est moi," but also "Le langage, c'est le mien."

In trying to clarify these relationships, linguistic science has been only moderately successful. Even in the Renaissance it was perfectly clear to serious students of the subject that the term "language" was associated with the rise of a nation to conscious unity and identity. George Puttenham wrote in his book *The Arte of English Poesie* (1589): "After a speach is fully fashioned to the common understanding, and accepted by consent of a whole country and nation, it is called a language." This kind of historical development, by which convergence was achieved at the expense of deviating varieties, was familiar to the men of that age. But the arbitrary tower-of-Babel approach to linguistic divergence was dispelled by the discovery, in the early 19th century, of historical regularity. The realization that languages have resulted from dialect-splitting gave a new content to the terms and made it possible to begin calling languages like English and German "dialects" of a Germanic "language."

But in the mid-19th century, when scientific study of the rural and socially disadvantaged dialects began, a generation of research was sufficient to revolutionize the whole idea of how a dialect arises. The very notion of an area divided into a given number of dialects, one neatly distinct from the next, had to be abandoned. The idea that languages split like branches on a tree gave way to an entirely different and even incompatible idea, namely, that individual linguistic traits diffused through social space and formed isoglosses that rarely coincided. Instead of a dialect, one had a "Kernlandschaft" with ragged edges, where bundles of isoglosses testified that some slight barrier had been interposed to free communication. Linguistics is still saddled with these irreconcilable "particle" and "wave" theories; this in effect involves the differing points of view from which any linguistic structure can be seen: as a unitary

structure (a "language"), or as one of several partially overlapping structures (the "dialects").

Without going into the problems raised by this conflict, we may simply state that the "particle" theory of language as a unified structure is a fruitful hypothesis, making it possible to produce an exhaustive and self-consistent description. But it excludes as "free variation" a great many inconsistencies within the speech of any informant, and it fails to account for the fact that communication is possible between users of identifiably different codes. Comparative grammar succeeded in reconstructing the common structure from which "dialects" could be derived. Contrastive grammar has tried to program the differences between languages in order to ease the learner's task or, on a higher theoretical plane, to arrive at a linguistic typology. But there is still no calculus that permits us to describe the differences between languages in a coherent and theoretically valid way.

Our discussion has shown that there are two clearly distinct dimensions involved in the various usages of "language" and "dialect." One of these is *structural*, that is, descriptive of the language itself; the other is *functional*, that is, descriptive of its social uses in communication. Since the study of linguistic structure is regarded by linguists as their central task, it remains for sociologists, or more specifically, sociolinguists, to devote themselves to the study of the functional problem.

In the *structural* use of "language" and "dialect," the overriding consideration is genetic relationship. If a linguist says that Ntongo has five dialects, he means that there are five identifiably different speech-forms that have enough demonstrable cognates to make it certain that they have all developed from one earlier speech-form. He may also be referring to the fact that these are mutually understandable, or at least that each dialect is understandable to its immediate neighbors. If not, he may call them different languages, and say that there is a language Ntongo with three dialects and another, Mbongo, with two. Ntongo and Mbongo may then be dialects of Ngkongo, a common ancestor. This introduces the synchronic dimension of comprehension, which is at best an extremely uncertain criterion. The linguist may attempt to predict, on the basis of his study of their grammars, that they should or should not be comprehensible. But only by testing the reactions of the speakers themselves and their interactions can he confirm his prediction (Voegelin and Harris 1951; Hickerson *et al.*, 1952). Between total incomprehension and total comprehension there is a large twilight zone of partial comprehension in which something occurs that we may call "semicommunication."

