STUDIES OF NORTH AMERICAN INDIAN LANGUAGES

Marianne Mithun

Department of Linguistics, University of California, Santa Barbara, California 93106

KEY WORDS: linguistics, discourse, language contact, genetic relationships, sociolinguistics

The study of North American Indian languages has been shaped by several circumstances. Since its beginning, most research has been based in fieldwork: Data have come from direct contact with speakers, usually in their own cultural settings, rather than from secondary sources. Although the number of languages indigenous to North America is large, several hundred, the number of scholars working with most of them has been relatively small, often only one or two individuals per language. A center of scholarly interaction has thus been the community of those studying languages all over the Americas, languages that are quite diverse genetically and typologically. These factors, the grounding in fieldwork and the composition of the scholarly community, have affected the kinds of work undertaken, the theoretical issues addressed, the nature of the explanations sought, and the applications made. The field is in many ways highly anthropological: strongly contextualized linguistically, culturally, and typologically.

The researcher who works directly with speakers of a little-documented language needs basic proficiency in all areas of linguistic structure. It is seldom possible to achieve insights about syntax, for example, without sufficient phonetic skill to hear accurately, sufficient understanding of phonology to construct a usable transcription, and sufficient knowledge of morphology to be aware of the grammatical functions signaled within words. The importance of broad competence has shaped the kinds of description and explanation that have come out of the field. Structures are usually considered in the context of other structures in the language rather than in isolation.
Few linguists working with North American Indian languages can ignore their rich cultural settings or, in many cases, the rapid changes they are undergoing. Good linguistic work in such settings has seldom been limited to the elicitation of grammatical paradigms or sentences. The rich but often fragile cultural settings of North American Indian languages have prompted a tradition of collecting texts of all kinds: religious or political oratory; legends; historical accounts; reminiscences; children’s stories; descriptions of ceremonies such as naming, marriage, burial, selection and installation of leaders, etc; various aspects of daily life, such as hunting, fishing, cooking, medicine, basketmaking, games, songs, etc; and now, with the availability of tape recorders, conversation as well. Such documentation has been valuable in itself, in many cases providing the only descriptions in the speakers’ own words of earlier events and customs that are now fading from common memory. It has also had an important effect on the study of language and its use.

The textual material provides linguistic and cultural context for the study of grammar. If the investigation of grammatical structures were limited to data consisting of isolated sentences translated from a contact language like English, much would be missed. Translations can easily distort the grammatical patterns of a language in ways that obscure their actual functions: Too often, aspects of the language under investigation are understood as perfect counterparts to their translations when in fact they are parts of very different systems. Much can also be lost when sentences are examined in isolation: The use of many constructions is governed by factors beyond the limits of a single sentence. Reliance on elicitation alone has a further drawback: If all analysis were based on elicited data, the analyst might never notice structures that he or she had not anticipated in advance. When grammatical structures are studied in the context of naturally occurring, connected speech, their precise functions, and their differences from similar structures in other languages, can be detected in ways that are often not possible when data are limited to isolated elicited translations of sentences.

Linguistic typology and American Indian languages have long enjoyed a special relationship. Since von Humboldt first brought glimpses of New World languages to European scholars, the exotic character of their structures has aroused interest. The work of Boas, Sapir, Bloomfield, and others assured the role of North American languages in the development of linguistic theory in this century, a role that continues, partly because these languages contain structures not often found in other parts of the world. At the same time, exposure to other languages has played a significant role in work with North American languages. The more experience one has with the structures of a variety of languages, the more effectively one can perceive patterns in an undescribed language. The major journal for those working with American
Indian languages, the *International Journal of American Linguistics*, and the primary scholarly meetings, the SSILA (Society for the Study of the Languages of the Americas) sessions at the annual meeting of the American Anthropological Association and elsewhere, consist of work on all of the languages indigenous to the Americas. Any North Americanist who reads *IJAL* and attends one or two SSILA meetings per year, as most do, is constantly exposed to languages genetically and typologically quite diverse. This exposure has been a constant source of fresh ideas.

Current work with North American Indian languages addresses a broad range of topics. Since few if any of the languages have been fully documented, and the opportunities for recording many of them will soon disappear, description continues to be a priority. Grammars, dictionaries, and collections of texts are being produced with increasing sophistication. More areas of linguistic structure are being investigated and understood in terms of more factors, including discourse patterns, social and geographic variation, and diachrony. In the following sections, a few of the main currents of recent work on the languages north of Mexico will be sketched.

### STRUCTURAL CONTEXTUALIZATION

Much current work on grammar, both morphology and syntax, demonstrates the importance of linguistic context beyond the bounds of the sentence. One area of interest to linguists working within most theoretical models is reference. North American Indian languages have much to contribute to theories of reference because their repertoires of referential devices are not always isomorphic with those of more familiar languages.

#### Referential Alternatives

Much general theoretical work on reference has concentrated on coreference relations among nominals (full noun phrases, pronouns, etc) within sentences in Indo-European languages. Recent work on reference in a number of North American Indian languages has shown that referential systems can be much more varied and complex than might be supposed on the basis of Indo-European alone. Furthermore, most can be fully understood only when their use is examined in larger stretches of natural speech.

