

BOLIVIAN INDIAN GRAMMARS: TWO

PUBLICATION NUMBER 16

Benjamin F. Elson

EDITORIAL COMMITTEE

Kenneth L. Pike

Robert E. Longacre

Viola Waterhouse

BOLIVIAN INDIAN GRAMMARS: II

Prepared under the direction

of

ESTHER MATTESON

A Publication of the
SUMMER INSTITUTE OF LINGUISTICS
OF THE
UNIVERSITY OF OKLAHOMA
Norman

© Summer Institute of Linguistics 1967
octubre, 1967 primera edición
Esta edición consta de 550 ejemplares

Derechos Reservados

por el

Instituto Lingüístico de Verano, A.C.
Hidalgo 186, México 22, D.F.
Impreso en México Printed in Mexico
5.5C205 6-109

Introduction

Ten grammars of indigenous Bolivian languages are presented in the two volumes of Bolivian Indian Grammars. The data were gathered and analyzed by members of the Bolivia Branch of the Summer Institute of Linguistics, which has carried on linguistic investigations in this country since 1955.

The corpus on which each grammar has been based was closed far short of an exhaustive treatment. Nevertheless each analysis is based on a minimum of 200 pages of text dictated spontaneously. Gaps in the patterns have been filled in by elicited materials, but no patterns have been established on the basis of elicited materials alone. Most of the linguists have not only gathered far more than the minimum amount of texts, but have also become fluent in speaking the languages, and intimately acquainted with the culture through years of living in the tribal villages.

The grammars are presented in the tagmemic model¹ of Pike with adaptation of the matrices² and syntax paradigms³ developed by him, as well as the basic concepts which characterize the tagmemic model: (1) the postulation of three hierarchies, the phonological, grammatical, and lexical or semantic; (2) the segmentation of the construction occurring on a given level into a string⁴ of tagmemes manifested by lower-level constructions or units; (3) the definition of the tagmeme as a class-slot correlative with an obligatory versus optional (+ or ±) characteristic; (4) the assumption that units and patterns exist in a language; (5) the stipulation that a well-defined unit be described in terms of identificational-contrastive features, variants, and distribution; and (6) the establishment of emic (versus etic) units.

¹ Pike, Kenneth L. 1954, 1955, 1960. *Language in relation to a unified theory of the structure of human behaviour*, Part I, II, III. Glendale, Summer Institute of Linguistics.

² Pike, Kenneth L. 1962. "Dimensions of Grammatical Constructions," *Language*, Vol. 38:3, part 1, July-September.

³ Pike, Kenneth L. 1963. "A Syntactic Paradigm," *Language*, Vol. 39, No. 2, April-June.

⁴ Longacre, Robert E. 1960. "String Constituent Analysis," *Language*, Vol. 36, No. 1, January-March.

The grammars begin with a description of the highest level for which a formal definition has been described, and proceed from level to level, ending with roots and affixes. Each construction which manifests a tagmeme in the formulas is analysed on a lower level, or, in case of recursive constructions, on the same or a higher level. Rewrite operations⁵ for the generation of sentences from the formulas are not explicit, but are implicit in the grammars. By substituting each formula or lexical unit for its symbol, from the highest level to the root and affix, the grammar generates the grammatical sequences of the language insofar as the grammar and lexicon are correct and complete.

In these volumes of grammars which are uniform in underlying theory and general form of presentation, a comparison of sets of matrices might serve as a guide or outline for typological studies. For example, a glance at the clause matrices of Chacobo and Quechua immediately calls attention to major structural differences between the two languages: the basic cleavage between transitive and intransitive, and between complete and incomplete in the Chacobo, whereas these categories are relatively unimportant in Quechua, which emphasizes an affirmative-negative set of contrasts. Both of these languages have very complex clause structures, but the complexity in the Chacobo is in markedly different areas than that in the Quechua. A comparison of the clause matrices of these two languages with that of Movima again emphasizes the differences in the general structure of the languages.

On the other hand a comparison of sets of matrices for the purpose of establishing or corroborating a genetic relationship is sometimes or in some areas fruitful, sometimes not. It would be difficult to argue a relationship between the Tupi-guaranian Sirionó and Guaraní on the basis of Chart I of the Guaraní grammar (Guaraní 1.1) and Chart I, which most nearly corresponds to it in the Sirionó grammar (Sirionó 2.3.1), although minor features of the grammars are strikingly alike. The combination of function words on the clause level (Sirionó 2.3.3. and 4.12) and of particles in the verb phrase (Sirionó 3.1.2) which are peculiar to Sirionó corresponds very closely to the clutter of clitic particles of the verb phrase periphery which is characteristic of Guaraní (Guaraní 2.1.3). Likewise the series of predicates P1, P2, P3, P4 and the anteverb within one clause in Guaraní (1.1.2) are parallel to the series of dependent clauses filling

⁵ Longacre, Robert E. 1964. *Grammar Discovery Procedures, a Field Manual*. The Hague, Mouton and Co.

Sat slots in the content string (Sirionó 2.3.1.2) and the preverb in the verb phrase (Sirionó 3.1.2) of Sirionó.