In the *functional* use of "language" and "dialect," the overriding consideration is the uses the speakers make of the codes they master. If a sociolinguist says that there is no Ntongo language, only dialects, he may mean that there is no present-day form of these dialects that has validity beyond its local speech community, either as a trade language or as a common denominator in interaction among the various dialect speakers. A "language" is thus functionally

defined as a superposed norm used by speakers whose first and ordinary language may be different. A "language" is the medium of communication between speakers of different dialects. This holds only within the limits established by their linguistic cognacy: one could not speak of Ntongo as a dialect of English just because its speakers use English as a medium of intercommunication. The sociolinguist may also be referring to the fact that the "language" is more prestigious than the "dialect." Because of its wider functions it is likely to be embraced with a reverence, a language loyalty, that the dialects do not enjoy. Hence the possibility of saying that "Mbongo is only a dialect, while Ngkongo is a language." This means that Ngkongo is being spoken by people whose social prestige is notoriously higher than that of people who speak Mbongo. When used in this sense, a dialect may be defined as an undeveloped (or underdeveloped) language. It is a language that no one has taken the trouble to develop into what is often referred to as a "standard language." This dimension of functional superiority and inferiority is usually disregarded by linguists, but it is an essential part of the sociolinguist's concern. It becomes his special and complex task to define the social functions of each language or dialect and the prestige that attaches to each of these.

What is meant by an "undeveloped" language? Only that it has not been employed in all the functions that a language can perform in a society larger than that of the local tribe or peasant village. The history of languages demonstrates convincingly that there is no such thing as an inherently handicapped language. All the great languages of today were once undeveloped. Rather than speak of undeveloped languages as "dialects," after the popular fashion, it would be better to call them "vernaculars," or some such term, and limit "dialect" to the linguist's meaning of a "cognate variety." We are then ready to ask how a vernacular, an "undeveloped language," develops into a standard, a "developed language." To understand this we will have to consider the relation of language to the nation.

The ancient Greeks and Romans spread their languages as far as their domains extended, and modern imperialists have sought to do the same. But within the modern world, technological and political revolutions have brought Everyman the opportunity to participate in political decisions to his own advantage. The invention of printing, the rise of industry, and the spread of popular education have brought into being the modern nation-state, which extends some of the loyalties of the family and the neighborhood or the clan to the whole state. Nation and language have become inextricably intertwined. Every self-respecting nation has to have a language. Not just a medium of communication, a "vernacular" or a "dialect," but a fully developed language. Anything less marks it as underdeveloped.

The definition of a nation is a problem for historians and other social scientists; we may accept the idea that it is the effective unit of international political action, as reflected in the organization of the United Nations General Assembly. As a political unit it will presumably be more effective if it is also a social unit. Like any unit, it minimizes internal differences and maximizes

external ones. On the individual's personal and local identity it superimposes a national one by identifying his ego with that of all others within the nation and separating it from that of all others outside the nation. In a society that is essentially familial or tribal or regional it stimulates a loyalty beyond the primary groups, but discourages any conflicting loyalty to other nations. The ideal is: internal cohesion—external distinction.

Since the encouragement of such loyalty requires free and rather intense communication within the nation, the national ideal demands that there be a single linguistic code by means of which this communication can take place. It is characteristic that the French revolutionaries passed a resolution condemning the dialects as a remnant of feudal society. The dialects, at least if they threaten to become languages, are potentially disruptive forces in a unified nation: they appeal to local loyalties, which could conceivably come into conflict with national loyalty. This is presumably the reason that France even now refuses to count the number of Breton speakers in her census, let alone face the much greater problem of counting the speakers of Provençal. On the other hand, a nation feels handicapped if it is required to make use of more than one language for official purposes, as is the case in Switzerland, Belgium, Yugoslavia, Canada, and many other countries. Internal conflict is inevitable unless the country is loosely federated and the language borders are stable, as is the case in Switzerland.

Nationalism has also tended to encourage external distinction, as noted above. In language this has meant the urge not only to have one language, but to have one's own language. This automatically secludes the population from other populations, who might otherwise undermine its loyalty. Here the urge for separatism has come into sharp conflict with the urge for international contact and for the advantages accruing both to individual and nation from such contact. Switzerland is extreme in having three languages, no one of which is its own; Belgium has two, both of which belong to its neighbors. The Irish movement has faltered largely under the impact of the overwhelming strength of English as a language of international contact. The weakness of the New Norwegian language movement is due to the thorough embedding of Danish in the national life during four centuries of union; what strength the movement has had is derived from the fact that Danish was not one of the great international languages.