Watkins (92) shows that the relative frequencies of full noun phrases and zero anaphora in Kiowa differ radically across different kinds of texts. In stories centered around a single protagonist, this character is typically named only once early in the narrative, then subsequently not mentioned at all (zero anaphora). Such texts may contain as many as 20 clauses without a single full noun phrase. In stories with several characters of comparable importance,
who interact significantly, each may be identified by a full noun phrase in nearly every clause, emphasizing the opposition.

Describing Central Pomo narrative and conversation, Mithun (69) shows that alternations among several referential devices are determined by the status of the referent within the consciousness of the speaker. Full noun phrases introduce new referents into consciousness. Pronouns derived from demonstratives bring referents back into consciousness after any discontinuity in discussion. Referents already within consciousness are not overtly reidentified at all (zero anaphora). A special set of empathetic pronouns provides another kind of distinction, encoding referents whose point of view is represented.

A related language, Northern Pomo, contains regular third person pronouns, in addition to empathetic pronouns, demonstratives, and the option of zero anaphora. Masculine and feminine gender, and singular and plural number are distinguished by the pronouns. O'Connor (73), examining conversation, demonstrates how the alternation between these personal pronouns and the demonstratives reflects not only a human/nonhuman distinction, but also the social context and the speaker's attitude toward third persons under discussion, and interacts with the indication of evidence for the information presented.

Kinkade (49) discusses the relation between discourse topicality and agency in a number of Salishan languages. In all of the languages, transitive verbs may be passivized if their patients are more topicworthy within the discourse than their agents. At least 6 of the 23 Salishan languages contain an additional device; a special topical object marker sets off grammatical objects of special importance within a stretch of discourse containing multiple third persons. Kinkade notes that the neighboring but unrelated Sahaptin also has such a marker, which may be the source of the Salishan suffixes. Rude (79) describes a similar phenomenon in Klamath.

Goddard (29) details the function of the proximate-obviative distinction in Fox narrative: “The proximate is the unmarked third person category; if there is only one third person referent in a context, it can only be proximate. . . . In contexts that have a third person animate (noun or pronoun) and, in addition, another third person, the higher-status or more central third person is proximate and the lower-status or less central third person is obviative.” Goddard explores the points at which the proximate category shifts to new referents in discourse. While proximate or obviative status is determined in certain contexts by syntactic factors, the alternation most often functions as a discourse device, “a highly significant aspect of the structuring and resulting narrative texture of discourse.”

All of these referential distinctions—full noun phrases/zero, pronouns/zero, demonstratives/pronouns, empathetic/personal pronouns, topical ob-
jects/zero, and proximate/obviative—clearly depend on factors beyond the scope of the sentence. Most involve topicality or point of view in some sense, but they differ in important ways that are only beginning to be understood.

The Forms of Referential Devices and Their Structural Ramifications

Of course all languages contain full noun phrases, although the same kinds of distinctions are not always encoded, such as definiteness or number. All languages also contain pronouns, but their forms and uses vary in interesting ways from language to language, ways that can have wider ramifications in the syntax and discourse.

Pronouns in some languages are much like English pronouns: separate words that alternate with full noun phrases to refer to identifiable persons, objects, etc. In many North American Indian languages, however, regular pronouns are verbal affixes. They appear with every verb whether an additional noun phrase is present in the sentence or not. The presence of a full set of pronominal affixes, in three persons, is often correlated with certain other grammatical characteristics (43, 64). The pronominal affixes function as the primary arguments of clauses, so any coreferent noun phrases in the clause typically have a somewhat loose syntactic connection with the verb. Accordingly, the relative order of noun phrases and verbs in many of these languages may reflect not syntactic relations but rather the pragmatic status of the information they represent within the discourse (65, 76, 86). Information that is new and/or important often occurs early in the clause, with more predictable and incidental information appearing later. Noun phrases themselves may not have internal constituency like that in European languages (66). Subordination may not be as strongly grammaticized (44, 62). Languages of this type have sometimes been referred to as “nonconfigurational” or “pronominal argument” languages.

Discourse pragmatic factors in the word orders of other kinds of languages have also received attention. Eastman (16), Eastman & Edwards (17), Edwards (18, 19), and Enrico (20) discuss the roles of focus and topic in Haida, a basically verb-final language. Chafe (13) points out the role of idiomaticity of Subject-Verb and Verb-Object complexes in word order in Caddo.

The Roles of Referents: Voice

Another area in which a discourse perspective has proven important is that of voice. As long as voice alternations are compared only in pairs of sentences out of context, it may be possible to isolate their forms but difficult to determine their full functions. Several studies have demonstrated the importance of examining voice phenomena in discourse.

Rude (80) investigates the discourse-functional contexts that trigger al-
ternations among antipassive, ergative (transitive), and passive constructions in Nez Perce. He concludes that the alternation is best understood in terms of the comparative topicworthiness of agents and patients. "In the antipassive, the agent far outweighs the patient in topicality, in the passive the agent is completely suppressed, and in the ergative construction the patient is clearly a secondary; in terms of cataphoric continuity it equals the agent."

