Mayan Master Speakers – The Archive of the Indigenous Languages of Chiapas

John B. Haviland¹,²

¹ Departments of Linguistics and Anthropology, Reed College, Linguistics, Portland, USA
² CIESAS – Sureste, Center for Research and Advanced Studies in Social Anthropology,
   San Cristóbal de las Casas, Chiapas, Mexico

ABSTRACT

The Archive of the Indigenous Languages of Chiapas is an electronic database documenting the three principal Indian languages of Chiapas, Mexico. This report describes the design philosophy behind the archive, intended to distribute its results in digital form via the Internet. It illustrates some of the products of the Archive, ranging from standard linguistic description and lexicography, through semi-experimental elicitation, to ethnographically situated interaction characterized by different sorts of speech genre. It also discusses representational and ethical issues derived from electronic distribution of digital media in linguistic documentation.

Key words: archive, indigenous languages, Chiapas, Mexico, Tseltal, Tzotzil, Chol

Introduction

In this brief report I describe the ongoing creation of a linguistic database of indigenous languages of Chiapas, Mexico. The project has been funded by CONACYT, the Mexican National Council on Science and Technology, and both the research and the archived results are based at CIESAS (the Center for Research and Advanced Studies in Social Anthropology) in its southeastern branch in San Cristóbal de las Casas, Chiapas. I concentrate on central features of the design philosophy for this database, a philosophy which in turn derives from the specific ethnographic conditions and linguistic traditions in Indian Chiapas. The archive intends to document the best speakers of the languages in question, recognizing ethno-linguistic standards not just for “knowledge” of a language, but for true expertise or excellence at speaking (and these activities achieved through speech).
The Archivo de los Idiomas Indígenas de Chiapas (AIC) is somewhat grandiloquent for a digital database which documents the three most widely spoken Indian languages of Chiapas, Tzotzil, Tzelta, and Chol. The two largest of these, Tzotzil and Tzelta, have well over a quarter of a million speakers each, and the smallest, Chol, has around 140,000 speakers according to official 2000 census figures (INEGI, Censo General de Población y Vivienda del 2000) (Figure 1). All are nonetheless endangered, in slightly different ways and with different degrees of urgency. Children continue to learn all three languages as first languages in some communities, whereas in all three cases there are also communities where children acquire Spanish despite the fact that their parents are native speakers of the Indian language. There are also instances of multilingual communities where several Indian languages are spoken, along with Spanish, or where quite divergent dialects of a single language must accommodate to each other, producing new hybrid forms at least in the short term.

The AIC explicitly incorporates multimedia, including text, relational databases, and audiovisual media including both audio and video. In order to document not only the essential linguistic and lexicographic facts of the target languages but also central speech genres and typical contexts of speech, we have opted for an ethnographic approach to language description, starting with situations and speakers that exemplify what native speakers view as especially noteworthy occasions for talk or special speaking skills. The AIC accordingly has emphasized the education and involvement of native speakers in the compilation of the database, seeking out young speakers of the three languages who have both acute linguistic ears and who are prepared to help compile a broad ethnographic sample of speech by enlisting the cooperation of other speakers in recording, transcribing, and ultimately analyzing the resulting corpus of materials. From the beginning, the aim of the AIC has been to make its results available digitally, via the Internet, supplemented by other forms of digital distribution (for example, via CD-ROMs or traditional print media). The final section of this paper lists the currently available Internet links or URLs where the AIC can be consulted electronically. There were doubts about such a means of distribution and access at the beginning of the project, given the general lack of Internet access among speakers of Chiapas Indian languages at the time. However, in the intervening years such access has ballooned, and even small centers of population in Indian Chiapas now have available Internet connections, and many educated speakers routinely communicate via electronic mail.

**Bases, Objectives and Methods**

Certain theoretical and practical precepts have guided the creation of the AIC. First, we believe in the centrality of descriptive and typological linguistics within linguistic theory. That is, an abstract theory of human linguistic capacity can only succeed if it is based on a descriptively adequate and typologically varied account of human language behavior. From a wider anthropological perspective, the design of the AIC also reflects our conviction that language is a constitutive force in social life in general, that people construct, maintain, and transform their social relations with others in large part through talk. More particularly, the creation of the AIC was meant to emphasize the importance of language as a resource in processes of sociopolitical change in modern Chiapas, where demands for greater recognition of Indian rights and needs have repeatedly underscored the central practical and symbolic role of Indian languages in constituting a new relationship among Indian communities and between Indians and non-Indians. Finally, the creation of the AIC reflects our perception of a need for a developed technology of knowledge in Mexico (and especially in Indian Mexico) — a technology that allows rapid and flexible transformation and communication of different ways of understanding the world. We have chosen, in particular, to try to apply such technologies to the spoken and written word. The project involves, in the first instance, linguistic description, ranging from the basic formal facts of the languages, to their notable typological features, and aspects of sociolinguistic and pragmatic usage in different speech situations. First, in addition to basic facts of grammar, lexicon, and morphology, there are typologically striking issues: the encoding of motion, the class of roots traditionally called "positional" in Mayan linguistics (see below), the class of relational nouns, the facts of verbal voice, and multiple uses of grammatical possession. From sociolinguistic studies of variation (between dialects, within interactions, or across speech genres), we have also investigated the pragmatics of gesture. From the beginning, with varied success, we have emphasized both what is called in Spanish la formación de recursos hu-
Some Examples of Material from the Archive