Whenever any important segment of the population, an élite, is familiar with the language of another nation, it is tempting to make use of this as the medium of government, simply as a matter of convenience. If this is also the language of most of the people, as was the case when the United States broke away from England, the problem is easily solved; at most it involves the question of whether provincialisms are to be recognized as acceptable. But where it is not, there is the necessity of linguistically re-educating a population, with all the effort and disruption of cultural unity that this entails. This is the problem faced by many of the emerging African and Asian nations today (Le Page 1964). French and English have overwhelming advantages, but they sym-

bolize past oppression and convey an alien culture. The cost of re-education is not just the expense in terms of dollars and cents, but the malaise of training one's children in a medium that is not their own, and of alienation from one's own past.

The alternative is to develop one's own language, as Finland did in the 19th century, or Israel did in the 20th. Different languages start at different points: Finland's was an unwritten vernacular, Israel's an unspoken standard. Today both are standards capable of conveying every concept of modern learning and every subtlety of modern literature. Whatever they may lack is being supplied by deliberate planning, which in modern states is often an important part of the development process.

It is a significant and probably crucial requirement for a standard language that it be written. This is not to say that languages need to be written in order to spread widely or be the medium of great empires. Indo-European is an example of the first, Quechua of the Inca Empire an example of the second (Buck 1916). But they could not, like written languages, establish models across time and space, and they were subject to regular and inexorable linguistic change. It is often held that written language impedes the "natural" development of spoken language, but this is still a matter of discussion (Zengel 1962; Bright and Ramanujan 1964). In any case the two varieties must not be confused.

Speech is basic in learning language. The spoken language is acquired by nearly all its users before they can possibly read or write. Its form is to a great extent transmitted from one generation of children to the next. While basic habits can be modified, they are not easily overturned after childhood and are virtually immovable after puberty. The spoken language is conveyed by mouth and ear and mobilizes the entire personality in immediate interaction with one's environment. Writing is conveyed by hand and eye, mobilizes the personality less completely, and provides for only a delayed response. Oral confrontation is of basic importance in all societies, but in a complex, literate society it is overlaid and supplemented by the role of writing.

The permanence and power of writing is such that in some societies the written standard has been influential in shaping new standards of speech. This is not to say that writing has always brought them into being, but rather to say that new norms have arisen that are an amalgamation of speech and writing. This can of course take place only when the writing is read aloud, so that it acquires an oral component (Wessén 1937). There is some analogy between the rise of such spoken standards and that of pidgin or creole languages (Meillet 1925:76; Sommerfelt 1938:44). The latter comprise elements of the structure and vocabulary of two or more languages, all oral. They have usually a low social value, compared to the oral standards, but the process of origin is comparable. The reawakening of Hebrew from its century-long dormant state is comprehensible only in terms of the existence of rabbinical traditions of reading scripture aloud (Morag 1959). Modern Hebrew has shown a rapid adaptation to the underlying norms of its new native speakers, so that it has

become something different from traditional Hebrew. Similarly with the standard forms of European languages: one is often hard put to say whether a given form has been handed down from its ancestor by word of mouth or via the printed page. "Spelling pronunciations" are a well-known part of most oral standards, even though purists tend to decry them.

While we have so far spoken of standard languages as if they were a clear and unambiguous category, there are differences of degree even among the well-established languages. French is probably the most highly standardized of European languages, more so than, for example, English or German. French, as the most immediate heir of Latin, took over many of its concepts of correctness and its intellectual elaboration. French in turn became a model for other standard languages, and its users were for centuries nothing loth to have it so considered. When English writers of the 18th century debated whether an English academy should be established to regulate the language, the idea of such an institution came from France. The proposal was rejected largely because the English did not wish to duplicate what they regarded as French "tyranny."

In France, as in other countries, the process of standardization was intimately tied to the history of the nation itself. As the people developed a sense of cohesion around a common government, their language became a vehicle and a symbol of their unity. The process is reasonably well documented in the histories written for the older European languages. But the period since the French Revolution has seen a veritable language explosion, which has been far less adequately studied. In many countries a process that elsewhere took centuries of effort on the part of a people and its writers has been compressed into a few short years or decades. In a study of the new standards developed since 1800 for Germanic languages, Heinz Kloss has suggested that there may be a typical profile for what he has called the "Ausbau" of a new language (Kloss 1952:28). First comes its use for purely humorous or folkloristic purposes. Then lyric writers may adopt it, followed by prose narrators. But it has not reached a crucial stage of development until success is achieved in writing serious expository prose, or what he calls "Zweckschrifttum." Beyond this comes the elaboration of the language for purposes of technical and scientific writing and government use. Each of these "domains" (as Fishman [1964] has called them) constitutes a challenge for the language in its attempt to achieve full development.