Whisder (93) examines a construction sometimes labeled "passive" in Nootkan languages and concludes that the phenomenon is better described as inverse marking, whose function is best understood in terms of the thematic organization of texts at the "paragraph level." "Constituent clauses of a paragraph are marked direct or inverse depending on the ACTOR versus GOAL status of a thematic participant."

Thompson (89) shows that the inverse construction in Koyukon Athabaskan (the famous yi/bi alternation) differs significantly from passive and impersonal constructions in that the agent is seldom suppressed, and when it is, this is because it is anaphoric. The inverse should be viewed functionally as "a marker of a deviation from the normal topicality relations."

The research described here constitutes only a small portion of work currently being carried out on the structures of North American Indian languages. It is, however, representative of the current interest in considering grammatical structures in their larger context of natural connected speech of a variety of styles.

SOCIOCULTURAL CONTEXTUALIZATION

The fact that much work on North American Indian languages has been based on texts (narrative or conversation) recorded in their cultural settings has meant that most studies of the forms of the languages are not isolated from the communicative and social functions they serve. Much work has taken into account the variety of registers available to speakers, as well as the social contexts in which they are spoken: geographical, social, and temporal.

Registers

Linguistic structures vary according to the cultural functions they are called upon to serve. Some recent work on such stylistic alternations, or registers, has examined the special speech forms used with children, in ritual, and in writing, among others.

In many cultures, adults speak differently to young children and sometimes to pets than to other adults. The speech forms themselves, as well as speakers' attitudes toward them, are interesting both culturally and linguistically. In some cultures, they are treasured as a way of showing affection or considered an aid for children in learning their first language; in others, they are dis-
dained as an impediment to children’s proper acquisition of the language. Speakers generally consider such forms to be simpler, but ideas of simplicity are not necessarily uniform from one group to the next. Many North American Indian languages provide an especially good basis for study of such notions of simplicity owing to their relatively complex phonetic, phonological, or morphological structures.

Earlier in this century, Sapir described such special forms in Nootka, a Wakashan language of British Columbia (81, 82). More recently, Kess & Kess (45) have reexamined this speech register in another Nootkan dialect, Ahousaht. Ahousaht has a large inventory of consonants, including glottalized series of obstruents and resonants (p’, t’, c’, č’, etc. and m’, n’, y’, w’, etc.), distinctive velar (k, k’, x), labiovelar (kʷ, kʷ’, xʷ), uvular (q, q’, č), labio-uvular (qʷ, qʷ’, čʷ), pharyngeal (ţ, h) and laryngeal (ʔ, h) series. As might be expected, some of the articulatorily more complex sounds do not appear in baby talk (motherese): q’, qʷ’, č, ş, sounds that are in fact somewhat infrequent in adult speech as well. Consonant clusters do not appear in baby talk, although they do in normal adult speech. There is morphological simplification in the lack of imperative suffixes, but as in much baby talk, reduplication and diminutive suffixes are common. Thompson (90) has examined the ways speakers of Twana, a Salish language of Western Washington, address young children. Characteristics of this register are compared to the Twana women’s speech register.

A quite different kind of register, highly developed in many North American Indian languages, is that of ritual language. Du Bois (15) has surveyed the linguistic forms of such speech that serve to convey authority, to make the ritual utterance appear self-evident. He points out that the ritual register is often characterized by a marked voice quality and by stylized intonation patterns, often a much more restricted set than those found in colloquial language. They are sometimes described as “singsong” or “chanted.” Ritual speech is typically more fluent than colloquial speech, without the false starts, afterthoughts, or other disfluencies characteristic of everyday spoken language. The vocabulary often contains many archaic, borrowed, or tabooed terms, as well as euphemistic or metaphorical circumlocutory expressions. Certain structural patterns are common. Many ritual languages are characterized by couplet structure—pairs of lines that are semantically and/or grammatically parallel. Evidential markers, lexical and grammatical markers of the source and reliability of the information conveyed, often play a strong role in ritual speech.

Still another special register may be found in written language. Most of the languages indigenous to North America do not have lengthy literary traditions, although some traditions are well known, including those of the Massachusetts, the Micmac, and the Cherokee. Massachusetts literacy, which
spanned more than a century during colonial times, is beautifully chronicled in Goddard & Bragdon's (30) two-volume work, which contains all known documents written by the Massachusetts, with word-by-word translation and extensive notes. Micmac writing is traced from aboriginal ideographic origins to a current roman-based system by Battiste (2). An entire volume of articles on native writing systems in Canada has been assembled by Burnaby (7).

As more North American Indian languages are written by their speakers, it is possible to observe the effect of the written medium on the register itself. Several papers have investigated the special characteristics of the new written medium. Mithun (63) discussed the development of a written Mohawk style. When Mohawk speakers first began to produce written Mohawk texts, their work clearly showed the influence of English literary style, a style familiar to all of these writers. Spoken Mohawk consists largely of verbs, but early written texts contained a noun/verb ratio similar to that of English. Word order in spoken Mohawk is purely pragmatically determined, but early written texts exhibit general SVO order. An unusually high proportion of certain particles appeared, on the model of English definite articles and conjunctions. There was little noun incorporation. As writers became more experienced, however, they were able to take advantage of the luxury of time to produce a uniquely rich Mohawk style, characterized by effective use of elaborate morphological constructions, including stylistic word order and noun incorporation.