One example of transforming an existing resource from previous research by other scholars by adapting it to electronic presentation is the tagging for automatic retrieval of existing dictionaries. Significant collections of lexical material exist for all three languages in the Archive, and we are exploring different representational formats for them. One example is a dictionary of the Chol of Tumbaid published in both print and machine-readable form by the Summer Institute of Linguistics. For purposes of the Archive this dictionary has been restructured in a more widely used practical orthography and tagged in the style of the lexical database program Shoebox**, making it available for systematic correction, consultation, and extension. Here is a sample entry from the dictionary:

\begin{verbatim}

\texttt{ach' esxibl \textbf{fiki} lam cha'a \textit{ja} al} =

\texttt{Lo \textit{tierra} \textit{est}a \textit{mojada} por la lluvia.}

(The ground is wet from the rain.)

\texttt{sub 2}

\texttt{cat adj}

\texttt{leesp \textit{re}x\textit{ada} 'eng watered}

\texttt{see Ach esxibl \textit{fini} pak' \textit{ab}al} =

\texttt{La \textit{hort}aliza \textit{est}a \textit{re}x\textit{ada}.}

(The garden is watered.)
\end{verbatim}

Each field here is labeled with a tag that identifies the nature of its contents: the original entry or morpheme (\texttt{lem}), a sub-entry (\texttt{sub}) corresponding to one phrasal gloss, the first of which belongs to the word category (\texttt{cat}) - 'adjective', with Spanish (\texttt{leesp}) and English (\texttt{en}) glosses as shown, and an example (\texttt{ex}) sentence from the original dictionary, and so on, for a second sub-entry.

A second example that better illustrates the strengths of a multimedia format is the presentation of basic verbal paradigms. Mayan linguistics traditionally distinguishes two sets of verbal affixes, set A - usually prefixes which vary with the phonological shape of the verb stem and show person, indexing an ergative argument (and also a possessor on nominals) - and set B, which cross-indexes an absolutive argument. The following sentences from the Tzeltal dialect of Petalingo illustrate set A of prefixes in this language:

\begin{verbatim}
\texttt{lej k-il k-na 'I saw my house'}
\texttt{lej aw-il k-na 'You saw your house'}
\texttt{lej \textit{aw} \textit{il} k-na 'Juan saw his house'}
\texttt{lej aw-ilh k-na 'You (all) saw your house'}
\texttt{lej k-ilh yejil k-na 'We saw our house'}
\texttt{pisil lej k-il k-na 'We all saw our house'}
\end{verbatim}


** Shoebox is another product of the SIL, one of their tools for computationally aided linguistic description, for which information is available at http://www.sil.org/computing/shoebox.

---

\texttt{Ja'\textit{ik} lej yil s-najik 'They saw their house'}

However, in addition to simply presenting the written paradigm, on the Web we are able to show a Tzeltal speaker pronouncing the forms, with a subtitled sound video (Figure 2).

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure2.png}
\caption{Tzeltal speaker pronouncing verbal paradigm.}
\end{figure}

More traditional linguistic description is taken from the Master's thesis of a Chol speaking linguist whose work was supported by a grant from the AIIC. Juan Jesús Vázquez A., from the Chol-speaking town of Tila, Chiapas, finished a thesis on the morphology of the Chol verb for the \textit{Maestra de Lenguisticas Indigenas} of CIESAS, and the entire thesis is available online at the AIIC cite. Here are a couple of Vázquez's examples, which illustrate a difference in basic word order in sentences whose arguments differ in animacy and definiteness:

\begin{verbatim}
mi i-xal \textit{a w\textit{i}nik \textit{i} i-xal}
\end{verbatim}

