DOCUMENT RESUME

ED 036 782 AL 001 905

AUTHOR Hardman-de-Bautista, M. J.
TITLE Computerized Archive and Dictionary of the Jaqimara Languages of South America.
NOTE 10p.; Paper delivered at the Southeast Conference on Linguistics, Florida State University, March 1969

EDRS PRICE EDRS Price MF-$0.25 HC-$0.60
DESCRIPTORS *American Indian Languages, Archives, *Computational Linguistics, *Computer Programs, *Dictionaries, Grammar, Material Development, Morphology (Languages), Phonemes, Phonology
IDENTIFIERS *Jaqi Languages

ABSTRACT

The three extant members of the Jaqi (Jaqimara) family, Aymara, Jaqaru and Kawki, are spoken by over one million people primarily in Peru and Bolivia, but earlier members of the Jaqimara family were probably spoken throughout the whole area of present-day Peru. This paper gives an outline of some of the salient structural features of these languages and reports on progress in the preparation of a computerized dictionary of the Jaqi languages presently being adapted for the computer at the University of Florida. Each item of the Jaqi languages is encoded according to a number of categories any one of which or any combination of which may then serve as indexers for machine printouts: machine item number, language name, entry, allomorphs, derivations, grammatical class or classes, historical origin, dialect of item, source of information, English glosses, Spanish glosses, citations, dialectal and/or language cognates, coded dialect and/or language of cognates, reconstructed proto forms, variant spellings, date of entry, and name of transcriber. It is hoped that the dictionary will be useful as a source for bilingual dictionaries for Peace Corps or government use as well as for work in the reconstruction of Proto-Jaqi. Further plans include a computerized archive to include as much as possible of the structure of the Jaqi languages. (FWB)

Dr. M. J. Hardman-de-Bautista
Dept. of Anthropology
University of Florida

COMPUTERIZED ARCHIVE AND DICTIONARY OF THE JAQIMARA LANGUAGES OF SOUTH AMERICA

The Jaqimara languages are spoken in the Andes mountains of South America, primarily in Peru and Bolivia. There are now three known extant members of the family:

Aymara, spoken by approximately one million speakers in Peru and Bolivia, is of political, social, and economic importance in Bolivia.

Jaqaru, spoken by about two thousand people, is little known. It is spoken in a town only some 200 miles from Lima, the capital city of Peru. However, since the children are still learning Jaqaru, now along with Spanish, no date can be predicted for its demise.

Kawki is spoken in the village located in the valley over the mountain from where Jaqaru is spoken. There remain only some twenty speakers of the language, and, since they are all over sixty, the language will probably be completely gone within twenty years.

From toponymical evidence it appears that the Jaqimara family formerly had a much wider extension, most likely throughout all of the area that is today the country of Peru.

The dominant non-Indo-European language of the area is Quechua, the language of the Inca empire. Rather than a single language, some investigators consider this a family of languages. Whether or not the two families, the Jaqimara and the Quechua, are related or not, is still an open question. The work now in progress on the Jaqimara family is a prerequisite to any reasonnable statement as to their relationship. The two families have been in intimate contact for at least 2000 years, and probably more. From archeological and mythical evidence, it appears that the groups have been on top of each other alternately, which has doubtless led to layerings of borrowings, first heavily in one direction and then in the other. Therefore, the reconstruction of proto-Jaqimara and of proto-Quechua must be completed independently and then the layers of borrowings sorted out, before their degree of relationship, if any, can be determined.

Before explaining the computerized work in which we are now involved, I would like to give you a brief outline of some of the salient structural features of the Jaqimara languages.

The languages differ more phonologically than they do in grammatical patterning, although there are some striking differences there also.

Jaqaru and Kawki are more similar to each other than either is to Aymara, but Kawki is closer to Aymara than is Jaqaru.