Watkins (91) compared four Kiowa styles: (a) informal spoken (spontaneous anecdotes), (b) planned spoken (formal stories), (c) informal written (letters), and (d) formal written (written stories). Four structures were examined: right dislocation, incorporation, relativization, and subordination. She discovered that right dislocation appears only in spoken texts, representing given material. The other three structures appear in both spoken and written styles, but their frequency of occurrence increases significantly from type a to type d, with the time available for planning.

Rhetorical Structure

An area that has aroused special interest in recent years has been rhetorical structure. It has become clear that the presentation of oral texts in simple paragraph form can obscure much of their inherent internal structure. Two kinds of patterning have been explored in particular, that signaled by linguistic form and that marked by intonation. Hymes (37), working with Chinookan and other North American Indian texts, many transcribed before the advent of tape recorders, has argued for the presentation of textual material according to lexical markers, syntactic structure, and content. Chafe (10) and others have looked at the textual structure expressed intonationally. The division of texts into intonation units separated by pauses of various kinds, and the kinds of
prosodic contours (involving pitch and rhythm) associated with these intonation units, reveal much about the systematic ways information is presented.

In some cases, the textual organization signaled by lexical and grammatical cues coincides with that expressed by intonation and prosody. In others, the two approaches reveal different kinds of organization. These kinds of rhetorical structure and their relationships in North American texts have been discussed in a number of works. Woodbury (96) investigates rhetorical structure in Central Alaskan Yupik discourse, “prosodically and intonally signaled phonological phrasing along with whatever other significant formal features consistently pattern or interact with it.” In addition to their basic structural functions, he describes four additional communicative functions associated with these rhetorical devices: organization of information, expression of affective meaning, indexing of genre, and regulation of dialogic interaction. Other work on rhetorical structure in North American native discourse includes, among others: Bright on Karok (4, 6), Chafe on Seneca (11), Goddard on Unami (25), Golla on Hupa (31), Kinkade on Upper Chehalis (47), Kroskrity (52) on Tewa, McLendon on Eastern Pomo (58, 59), Rood on Wichita (78), Sherzer on Kuna (83, 84), Tedlock on Zuni (88), and especially Woodbury on Central Alaskan Yupik Eskimo (95, 97). This work has had far-reaching consequences not only for the visual presentation of oral discourse but also for our understanding of linguistic structure.

**Language Contact**

Investigation of the linguistic effects of contact among speakers of different languages has always been a part of the study of North American Indian languages. North America contains some notorious linguistic areas, geographic regions in which unrelated languages share sets of traits, such as Northern California. The distribution of such traits continues to be investigated, because it will have much to tell us about the kinds of diffusion that are possible, and in many cases constitutes an important piece of deep genetic puzzles. Among such areal studies are a discussion of the distribution of sibilants in aboriginal California by Bright (3) and of dental and alveolar apicals in California languages by Langdon & Silver (56).

Because of the intense contact situations all over North America, the effects of encroaching European languages on the indigenous languages continue to be documented. Effects of French can be seen in languages of the Northeast, such as Mohawk, and the influence of Russian can be seen not only in Alaska but as far south as Kashaya in California (74). The impact of English is of course pervasive. The influence of Spanish continues to be documented for languages all over the Southwest and California. [See, among others, Kroskrity & Reinhardt (54) and Gamble (22).] Loanwords indicate the nature of
contact, such as names of introduced cultural items including tools, domestic animals, foods, clothing, etc. The shapes of these words in the modern languages also indicate that the borrowing was not necessarily directly from Europeans themselves, but often through other indigenous languages (5).

When contact becomes particularly intense, and most speakers of the indigenous language become bilingual, the original language may be used in fewer and fewer contexts until its sphere of use all but disappears. Some effects of such intensive contact have been described especially insightfully by Hill & Hill (36). They point out that bilingual situations can reflect tremendously creative management of the double resource of two linguistic systems, in this case a Uto-Aztecan language and Spanish.

If speakers no longer use a language on a regular basis, how accurately does their speech represent the full original system? Cool (14) notes that two of the most conspicuous characteristics of dying languages are (a) structural and stylistic simplifications and (b) dramatic increases in variability. He describes variations resulting from simplifications, including syllable reduction, phonemic mergers, and reduction of allomorphy, in two Athabaskan languages, Chipewyan and Sarcee. He proposes viewing these changes as the result of impeded language acquisition, since none of the semispeakers ever reaches the highest level of proficiency in the traditional grammatical system.