While making a survey of the world's standard languages, Ferguson proposed (1962) to classify them along two dimensions: their degree of standardization (St. 0, 1, 2) and their utilization in writing (W 0, 1, 2, 3). Zero meant in each case no appreciable standardization or writing. St. 1 meant that a language was standardized in more than one mode, as is the case, for example, with Armenian, Greek, Serbo-Croatian, and Hindi-Urdu. He also included Norwegian, but it is at least arguable that we are here dealing with two languages. St. 2 he defined as a language having a "single, widely accepted norm which is

felt to be appropriate with only minor modifications or variations for all purposes for which the language is used." W 1 he applied to a language used for "normal written purposes," W 2 to one used for "original research in physical sciences," and W 3 to one used for "translations and résumés of scientific work in other languages."

These categories suggest the path that "underdeveloped" languages must take to become adequate instruments for a modern nation. The "standardization" to which Ferguson refers applies primarily to developing the form of a language, i.e., its linguistic structure, including phonology, grammar, and lexicon. We shall call this the problem of *codification*. Ferguson's scale of "utilization in writing" applies rather to the *functions* of a language. We shall call this the problem of *elaboration*, a term suggested by a similar usage of Bernstein's (1962) and corresponding to Kloss's *Ausbau*. As the ideal goals of a standard language, codification may be defined as *minimal variation in form*, elaboration as *maximal variation in function*.

The ideal case of minimal variation in form would be a hypothetical, "pure" variety of a language having only one spelling and one pronunciation for every word, one word for every meaning, and one grammatical framework for all utterances. For purposes of efficient communication this is obviously the ideal code. If speakers and listeners have identical codes, no problems of misunderstanding can arise due to differences in language. There can be none of what communication engineers call "code noise" in the channel (Hockett 1958:331–332). This condition is best attained if the language has a high degree of stability, a quality emphasized by many writers on the subject (e.g., Havránek 1938). Stability means the slowing down or complete stoppage of linguistic change. It means the fixation forever (or for as long as possible) of a uniform norm. In practice such fixation has proved to be chimerical, since even the most stable of norms inevitably changes as generations come and go. At all times the standard is threatened by the existence of rival norms, the so-called "dialects," among its users. It is liable to interference from them and eventually to complete fragmentation by them.

Apparently opposed to the strict codification of form stands the maximal variation or elaboration of function one expects from a fully developed language. Since it is by definition the common language of a social group more complex and inclusive than those using vernaculars, its functional domains must also be complex. It must answer to the needs of a variety of communities, classes, occupations, and interest groups. It must meet the basic test of *adequacy*. Any vernacular is presumably adequate at a given moment for the needs of the group that uses it. But for the needs of the much larger society of the nation it is not adequate, and it becomes necessary to supplement its resources to make it into a language. Every vernacular can at the very least add words borrowed from other languages, but usually possesses devices for making new words from its own resources as well. Writing, which provides for the virtually unlimited storage and distribution of vocabulary, is the technological

device enabling a modern standard language to meet the needs of every specialty devised by its users. There are no limits to the elaboration of language except those set by the ingenuity of man.

While form and function may generally be distinguished as we have just done, there is one area in which they overlap. Elaboration of function may lead to complexity of form, and, contrariwise, unity of form may lead to rigidity of function. This area of interaction between form and function is the domain of *style*. A codification may be so rigid as to prevent the use of a language for other than formal purposes. Sanskrit had to yield to Prakrit, and Latin to the Romance languages, when the gap between written and spoken language became so large that only a very few people were willing to make the effort of learning them. Instead of being appropriate for "all purposes for which the language is used," the standard tends to become only one of several styles within a speech community. This can lead to what Ferguson (1959) has described as "diglossia," a sharp cleavage between "high" and "low" style. Or it may be a continuum, with only a mild degree of what I have called "schizoglossia," as in the case of English (Haugen 1962). In English there is a marked difference between the written and spoken standards of most people. In addition, there are styles within each, according to the situation. These styles, which could be called "functional dialects," provide wealth and diversity within a language and ensure that the stability or rigidity of the norm will have an element of elasticity as well. A complete language has its formal and informal styles, its regional accents, and its class or occupational jargons, which do not destroy its unity so long as they are clearly diversified in function and show a reasonable degree of solidarity with one another.