Jaqaru and Kawki have 39 segmental phonemes, 36 of which
HANDOUT are consonants, 24 of which are voiceless stops. Aymara has 9 voiceless stops less and also does not distinguish velar and alveolar nasals nor alveolar and palatal sibilants. It does, however, distinguish pharyngeal and velar voiceless fricatives. Aymara has also extended greatly the use of vowel length which in Jaqaru and Kawki is restricted to the first syllable of roots. In Aymara it has important grammatical functions, and
HANDOUT one morpheme, first person future, consists entirely of vowel length. Stress, on the penultimate syllable of a word, is not phonemic in any of the languages,

All of the languages show three basic classes of roots: nouns, verbs, and particles. There are subclasses of nouns. The particle class is small in all the languages, consisting primarily of greetings and exclamations. The languages are all suffixing. Suffixes are of three types: nominal, verbal, and independent. These last express syntactical relationships. Although they do occur on a particular word within the sentence, they are syntactically distributed and related to the sentence as a whole rather than to the word with which they happen to occur. Verbal suffixes are divided by all the languages into derivational and inflectiona.

The Jaqaru and Kawki word structure is as seen on the
HANDOUT handout. Aymara, however, has a class of independent suffixes which apprently has no cognate in Jaqaru and Kawki and which affects the structure of the word. The Aymara word structure is seen below that of Jaqaru.

All the languages have a basic four person system reflected both in nominal and in verbal suffixes. Number is not indicated within the person system.

The persons are:
1 - speaker involved; addressee excluded;
2 - addressee involved, speaker excluded;
3 - neither speaker nor addressee involved;
4 - both speaker and addressee involved.

HANDOUT The fourth person, although apparently plural to English or Spanish ears, is not so withing the system. All of the languages have optional plural, both nominal and verbal, that will emphasize great quantity or all inclusivess. These may be added to any of the four persons. These suffixes, moreover, are not cognate between the languages, while the four person system is. The nominal plurals are Jaqaru /-kuna/, Aymara /-naka/, verbal plurals Jaqaru /-rqaya/, Aymara /-pxa, -pka/.

The nominal suffixes are of the pure relational type; several suffixes may co-occur on the same stem.

The verbal suffixes are either derivational or inflectional. The derivations are quite extensive, and are largely cognate. They tend to specify movement or direction or aspect. The

HANDOUT

inflectional system shows both actor and goal and thus the persons in the verbal paradigm are ten. Aymara lacks one of the ten persons - apparently 2-1 and 2-4 have fallen together. I would not be surprised if the missing form showed up in some dialect of Aymara. Tenses are distinguished, including remote, past, future, imperative (with all ten persons), etc. Aymara differs from Jaqaru in than the unmarked tense is a non-future, i.e., present or past. Jaqaru lacks the unmarked tense, and the tense marked present (which is lacking in Aymara) is general, present, or near future. Verbs are inflected in all the languages for occurrence in principal or subordinate clauses, although Aymara does not indicate person in the subordinate, while Jaqaru does.

The independent suffixes show the nature of a sentence, whether statement or question - several types of each are distinguished. Also, degrees of probability are indicated and degrees of confidence in the source of information.

The comparative work involving the three languages has just begun. It is hoped that the dictionary that we are now developing will be of great help.

The dictionary now being prepared was started while I was at Indiana University. They are continuing work there on a similar project for Quechua. We are now adapting the programming for the computer at the University of Florida.

The programming is as follows: Each item of the Jaqimara language is encoded according to a number of categories (currently 18) any one of which or any combination of which may then serve as indexers for machine printouts.

HANDOUT

The categories encoded are the following:

1. Machine item number. This is required by the programming to identify uniquely each item. We use it redundantly with the language name so that each language, and later each investigator, ban be assigned a 'million number' and then number his entries consecutively. For example, 5 is Kawki, 7 is Jaqaru, 9 is Aymara. -91 could be Bertonio's dictionary, etc.

2. The language name.

3. The item for entry. This may be a root, a stem (derived form), a complex form (place names tend to be very complex), or an affix - only suffixes here, but others could be added if ever needed.