Moore (71) suggests that contemporary Wasco speakers view language primarily as a collection of words, whose use is primarily display. Investigating Gros Ventre (Atsina), a dialect of Arapaho (Algonquian), Taylor (87) discovered that as the language was replaced by English, analogical leveling regularized singular and plural forms of some nouns and verbs. Comparing two dialects of Cayuga, one still spoken by a substantial community, the other seldom used at all, Mithun (67) noted that in the little-used dialect, a complex phonological phenomenon involving the spreading of laryngeal features over odd-numbered syllables had been simplified. As with any language, much vocabulary had been forgotten. The productivity of derivational morphology had also diminished, including the capacity to form new verbs with incorporated nouns. Overall morphological complexity was reduced, so that individual affixes might appear with particular words, but complex combinations of them were rare.

Comparing the speech of speakers of Central Pomo with varying degrees of fluency, Mithun (68) discusses the effect of language obsolescence on linguistic description. While those who have not used their first language for a number of years may forget considerable vocabulary and exhibit a reduced range of stylistic alternatives, they can still be of importance to investigators. They can set the stage for more fluent speakers to exhibit their own virtuosity and provide lexicalized expressions, which themselves may reflect complex derivational structures.
DIACHRONY

Partly because of the large number of languages indigenous to North America, many only recently documented, diachronic work continues to be of great interest to North American Indian linguists. There is perhaps less separation here than in some language areas between those doing diachronic and synchronic work. Most scholars currently engaged in reconstruction, subgrouping, and the search for deeper genetic relations, have done extensive field work themselves with many or all of the languages of at least one family, usually more. Work in each domain influences the other. Particular diachronic problems are less likely to be considered in a structural vacuum; at the same time, a diachronic perspective pervades the description and explanation of synchronic patterns.

Over a decade ago, scholars working with most of the major language families in North America met to share notes on the current state of knowledge in their fields (9). Since that time, diachronic work of all kinds has continued at an energetic rate. As more and better data have become available, our understanding of the interrelationships among languages and of their earlier states has continued to grow. A number of cooperative projects are in place, for families such as Siouan, Uto-Aztecan, and Yuman, in which specialists in related languages are pooling their data in large computer databases.

Processes of Language Change and Reconstruction

Study of language change on all levels continues. Much of this work, like the synchronic work described earlier, is highly contextualized: The development of structures is traced in the context of their relationships with other structures in the languages and their social and communicative functions. One example of such work is Goddard's discussion of the Eastern Algonquian subordinative mode (24). Mithun (70) traces the simultaneous grammaticization of conjunction in seven Northern Iroquoian languages under the influence of English and French. Kroskrity (51) examines the discourse pragmatic factors influencing syntax in his discussion of negation and subordination in Arizona Tewa. Kroeber (50) considers discourse and functional factors in the development of ergative case marking from passive constructions in Southern Interior Salish languages. Goddard (27) shows that apparent rapid change in Fox during this century can be traced to the interaction of two styles: deliberate and casual.

As more lexical reconstruction becomes possible, so does linguistic paleontology, the reconstruction of words for cultural elements. Fowler (21) suggests an original location of the Uto-Aztecan homeland stretching across modern Arizona, down into Mexico, and perhaps into Southern California.
This was followed by a north-south split. There occurred a gradual deeper penetration of some of the southern languages into central Mexico, northward and eastward expansion of the Numic languages, and movements into adjacent deserts by Takic and Pimic groups. Mithun (61) reconstructs Proto-Iroquoian terms for flora, fauna, hunting, agricultural, and aquatic complexes, other aspects of material culture, and some social traditions. Kinkade (48) investigates the origin of the Salish homeland. The Salish have occupied two markedly different ecological zones, the temperate North Pacific coast from northern Oregon to central British Columbia, and the arid interior plateau. Boas had suggested that these people originally inhabited the interior and then migrated to the coast; more recently, Suttles suggested a coastal origin. Kinkade’s investigation of terms for flora and fauna supports the coastal origin, except for Bella Coola, which contains evidence of an interior origin. He points out that this conclusion corresponds with that reached by archaeologists, an agreement he realized only after completing the work.

Subgrouping

As reconstruction has progressed, progress has been made in subgrouping as well, determining how languages within families are related to each other. Recent work on subgrouping includes among others that on Yokuts (23, 94), on Uto-Aztecan (60), on Muskogean (46, 72), on Siouan (75), and on Algonquian (77).

Deeper Genetic Relations

Demonstrable genetic relationships among families on the level of Indo-European have essentially been established in North America, although work continues to refine reconstruction and subgrouping. The same scholars involved in synchronic and family-level diachronic work continue to investigate deeper genetic relationships among families. Several of the groupings originally hypothesized by Kroeber and Sapir on the basis of superficial structural similarities continue to be explored as more data become available for comparison. Langdon (55) and Jacobsen (39) assess various aspects of the Hokan hypothesis. More recently, ambitious unpublished work by T. Kaufman provides mounting lexical evidence of relationships among some of the language families grouped as Hokan, but suggests that others, such as Chumash, do not belong with this group. Silverstein (85) assesses the status of the Penutian hypothesis. Investigation continues in this area as well in largely unpublished work by S. DeLancey, V. Golla, and K. Whistler, among others.