The comparison of clause matrices in the Arawakan languages is more encouraging. Those of Baure (Chart II, 2.1.1) and Ignaciano (Chart I, 2.1.1.1) show noteworthy similarities as well as differences. In the Tacanan family the matrices of the clause, verb phrase, and verb of Tacana and Eseejja suggest similarity of structure, although some major differences appear (Eseejja Chart I, 2.1.1, Chart II, 3.1.1, and Chart III, 4.1.1; Tacana Chart I, 2.1, Chart II 3.2;1, and Chart IV, 4.1.1.1).

In noting both contrasts and similarities consideration must be given to differences in presentation which reflect the linguist's preferences rather than structural contrasts. An instance is the apparent difference between the Eseejja and Tacana verb charts, due to the presence of the quotative verbs 40 and 50 in Eseejja and the auxiliary verbs 50 in Tacana. This difference is easily determined to be primarily one of presentation by comparing the roots of the two sets of verbs (Eseejja 4.1.6.1 and Tacana 4.1.4), and their ultimate distribution in the quotative clauses (Eseejja 2.1.2 and 3.1.2.1; Tacana 2.2.2).

The comparison of the independent, dependent, and included clause matrices of Quechua (Chart I, 2.1.1 with those of Aymara (unpublished grammar notes by Warkentin, Heaslip, McNeil of the Summer Institute of Linguistics) shows a very striking similarity.

A few miscellaneous features which are particularly exotic or which may be of theoretical interest in the various grammars of this volume are listed below:

(1) In Quechua, included clauses whose distribution class is determined by the verb class (ditransitive, transitive, etc.) and by the noun case (subjective, objective, locative, etc.) of the verbal noun in the filler of the predicate slot result in an exceedingly complex system (Quechua 2.1, 2.4, 4.1, 4.2, et al).

(2) In Guaraní, a supersegmental morphophoneme of nasalization occurs (Guaraní 4.1).

(3) In Sirionó there is a unique morphophonemic span (Sirionó 5.2).

(4) Itonama and Guaraní are characterized by complex levels of morphophonemics (Itonama 5.2, Guaraní 4.2).

(5) In Movima are found noun phrases of extraordinary recursiveness, traced through at least eight layers, which are of high frequency in text and in conversation (Movima 2.3).

(6) Of the great variety of rare semantic categories which cor-

respond to grammatical classes in the various sections of these volumes, those which are described as "rhythm" syffixes classifying action in Itonama are perhaps the most unusual (Itonama 4.1, Classes 1220 and 1230).

Conventions of analysis and symbolism

In the following paragraphs various conventions and symbols are explained by means of notes and illustrations.

Generally in the tagmemic model units are considered to be in emic contrast on the basis of two structural differences⁶. These may be two differences in composition (one of which may be recognized by difference in transform potential) which necessitate the writing of separate formulas, or, following Pike, they may be one structural difference and one difference in distribution. For the purposes of these volumes we have further defined the basis of emic distinction as follows: A difference in distribution together with a difference in class of slot fillers which is relevant on different levels and/or is correlated with other emic distinctions (such as the difference in the lists of transitive and intransitive verbs in most languages) establishes emic contrast; however, a difference in distribution plus a difference in lists of slot fillers which are primarily semantically determined does not establish emic contrast. For example, we may describe one phrase, N 1, as manifesting both time and location tagmemes, although the time slot is filled only by a list of those phrases which include a word indicating time, while the location slot is filled by a different list of those phrases which include a word indicating place:

...+T:N 1 +L:N 1 ...

When only one difference between the formulas of two emic classes is apparent, a difference in distribution is to be assumed.

In two of the grammars, the Chacobo and Baure, it was found economical to describe the nuclei of the clauses separately from the margins. The resultant distinction between clause and clause nucleus is not considered to be a bona fide difference in level. The clauses are therefore shown as emically different, although the formulas combining nuclei and margins show that only the nuclei contrast.

We have not tested all combinations implied in the formulas. In a language with a great many clause classes with numerous marginal tagmemes each manifested by a variety of classes of constructions, the potential number of combinations is astronomical. In such cases when no restriction is apparent we have assumed that all possible combinations occur.

⁶ Pike, 1962 and Longacre, 1964.

Wherever it has been a matter of choice we have kept the higher levels simple at the expense of complexity on the lower levels.

Pike's technique of multiplication⁷ is frequently used; that is, each unit of a matrix (or each formula of a paradigm) is modified in the same way so that a corresponding matrix is established.

Co-occurrent tagmemes are said to be in portmanteau relationship when they are manifested simultaneously by the same construction.

When the formulas of a tagmemic paradigm are identical with the formulas of the tagmatic paradigm except for the slot fillers, the fillers are shown in the tagmemic paradigm, and the tagmatic paradigm is omitted.

When there is variation of order of tagmemes the most frequent order is shown in the formula, and the variations are described.

In tagmemic grammars a phrase is generally described as a construction involving an obligatory tagmeme and at least one other tagmeme, obligatory or optional. For the grammars in these volumes the phrase is better defined as a unit of the level between clause and word, which characteristically manifests a clause-level tagmeme. For economy of description in some instances we have treated as a phrase a single tagmeme manifested by one word, a unit which cannot be expanded, when it shares the distribution of a class of phrases.