* Note, of course, that many of the original entries in this dictionary, including this one, are multi-morphemic.
Without giving full details of the morphology and syntax, we can show how the Tzotzil narrator makes careful use of a different formal devices to maintain a specific perspective on the argument structure of this mini-disscourse. There are two participants, the boy (k'ox small krem 'boy') and his dog (23rd person possessive tz'e 'dog'). The narrator adopts the boy as principal protagonist. In line 1 he introduces the boy as topic, followed in 2 by a normal active transitive clause (Verb followed by Object), 'he was putting on his shoe.' In 3 he continues with a further transitive verb, 'he was holding his dog this way,' and in 4 he adds a secondary predicate, the positional ja'awl 'belly up', that describes how he was holding the dog (Figure 5). In 5 he wants to describe the fact that the dog licks the boy on the cheek. To maintain boy as the topic of the resulting clause he must perform two syntactic operations: he must transform the verb into a passive form, via the suffix -nt, denoting the licking dog to an oblique agent ('the boy's cheek was licked by the dog'), and he must apply an applicative suffix -b to the verb to promote the possessor of the cheek - the boy, that is - to the position of grammatical subject ('the boy had his cheek licked by the dog') (Figure 6). The nominal expression anos sat (literally, 'the rock of his face', i.e., the cheek) becomes a grammatical clouvre.

Note that by including video representations in the digital database, the AIC also makes possible analysis of important cultural and linguistic aspects of spoken interaction that might otherwise be lost. Language evolves and is acquired in the context of face-to-face interaction, and it is thus by its very nature multi-dimensional and cross-modal, involving a variety of signaling channels and modalities at once. One of our interests in the case of Chiapas languages has been the use of space as an expressive resource, a central part of deixis reference, for example, and prototypically linked to gesture. A simple but telling example is the iconic gesture, adopting the viewpoint of the little boy protagonist, the Tzotzil narrator uses to illustrate the positional secondary predicate ja'awl (Figure 7). (Note that in Figure 4, too, the narrator gesturally adopts the...
perspective of the boy in his story: he performs the licking as though it were happening to his own cheek.> Positional roots in Mayan languages are formally distinguishable from other roots by the sorts of stem derivations they accept; semantically, they are a hypertrophied class frequently denoting specific positions and configurations of different sorts of objects; they often appear as secondaries predicates with reduced inflection and syntactically governed agreement patterns. Expressively, they also seem frequently to co-occur with iconic gestures, as though the positions they denote are communicatively salient to speakers.

Speech Genres

The final examples I present reflect the AIC’s wider interest in linguistic knowledge, ability, and mastery: not just the minimal shared competence that all native speakers are presumed to have, but the expert linguistic knowledge achieved by specialists, performers, and cultural virtuosos, what we have been calling “master speakers.” Specialized linguistic expertise is also associated with different kinds of talk; genres linked to specific sorts of activities and social situations, often with quite specific lexicons and constructional features. Among labeled ethnomusicological speech genres in highland Chiapas, one can distinguish such categories as prayer, scolding, greetings, gossip, denunciation, requests, jokes, stories, riddles, proverbs, and ridicule, among others.

Of particular interest is the language of prayer. Throughout Mesoamerica certain sorts of ritual speech are structured in parallel, with utterances in matched pairs which differ only in a single lexical element; taken together the two matched lexical elements form a standard pair, incorporating a single cultural image. In the first line, for example, illustrated in Figure 8, the curer refers to her patient as manbil vinik // tañibil vinik —the bought man // the paid-for man. The paired words are man / itsaj, two transitional verb roots for ‘buy, pay.’ The reference, in turn, is to the idea that we humans here on earth have had our souls redeemed — bought and paid for — by the crucifixion of Christ. The standard way to refer to the care of a saint is to place the object of his care under his hands / feet (line 2). The ancestral deities who have the power to cure illness are conceived of as holy fathers / mothers (line 3), who can act effectively by agreeing, that is, having a shared word / mouth (line 4). In this highly marked linguistic variety of prayer there remain many archaic words, as well as loans from the earliest period of Spanish contact with the Maya languages of the region. In these speech communities prayer is perhaps the prototype of ‘good speech,’ thought of as elegant, expert, powerful, and efficacious: it is, in the end, what cures people, and the ability to produce it is thought, in the Tzotzil community of Zacapantla at least, to be bestowed on shamans in prophetic dreams.

This observation, finally, brings us to another special sort of narrative, the retelling of dreams. In Indian communities of highland Chiapas, dreams are accorded special significance. They are routinely interpreted as portents of both future events, and of spiritual or interpersonal crises. The standard conception is that in dreams human beings have contact with the adventures of the ch’ulel or ‘soul’; the events experienced in dreams take place not ta sbo balamul ‘on the face of the earth’ but in some other realm where the soul can travel freely and interact with the souls of others, both alive and dead. I will comment here on a single notable feature of narrated dreams, the linguistic marking of the fact that the events experienced in dreams come from a non-ordinary realm of existence.