In 1987, J. Greenberg published a book, Language in the Americas, in which he proposed that all of the indigenous languages of the Americas fall into only three genetic groups: Eskimo-Aleut, Na-Dene, and what he calls “Amerind.” His work has not generally met with success among specialists in
the field; see discussions by Adelaar (1), Campbell (8), Chafe (12), Goddard (26, 28), Golla (32), Jacobsen (40–42), Matisoff (57), and a reply by Greenberg (35). Both his methodology and its application have been severely criticized.

The initial step in comparative work has always been a survey of languages for obvious similarities. The better one knows the languages, of course, the more similarities one can perceive. When superficial resemblances are detected among a set of languages, one can begin work to determine whether these are due to chance, to borrowing, or to common inheritance. Greenberg’s methodology is essentially the first step without the second. His technique consists of what he terms “multilateral comparison”: “looking at the basic vocabulary and concrete grammatical markers of a large number of languages simultaneously” (34:648).

The methodology presupposes its conclusion: Some genetic relationship falls out of it automatically in any case. Those languages that share the most superficial resemblances are assumed to be the most closely related. In his first chapter, Greenberg states that “Basically, the wrong question has been asked, namely, when are languages genetically related? . . . What should be asked is, how are languages to be classified genetically?” (33:3). He clearly expects that all languages in the world will ultimately be demonstrated to be genetically related by his method. “The ultimate goal is a comprehensive classification of what is very likely a single language family” (33:337).

Greenberg feels that the genius of the method lies in the large numbers of languages considered simultaneously. It is certainly true that more similarities will be found among ten languages considered simultaneously, for example, than between any two. When the number is raised to 2000, all the more similarities will become obvious. He seems to have failed to take into account the mounting role of chance. The more languages considered, the more chance resemblances will appear. Mohawk, for example, contains 9 consonants: \(t, k, s, n, w, r, y, h, ?\). The verb root for ‘eat’ is \(-k-\). One would probably not have to look too far to find another language with a verb containing \(k\), or perhaps a sound somewhat like it such as \(g\) or \(x\), whose meaning is something like ‘consume,’ ‘eat,’ ‘bite,’ ‘chew,’ etc. Unfortunately, the methodology stops short of separating the role of chance from that of genetic relationship in producing resemblances.

It might be hoped that the data presented in support of the proposals could at least be utilized by other scholars to determine their validity. Unfortunately, as most reviewers have pointed out, the data are so riddled with serious problems that they could not provide a reliable basis for any further work. The problems range from low-level clerical errors to the criteria used in determining sound and meaning correspondences.

Basic identifications of languages are notoriously unreliable. In many cases
it was not noticed that different word lists actually represent a single language. Kashaya and Southwest Pomo are in fact the same language, for example, as are Clear Lake and Eastern Pomo. These are simply well-known alternative names for the same speech communities, but they are presented as distinct languages. Campbell (8) lists cases where names of towns or even people are confused with languages. Particular languages are not always given the same affiliation. Catawba, for example, is sometimes grouped with the Iroquoian family (33:179), sometimes with the Siouan family (33:179, which is correct).

Greenberg notes that work as ambitious as his had not been attempted previously owing to the paucity and inferior quality of documentation of many of the languages of the New World. Inferior data present no problem with his method. “The fact is, the method of multilateral comparison is so powerful that it will give reliable results even with the poorest of materials” (33:29). The materials are indeed poor, but they need not have been. His data are often drawn from brief early notes made by explorers passing through an area for the first time, rather than the rich, technically excellent dictionaries and grammars now available.

In an attempt to increase the compatibility of the lists, he has retranscribed them into his own system, apparently without knowledge of the actual sound systems of the languages. Numerous errors have accordingly been introduced at every turn. The retranscription additionally renders it impossible to recover the original sources of the material, none of which are cited because “listing all these sources in a general bibliography would have added greatly to the length and cost of the work” (33:xv).

Identification of comparable forms is a serious problem. Many American languages are polysynthetic: Words are often long, composed of multiple meaningful parts. Accurate comparison depends on the correct identification of the parts, because it is these that will be comparable across diachronically related languages. Compare, for example, the two words below.

**Mohawk**

naienenhstaienthó:ko

‘that corn will be harvested’

**Tuscarora**

yahrdá'nyg‘

‘he sat there’

These are in fact related. They are based on cognate verb roots that have not even changed much in sound or meaning from their original forms, although the spelling conventions are slightly different. Mohawk -ient- and Tuscarora -yg- are both traceable to a Proto-Northern Iroquoian root *-yεga ‘set/lay.’ The Mohawk word means literally ‘that-one-corn-lay-cause-reverse-(punctual)’ = ‘that one would unplant corn.’ The Tuscarora word means literally ‘there-he-self-set-(punctual)’ = ‘there he set himself’. Accurate segmentation is obviously not an easy task with such languages if one is not
familiar with their grammars. Unfortunately, in Greenberg’s work, wrong
cuts are consistently made and the wrong parts of words compared.

Specialists in most language families have been perplexed to find in
Greenberg’s work spurious forms that do not resemble any known forms from
the language they are said to represent, even when miscutting and miscopying
are taken into account. This is unfortunate, in that it contributes to the
problem of an unreliable data base. One scholar remarked that not a single
form attributed by Greenberg to a language he works with is actually from that
language.