In many of the grammars there are hesitation words which may interrupt any construction on any level. We have described the occurrence of the hesitation only where it is frequent or characteristic.

In the matrices each vector is represented by a millenium, century, decade or unit, in an arrangement such that the item in each cell is designated by a combination of the numerical symbols of the row and vector. For example in Chart I, Emic Clause Classes, of Quechua (2.1), the Intr column is numbered 30, and the Ig corr row is numbered 02. The symbol C1 30 refers to any intransitive clause; that is, to any of the clauses 31a to 34b. The symbol C1 02 refers to any interrogative corroborative clause, that is, to C1 12, 22, 32, 42, or 52. The symbol C1 32 refers to the intransitive interrogative corroborative clause. Similarly a row symbol 103 may be combined with decade symbols of the columns to produce the symbols 113, 123, etc.

In general the order of formulas in the paradigms follows column by column the order of the corresponding matrix.

⁷ Pike, 1962.

In the abbreviations lower case letters symbolize units which fill slots on the morphological level. The symbols of all higher level units and of all tagmemes begin with capitals. The sentence is labeled Sent, and the clause, C1; a construction whose symbol begins with a capital and which has no other indication of level, may be assumed to be a phrase; i.e. N is to be read "noun phrase," n is to be read "noun." The list of abbreviations gives only the forms with capitals, except in those cases in which the form with capital is not employed.

In citations of the indigenous language, however, capital letters are morphophonemic symbols, which are explained in the morphophonemic section of each grammar. Citations are written morphophonemically only where the morphophonemic form is particularly relevant.

In citations of the indigenous languages, the intonation contours are approximately marked by punctuation. Beyond that the contours are identified only to the extent that they are essential to the description of the grammar.

Class and numerical symbols are generally given together, except that affixes are symbolized by the numerical symbol alone. When, however, an affix numbered under a particular word class occurs in the composition of a different word class, that affix is shown with class symbol as well as numerical symbol:

vb stem 12 = ...+Caus: 3217, but

n stem 4 = +...-Caus:vb affix 3217.

Distribution classes of roots and constructions as well as morphophonemic classes are to be marked in the lexicon of a language. Loans are generally treated as roots.

N not followed by a numerical symbol, symbolizes any noun phrase.

A series designated by a hyphen refers to any one member of the series unless otherwise specified. The symbol C1 11-34 refers to any one clause numbered from 11 to 34. The symbol vb A-L refers to any one verb with a letter symbol from A to L of the alphabet.

If the formula of C1 11-34 includes a corresponding series, for example if it contains the tagmeme +P: Vb 11-34, it is to be assumed that in each clause of the series the numerical symbol within the formula will be identical with the numerical symbol of the clause. This means that in the example above the predicate of C1 23 will be manifested by Vb 23.

$\underline{+}(+A \ +B)$ indicates that the occurrence together of the tagmemes A and B is optional but that neither occurs without the other.

$\underline{+}(+A \ \underline{+}B)$ indicates that the occurrence of tagmeme A or of the

two tagmemes together is optional, but that tagmeme B does not occur without tagmeme A.

$\pm(+A +B)$ indicates that the occurrence of either one of the tagmemes or of the two tagmemes together is obligatory.

$\pm/+/-$ indicates that the occurrence of a tagmeme is optional under stated conditions, and obligatory under other conditions, and that it does not occur under other stated conditions.

$(...)^n$ indicates that the parenthetical item is optionally repeated an indefinite number of times.

T^3 indicates that the tagmeme T optionally occurs up to three times in a construction. In morphophonemic writing a superscript merely distinguishes symbols.

T 1, T 2 symbolize two tagmemes which are formally contrastive although they are similar in function.

The formula $+Base: n\ r\ 1/4\ Res/8/vb\ r\ 10 +Cpd: n\ r\ 3\ AS$ is to be read as follows: the obligatory Base tagmeme manifested by a noun root of class 1, by members of a restricted list of noun roots 4, by a noun root of class 8, or by a verb root 10; plus the obligatory tagmeme Compound manifested by members of noun root class 3 which arbitrarily select fillers of the Base slot with which to occur.

The symbol Res is thus employed of an etic class which is relevant at only one point in the grammar. The list of $n\ r\ 2\ Res$ in one formula is assumed to be a different list from that of $n\ r\ 2\ Res$ in another formula in the same language.

In the tagmeme $+Base: n\ r\ 2\ (lmo)$, the form cited in the parenthesis is the only $n\ r\ 2$ which manifests the Base.

When corresponding slot symbols are differentiated by number in tagmemic formulas, the number may be omitted in the tagmatic formula when the fillers clearly show the differentiation: P 22 and P 23 in the tagmemic formulas may be written as P: Vb 22 and P: Vb 23 respectively in the tagmatic formulas.

Esther Matteson
Technical Consultant
Linguistic Department

Riberalta, Bolivia
December, 1965

Table of Contents

Quechua	27
Guaraní	99
Sirlonó	195
Itonama	257
Movima	353