4. Allomorphs. All of the alloforms of the item listed are given. These may then be listed separatly sending the reader to the total form for more information, or they may simply be given as information with the complete form.

5. Derivations. This included the derivations from the

item listed as well as the forms from which it is derived. This category may contain as many entries as necessary. This will allow later cross-referencing of derivational forms, e.g. all the forms containing the suffix -k^{h}a.

6. Grammatical class or classes. The listing may be as detailed as desired. We currently have 28 categories, but this may be expanded when needed. Printouts may also be general or detailed, e.g., we can request only and all nouns, or only number nouns. U is included for forms unknown as to grammatical classification. Because of the multiple listing possible, one may guess the classification and indicate this by also listing a form as U.

7. The historical origin. If the form is known to have been borrowed this is indicated as well as the source language.

8. Dialect of item. This is not relevant for Jaqaru, but will be for Aymara.

9. Source of information. The text or informant who provided the data item is listed. This will allow us to extract material from only one source, say Bertonio's 1612 dictionary, or having the material listed with all the rest.

10. English gloss(es). Care is taken to list as many glosses as necessary for the form to appear as often as desired. For instant 'to sew a shoe' /ch'illu/ (Jaqaru), would be listed under 'sew' and 'shoe'.

11. Spanish gloss(es).

12. Citations.

Thirteen and following give historical and comparative information.

13. Dialectal and/or language cognates. At present this is very loosely used, and may include items with only semantic similarity.

14. Coded dialect and/or language of the cognates listed in thirteen. The list must be one to one, thus permitting an unlimited number.

15. Reconstructed proto form. This is not yet in use because the proto forms have not yet been reconstructed.

16. Variant spellings. This may include anything the particular transcriber wishes. We require all forms entered to follow a particular transcription in order to allow greater cross-comparison. Here old spellings, favorite transcriptions, original spellings, popular or received spellings, may be listed.

17. Date of entry.

18. Name of transcriber.

This dictionary then will be flexible enough to allow constant updating, but in all cases items will be clearly identified as to source, so that the information will not be scrambled except to the degree desired, and when desired. The indexer language may be the Jaqimara language or Spanish or English. The target languages printed out will be minimally two of the others, but could included all of them. E.g. all of the equivalents for English 'head' in all of the Jaqimara languages could be given at once.

The possibilities for the use of this dictionary are many. The first, obvious, parctical use is simply as the source for traditional bilingual dictionaries that could be published for student, Peace Corps, government, etc. use. In addition the dictionary will be used in the development of the reconstruction of Proto-Jaqimara, by making comparisons easily available. As the correspondences are discovered, it is planned to program the machine such that it would be able to check them for me and pull out all examples.

Along with the dictionary we are planning to develop an archive to include as much as possible of the structure of the language. The dictionary itself will handle morphology rather well. Suffixes are filed carefully, and the glosses made up so that they will reflect each of the relevant grammatical categories for that suffix. For example, /-sama/, with allomorphs /-sma/ and /-sam/ is listed as having the following glosses:

2p to 3p
second person to third contrary to fact
person, second to third contrary to fact
contrary to fact, second person to third.

This will give us listings thus: all of the suffixes for 2p to 3p, whatever the tense; this suffix along with all other personal suffixes; this suffix along with all other contrary to fact suffixes in a paradigm. As we work more with this aspect of the dictionary, we may wish to improve it. Printouts can be made including or excluding affixes utilizing the grammatical classification index.

We are now in the preplanning stage for the phonological archive; the syntactical stage will have to wait. Much of this will come as we get the comparative dictionary into full functioning. It is hoped that eventually the Central Dictionary and Archive of the Andean Languages can become a center where scholars in the field can obtain the necessary information rapidly and thus themselves contribute to the growing knowledge of the Andean languages.