Greenberg’s groupings of the forms into “etymologies” are disappointing
for several reasons. It should be noted that in contrast with usual usage, in
which “etymology” refers to the demonstrated historical source of a form,
Greenberg’s “etymologies” are simply sets of forms that appear to him on
inspection to be similar in form and meaning. His notion of similarity is
considerably looser than most. One set of “morphemes” grouped into a single
“Amerind etymology,” for example (33:233-34), are variously glossed
‘mouth,’ ‘ear,’ ‘listen,’ ‘cave,’ ‘hear,’ ‘smell,’ ‘blow nose,’ and ‘sniff.’ The
form they share is some kind of sibilant, variously s, ʃ, ʃ', t', tʃ', č, x, plus
nasal, m, or n, although not all forms contain a sibilant, or a nasal, or either.
The vowel between the sibilant and nasal is variously i, e, ě, a, o, u, or no
vowel at all. Some of the forms contain material before this sequence, some
after, and some both before and after. While shifts in meaning between some
of the glosses are not impossible, and shifts between some pairs of sounds are
also not impossible, it must be admitted that if one were to examine all words
for ‘mouth,’ ‘ear,’ ‘listen,’ ‘cave,’ ‘hear,’ ‘smell,’ ‘blow nose,’ and ‘sniff’ in
2000 unrelated languages, the probability of finding several forms containing
a sibilant and/or nasal is not small. When such wide ranges of sound and
meaning are classified as equivalent, it simply becomes difficult to rule out
the role of chance.

Greenberg’s distribution of forms into “etymologies” shows numerous
inconsistencies. Often words from a single language family, known not to be
cognate, are listed as part of the same etymology. Under ‘sweet,’ for example
(33:178), are listed Seneca xen- and Mohawk oni-ete, supposedly related to
Proto-Keresan *?an’e:za ‘be tasty.’ The Seneca form (actually -en) is the
descendant of Proto-Northern-Iroquoian *-ran-. The Mohawk word ‘nieht’e’
(root *-nyeh- with neuter prefix o-) means ‘snow’. (Entirely different words
are used for ‘sweet’ and ‘sugar’. ) At the same time, forms known to be close
cognates within families are listed under different etymologies. Three differ-
ent “Amerindian etymologies” are given for words meaning ‘man,’ for ex-
ample (33:242-43). Under Man₁ is given, among others, Central Pomo čač
(actually ča:č’), related to Achomawi is, Chimariko ići, itri, Shasta ?is,
Tequistlatec ašans, Yana hisi, Wintun siw-ij, Plains Miwok sawwe, Chiti-
macha ?asi, Tunica ši, Atakapa ša, ši, Huave na-šej, and Quiche Mayan ačij. Under Man2 is Eastern Pomo ka:kʰ, related to Alakaluf hekaje, Aymara hake, Jaqaru haqi, Barbareño Chumash o xoix, Ynezeño uyiyi, Coahuilteco xagü, and Karankawa a hakx. In fact, the Central Pomo ča:čʰ and Eastern Pomo ka:kʰ are well-known, systematically related cognates. Of course there is an easy response. If all languages of the world are assumed to be ultimately related anyway, the different “etymologies” for ‘man’ may themselves be cognate.

Unfortunately, too many of the etymologies Greenberg provides as evidence for wider relationships between stocks actually consist only of cognates from languages whose genetic relationships have already been established. One of the subgroups proposed under “Amerind” is “Almosan-Keresiuan,” consisting of the Caddoan, Iroquoian, Keresan, Siouan-Yuchi, Wakashan, Chemakuan, and Salish families. The relationship of Caddoan, Iroquoian, and Siouan was established long ago and discussed in several works referred to by Greenberg. Yet a large proportion of the etymologies cited as proof of “Almosan-Keresiuan” contain forms from these families alone. These sets of forms actually constitute no more than hypothesized evidence for what is already known. In many cases, the number of forms under a single etymology is also small. Jacobsen (41) points out, for example, that of 281 comparative sets in Greenberg’s Amerind Etymological Dictionary, 107 contain only two members. The average number is 3.5.

Abundant use of the transitivity argument further complicates the weighing of evidence. It is assumed that if a word in language A is related to one in language B, and a word in language B is related to one in language C, then language A is related to language B, B is related to C, and A is related to C. In principle, such reasoning is valid. Transitivity chains can only be as strong as their weakest links, however. If one relationship fails, the whole chain falls. In Greenberg’s work, whole families and stocks are in many cases judged to be related to other stocks on the basis of a few similarities among a few words in a few languages.

In too many cases the particular forms chosen for comparison are disappointing. It is clear from his notebooks, on file at Stanford University, that Greenberg had access to publications containing full cognate sets from many families, complete with reconstructions and detailed listing of the sound changes undergone by each language. Many of his forms were drawn from these sources, although they are not cited in his publication. Yet in many cases he chose idiosyncratic forms that are known to be the result of recent innovations, because they were superficially more similar to other forms outside of the family.