Neither codification nor elaboration is likely to proceed very far unless the community can agree on the *selection* of some kind of a model from which the norm can be derived. Where a new norm is to be established, the problem will be as complex as the sociolinguistic structure of the people involved. There will be little difficulty where everyone speaks virtually alike, a situation rarely found. Elsewhere it may be necessary to make some embarrassing decisions. To choose any one vernacular as a norm means to favor the group of people speaking that variety. It gives them prestige as norm-bearers and a headstart in the race for power and position. If a recognized élite already exists with a characteristic vernacular, its norm will almost inevitably prevail. But where there are socially coordinate groups of people within the community, usually distributed regionally or tribally, the choice of any one will meet with resistance from the rest. This resistance is likely to be the stronger the greater the language distance within the group. It may often be a question of solidarity versus alienation: a group that feels intense solidarity is willing to overcome great linguistic differences, while one that does not is alienated by relatively small differences. Where transitions are gradual, it may be possible to find a central dialect that mediates between extremes, one that will be the easiest to learn and most conducive to group coherence.

Where this is impossible, it may be necessary to resort to the construction of a new standard. To some extent this has happened naturally in the rise of

the traditional norms; it has been the aim of many language reformers to duplicate the effect in new ones. For related dialects one can apply principles of linguistic reconstruction to make a hypothetical mother tongue for them all. Or one can be guided by some actual or supposed mother tongue, which exists in older, traditional writings. Or one can combine those forms that have the widest usage, in the hope that they will most easily win general acceptance. These three procedures—the comparative, the archaizing, and the statistical —may easily clash, to make decisions difficult. In countries where there are actually different languages, amounting in some African nations to more than a hundred, it will be necessary either to recognize multiple norms or to introduce an alien norm, which will usually be an international language like English or French.

Finally, a standard language, if it is not to be dismissed as dead, must have a body of users. *Acceptance* of the norm, even by a small but influential group, is part of the life of the language. Any learning requires the expenditure of time and effort, and it must somehow contribute to the well-being of the learners if they are not to shirk their lessons. A standard language that is the instrument of an authority, such as a government, can offer its users material rewards in the form of power and position. One that is the instrument of a religious fellowship, such as a church, can also offer its users rewards in the hereafter. National languages have offered membership in the nation, an identity that gives one entrée into a new kind of group, which is not just kinship, or government, or religion, but a novel and peculiarly modern brew of all three. The kind of significance attributed to language in this context has little to do with its value as an instrument of thought or persuasion. It is primarily symbolic, a matter of the prestige (or lack of it) that attaches to specific forms or varieties of language by virtue of identifying the social status of their users (Labov 1964). Mastery of the standard language will naturally have a higher value if it admits one to the councils of the mighty. If it does not, the inducement to learn it, except perhaps passively, may be very low; if social status is fixed by other criteria, it is conceivable that centuries could pass without a population's adopting it (Gumperz 1962, 1964). But in our industrialized and democratic age there are obvious reasons for the rapid spread of standard languages and for their importance in the school systems of every nation.

The four aspects of language development that we have now isolated as crucial features in taking the step from "dialect" to "language," from vernacular to standard, are as follows: (1) selection of norm, (2) codification of form, (3) elaboration of function, and (4) acceptance by the community. The first two refer primarily to the form, the last two to the function of language. The first and the last are concerned with society, the second and third with language. They form a matrix within which it should be possible to discuss all the major problems of language and dialect in the life of a nation:

	Form	*Function*
Society	Selection	Acceptance
Language	Codification	Elaboration

NOTES

[1] This paper was written as a contribution to the work of the Seminar on Sociolinguistics, held at the Indiana University Linguistic Institute in the summer of 1964, under the direction of Charles A. Ferguson. It has profited from extensive discussion with the members of the Seminar.