HANDOUT

M.J.Hardman-de-Bautista
Department of Anthropology, University of Florida

COMPUTERIZED ARCHIVE AND DICTIONARY OF THE JAQIMARA LANGUAGES

Examples will be given throughout for Jaqaru and for Aymara. Kawki is often omitted because there is not sufficient data.

Jaqaru phonemic system:

p	t	T	ʧ	č	ĉ	k	q
pʰ	tʰ	Tʰ	ʧʰ	čʰ	ĉʰ	kʰ	qʰ
pʔ	tʔ	Tʔ	ʧʔ	čʔ	ĉʔ	kʔ	qʔ
m	n	ñ				ŋ	
	s	š					
	l	ɬ					
w				y			
	r					h	

i u : (vowel length)

a

Aymara phonemic system:

p	t	č	k	q
pʰ	tʰ	čʰ	kʰ	qʰ
pʔ	tʔ	čʔ	kʔ	qʔ
m	n	ñ		
	s			
	l	ɬ		
w		y		
	r		x	h

i u : (vowel length)

a

Aymara: /sarata/ 'gone', /sara:ta/ 'you will go'

/sara:/ 'I will go'

Jaqaru word structure:

{root / stem} + {nominal suffixes / verbal suffixes} + independent suffixes

Aymara word structure:

{root / stem} + derivational suffixes (some verbal and all nominal) + Independent suffixes non-final +

inflectional suffixes + final independent suffixes

Nominal persons

Jaqaru	Aymara	/uta/ 'house'
-ŋa	-xa	1p
-ma	-ma	2p
-p^{h}a	-pa	3p
-sa	-sa	4p
utŋa	utaxa	'my, our house'
utma	utama	'your house'
utp^{h}a	utapa	'his, her, their house'
utsa	utasa	'our house'

Verbal persons:

	Jaqaru	Kawki	Aymara
1 - 3	-t^{h}a	-ta	-ta
2 - 3	-ta	-ta	-ta
3 - 3	-i	-i	-i
4 - 3	-tana	-tana	-tana
1 - 2	-ima	-ima	-sma
2 - 1	-uta	-ita	-ista
2 - 4	-ušta	-	
3 - 1	-utu	-itu	-itu
3 - 2	-tama	- :	-tama
3 - 4	-ištu	-stu	-istu

future:

1 - 3	-ŋa	-:
2 - 3	-mata	-:ta
3 - 3	-ni	-ni
4 - 3	-tana	-ñani
1 - 2	-mama	-:ma
2 - 1	-utumata	-ita:ta
2 - 4	-ištumata	
3 - 1	-utuni	-itani
3 - 2	-matama	-:tama
3 - 4	-ištuni	-istani

Sample dictionary imput:

×1× 7001579

×2× JAQARU

×3× YATXI

×4× YATX×

×5× YATXK+A, MEMORIZE, APRENDER DE MEMORIA×

×6× V ×

×7× NATIVE

×8× (not relevant here)

×9× DIMAS × MARKA ×

×10× LEARN × ACCUSTOMED, BECOME × KNOW ×

×11× APRENDER × ACOSTUMBRARSE × SABER ×

×12× YATXKTATXI, NO SABES; YOU DONOT KNOW × YATXKTKASWA, VOY APRENDIENDO, I AM LEARNING ×

×13× YATI. ×

×14× A ×

×15× (not yet established)

×16× (none here)

×17× TODAY

×18× HARDMAN

×1× 7001765

×2× JAQARU

×3× WAKA

×4× WAK ×

×5× WAK JAYRA, COW DANCE, BAILE DE VACA ×

×6× NN ×

×7× SPANISH

×8× (not relevant here)

×9× DIMAS × MARKA ×

×10× COW ×

×11× VACA ×

×12× WAK ILLIR, PARA VER LA VACA, TO SEE THE COW ×

×13× WAKA × WAKA ×

×14× A × K ×

×15× (not yet established)

×16× HUACA ×

×17× TODAY

×18× HARDMAN