Of course it would be a phenomenal achievement to deal with the number of languages discussed in this work without some inaccuracies in detail.
Greenberg feels that such details are unimportant, because of the sheer mass of the data. Unfortunately, most specialists who have examined the work feel that the mass of the material falls on the wrong side of accuracy. Goddard examined the 142 word sets that include Algonquian forms and discovered that "errors in the Algonquian data alone invalidate 93 of these equations. . . . There remain 49 word sets in which the shape and meaning Greenberg assumes for the Algonquian forms are approximately correct . . . . An evaluation of these proposed equations shows that they do not demonstrate or even suggest genetic relationships for Algonquian and must be considered to be the result of chance, if they require any notice at all" (26:656).

Most Americanists are themselves extremely interested in uncovering deeper genetic relations among the language families of the New World, and hypotheses are an important first step in the process. Unfortunately Greenberg has here presented hypotheses as fact. It might be suggested that the proposals simply need to be disproven. Unfortunately, there is no direct way to demonstrate the lack of a genetic relationship among languages. At best, one can note that there is insufficient evidence to posit a particular relationship. Although specialists are in a position to evaluate the evidence presented here, those in other disciplines, to whom such proposals may be of considerable importance, may not be. Most linguists currently working with North American languages would rather spend their time creatively, exploring actual relationships, than trying to disentangle mistakes. Only too soon, direct study of these languages will no longer be possible, and the number of scholars in a position to evaluate the evidence at all will have diminished drastically as well.

RESPONSIBILITIES

Most linguists working in Native American communities have grappled with the issue of their appropriate roles and responsibilities in these communities. There is great concern in most communities about the possible disappearance of their languages. When a language disappears, some of the most important parts of a culture disappear along with it. Ways of thinking can be lost along with ways of speaking. Often much of the culture is not handed down in the new language. Sometimes traditional legends and stories are not passed on, because they lose so much when rendered in a new language. Traditional etiquette and styles of humor may be lost in translation. Younger generations can be left without the same awareness of or feeling for their cultural heritage that their ancestors had. Such a loss may not have obvious immediate material ramifications, but it can have a profound effect on an individual's sense of self and community.

Decisions about what should be done as languages are threatened with
disappearance belong with the communities themselves, the speakers and their descendants. Opinions vary considerably, even within communities, concerning what actions should be taken to record the languages and to teach them to younger generations. Some people see little point in documenting old ways in a modern world, while others are intently aware of the great loss of knowledge that occurs every time a speaker passes away. Some point out that ability to speak an Indian language is economically useless, or worse, a handicap in an English-speaking world. Many of these people are themselves successful bilinguals, but they remember the misery of arriving at school without English and of being ridiculed and punished for speaking their mother tongue. Others feel strongly that they want to equip younger generations with the sense of self that can come from knowing about their heritage. Some believe that documentation of the language is worthwhile but attempts to teach the languages will be futile.

Education programs have the best success when the impetus for their creation comes from within. Many communities are now managing language programs but calling on linguists for technical support in certain areas. One of these is orthography development: devising a workable practical spelling system for a particular language. Another is the development of language curricula. Mohawk lessons based on the model of French will never make sense, but a linguist should be able to show teachers how to devise programs that will present the language in its own terms. A third is the training of Native teachers. Being able to speak a language is essential to being able to teach it, but it is not sufficient. Most language structures are unconscious to native speakers, especially speakers of languages without literary traditions. Although language teachers need not teach grammar as a subject, an awareness of the structures of their language can allow them to present more coherent classes. Finally, linguists can provide help in developing documentation of the language: dictionaries, texts, and perhaps grammars.

There are now annual meetings of specialists in most of the language families. At most of these, Native language teachers and linguists come together to discuss the issues involved in the maintenance of language programs as well as various aspects of the languages themselves. In recent years, the Athabaskan conference, the Salish conference, the meeting of the Friends of Uto-Aztecan, the Siouan conference, the Iroquois conference, the Algonquian conference, and undoubtedly others have all had sessions specifically devoted to language teaching, and will undoubtedly continue to do so.

CONCLUSION

The languages of North America offer a wealth of examples of the varieties of linguistic form. This variety is delightful in itself, and it is crucial to any
general understanding of the phenomenon of language. North American languages contain some of the most elaborate phonological, morphological, and syntactic structures in the world. Most study of these languages is securely grounded in data, usually gathered directly by the researcher in the field; but it is not necessarily atheoretical, although there is little interest in model building for its own sake. Because of the structural, social, typological, and diachronic contextualization of most research, more kinds of explanations are called into play. Describing such work, Kroskrity (52:325) aptly comments that "this more holistic approach emulates the anthropological ideal of understanding cultural phenomena not only in terms of contemporary structure and function, but also as the products of a unique history involving both continuation of cultural tradition and diffusion from neighboring groups."

Literature Cited

53. Langdon, M. 1979. Some thoughts on Hokan with particular reference to Pomoan and Yuman. See Ref. 9, pp. 592–649
73. O’Connor, M. C. 1990. Third person reference in Northern Pomo conversation: the indexing of discourse genre and social relations. IJAL 56:3
92. Watkins, L. J. 1990. Noun phrase ver-