Dreaming of a dead man

1. ta la, puch’ul yietel ti jmanwauj un
It seemed that Jesus Christ was lying there LA.

2. pero bixata
But what do you think?

3. anima la
It was a dead person LA.

Tzotzil makes use of an evidential clitic la, often called a ‘hearer’ clitic.’ More precisely, it marks the utterance in which it occurs as originating with a speaker (i.e., having as illocutionary source someone) other than the one who is actually speaking. In a declarative sentence, this has the effect of suggesting that what one is stating is attested by another person, i.e., that it is overheard or reported or claimed by someone else. (With interrogative and imperative sentences it has the effect of distancing the speaker from the question or the command: someone ELSE is asking or commanding.) In conjunction with other evidential clitics, quite subtle nuances of meaning, moral and epistemological commitment, and responsibility can be expressed

Fig. 8. A Tzotzil shaman prays to cure a male patient.

Fig. 9. Dreaming about a dead man.
The striking fact, for present purposes, is that careful Tzotil speakers, like my ritual kinsmann PV in the extract shown, treat events in dreams as if they are attested not by the dreamer but by some third party, removed from oneself: they are events not witnessed directly, but witnessed by one's soul, and therefore in need of evidential distancing. In the dream, PV has a series of encounters with disturbing characters, who show him a figure called Jnoanuay 'Buyer of Souls' who represents a Christ, and who has been an important saint in PV's personal history. In his dream, he approaches the figure, and, as he tells us, 'it did seem to be Christ lying there—but he adds, the evidential is suggesting that this appearance, to his soul in the dream, was possibly deceptive (as, indeed, most dream images are). Even worse, when he came close to the putative saint, it turned out it was really (according to the soul again, hence the evidential so once more) the body of a dead person, also a dangerous omen in dreams. The linguistic facts underscore formally what the culture believes implicitly: that the realm of souls is real and consequential, but that one must take the reports of the soul with a degree of epistemological skepticism.

The Database and Problems of Digital Representation

The AIC is a work in progress, depending on the collaborative efforts of many researchers and native speaker linguists, as well as the cooperation of other academics, and of the speakers themselves who contribute their time and knowledge to the database. The overall aims of the AIC are ambitious, since they encompass not only grammatical, lexical, and discursive descriptions of the three languages of the project, but also comparison of dialects within the languages, changes over time, and at least some comparisons with other languages of the Mayan family. Ideally, the materials of the archive—which include not only traditional academic formats but also multimedia recordings, 'talking transcripts,' etc.—will feed further practical applications, such as the production of didactic teaching materials for schools and other educational programs, in line with the needs and priorities of speech communities. Similarly, the perfect outcome of such a project would be for members of the speech communities themselves to take responsibility and initiative for maintaining and augmenting the Archive over time.

Digital access, especially via the Internet, also brings problems, many of which we have not yet resolved. There are ethical problems and issues of intellectual property that plague any database in which identifiable individuals appear, or in which experts reveal and discuss their expertise. Such issues are not easily resolved by the standard Western device of individual permissions and informed consent, since the local criteria and ethno-theories of knowledge, essential to the selection of materials for documentation in the first place, may consider property to be collective, or may otherwise clash with Western individualistic ideologies of possession, responsibility, and control. We are barely beginning to explore such issues, together with the native speakers who provide the raw material for a linguistic archive like the one in construction here. These speakers are in the end the ones who need the training, the credentials, the official recognition, and the power to promulgate and defend their own languages.

URLs

The following URLs can be consulted to see incremental progress in the creation of the AIC. The home page, in Chiapas, Mexico, is:

http://www.ciesasuneste.edu.mx/juan/Archive.html

A frequently updated mirror in the USA is:

http://academic.reed.edu/linguistics

REFERENCES


J. B. Haviland

Departments of Linguistics and Anthropology, Reed College, 3203 SW Woodstock, Portland, OR 97202, USA
e-mail: johnh@reed.edu

MAJANSKI GOVERNORI - ARHIV AUTOHTONIH JEZIKA CHIAPASA

SAŽETAK

Arhiv autohtonih jezika Chiapas elektronička je baza podataka koja sadrži jezične podatke triju glavnih indijskih jezika meksičkih Chiapas. U ovom je radu izložena struktura samog Arhiva čiji je krajnji cilj raspolaganje rezultata u digitalnom obliku putem interneta. Iznesu se nekoliko primjera arhivaštine, a standardni jezni opis i leksikografija, prvo poluexperimentalne elekticije u etnografsku interakciju koju karakteriziraju rečni govorni žanrovi. Dobu se također i specifični problemi koji proizlaze iz elektroničke prezentacije digitalnih medija u jezičnoj dokumentaciji.