REFERENCES CITED

BERNSTEIN, BASIL
1962 Linguistic codes, hesitation phenomena and intelligence. Language and Speech 5:31–46.

BLOOMFIELD, LEONARD
1933 Language. New York, Holt.

BRIGHT, WILLIAM, AND A. K. RAMANUJAN
1964 Sociolinguistic variation and language change. Proceedings, IX International Congress of Linguists. The Hague, Mouton. Pp. 1107–1114.

BRUN, AUGUSTE
1946 Parlers régionaux: France dialectale et unité française. Paris, Toulouse: Didier.

BUCK, C. D.
1916 Language and the sentiment of nationality. American Political Science Review 10:44–69.

CLIVIO, GIANRENZO
1964 Piedmontese: a short basic course. Mimeographed ed. Center for Applied Linguistics, Washington, D. C.

DICTIONNAIRE DE L'ACADÉMIE FRANÇAISE
1932 Paris, 8th ed.

FERGUSON, CHARLES A.
1959 Diglossia. Word 15:325–340.
1962 The language factor in national development. Anthropological Linguistics 4, 1:23–27.

FISHMAN, JOSHUA
1964 Language maintenance–language shift as a field of inquiry. Linguistics 9:32–70.

GUMPERZ, JOHN
1962 Types of linguistic communities. Anthropological Linguistics 4, 1:28–40.
1964 Hindi-Punjabi code switching in Delhi. Proceedings, IX International Congress of Linguists. The Hague, Mouton. Pp. 1115–1124.

HATZFELD, A., AND A. DARMESTETER
1920 Dictionnaire général de la langue française. Paris, 6. ed.

HAUGEN, EINAR
1962 Schizoglossia and the linguistic norm. Monograph Series on Languages and Linguistics. Georgetown University, Washington, D. C. No. 15, 63–69.

HAVRÁNEK, BOHUMIL
1938 Zum Problem der Norm in der heutigen Sprachwissenschaft und Sprachkultur. *Repr. in* A Prague school reader in linguistics. J. Vachek, ed. Bloomington, Ind. Pp. 413–420.

HICKERSON, H., G. D. TURNER, AND N. P. HICKERSON
1952 Testing procedures for estimating transfer of information among Iroquois dialects and languages. IJAL 18:1–8.

HOCKETT, CHARLES F.
1958 A course in modern linguistics. New York, Macmillan.

KLOSS, HEINZ
1952 Die Entwicklung neuer germanischer Kultursprachen von 1800 bis 1950. Munich, Pohl.

LABOV, WILLIAM
1964 Phonological correlates of social stratification. *In* The ethnography of communication. John J. Gumperz and Dell Hymes, eds. American Anthropologist 66, no. 6, pt. 2:164–176.

LE PAGE, R. B.
1964 The national language question: linguistic problems of newly independent states. London, Oxford University Press.

LITTRÉ, E.
1956 Dictionnaire de la langue française. Paris.

MARTINET, ANDRÉ
1964 Elements of general linguistics. Chicago, University of Chicago Press.

MEILLET, ANTOINE
1925 La méthode comparative en linguistique historique. Oslo, Institutet for sammenlignende kulturforskning.

MORAG, SHELOMO
1959 Planned and unplanned development in modern Hebrew. Lingua 8:247–263.

OXFORD ENGLISH DICTIONARY
1888 ff. James A. H. Murray *et al.*, eds. Oxford University Press.

SOMMERFELT, ALF
1938 Conditions de la formation d'une language commune. Actes du IV Congrès international de linguistes. Copenhagen. Pp. 42–48.

VOEGELIN, C. F., and Z. S. HARRIS
1951 Methods for determining intelligibility among dialects of natural languages. Proceedings of the American Philosophical Society 95:322–329.

WESSÉN, ELIAS
1937 Vårt rikksspråk: Några huvudpunkter av dess historiska utveckling. Modersmålslärarnas Förenings årsskrift. Pp. 289–305.

ZENGEL, MARJORIE S.
1962 Literacy as a factor in language change. American Anthropologist 64:132–139.