STUDIES IN MESOAMERICAN LINGUISTICS

Report #4
Survey of California and Other Indian Languages
Reports from the Survey
of California and Other Indian Languages

Edited by Alice Schlichter, Wallace L. Chafe, and Leanne Hinton

Report #4

copyright © 1983 by the Survey
of California and Other Indian Languages

cover design (Santa Barbara Chumash rock painting) by Leanne Hinton
Contents

Jon P. Dayley

Voice and Ergativity in Mayan Languages ....................... 5

Thomas W. Larsen

Aguacatec Syntax from a Functional Perspective ............... 120

Martha J. Macri

Maya Writing: Linguistic Evidence for Eastern
Mayan Influence ........................................... 220

Claudia Brugman

The Use of Body-Part Terms as Locatives in
Chalcatongo Mixtec ........................................ 235

Martha J. Macri

Two Noun Class Systems in Mixtec .............................. 291

Nicholas Paracis

Preliminaries to Tonemic and Tonomechanical
Analysis for the Chalcatongo Dialect of Mixtec ............. 307
Voice and Ergativity in Mayan Languages

Jon P. Dayley

Boise State University

In recent years there has been a good deal of interest in ergativity and voice in linguistics because of the implications these have for grammatical theory. This paper is basically a description of the verbal category of voice in the context of ergativity in Mayan languages (see Table 1). Mayan languages are morphologically ergative, and many of them display characteristic features of syntactic ergativity as well. In addition, split-ergative constructions are also found in many of the languages. The category of voice is well developed in the family, and many of the languages have rather complex voice systems which include active, one or more passives, one or more antipassives, and instrumental and referential voices. The main purpose of this paper is to bring together data on voice and ergativity in the family as a whole. Voice and ergativity are discussed in some detail for at least one language from each of the main subgroups of Mayan: Huastecan, Yucatecan, Cholan, Tzeltalan, Kanjobalan, Mamean, Quichean, Pocom, and Kekchi.

Before actually presenting descriptions of voice and ergativity for particular languages, a number of concepts and working definitions are introduced in the section below 'Some Concepts and Working Hypotheses'. The discussion is informal and not meant to be definitive. The theoretical approach that emerges is perhaps at the same time rather eclectic and idiosyncratic, although I hope not necessarily inconsistent.
After discussing some concepts and definitions, a general characterization of what Mayan languages are like is presented. Then a discussion of voice and ergativity in particular languages is given. In the conclusion, some generalizations on ergativity and voice are stated and some hypotheses on voice and ergativity in Proto-Mayan are proposed.

Some Concepts and Working Definitions
Transitivity and the notion of subject.\(^3\)

An a priori assumption made in this paper is that there are two types of basic and fundamental universal activities recognized in all languages: transitive and intransitive. Transitive activities are those which involve two participants: one is a 'doer' herein called the agent (A), and the other is a 'nondoer' called the patient (P). Intransitive activities are those which involve only one participant called a subject (S). The S may be a 'doer' or a 'nondoer', the distinction isn't necessarily important. All languages, therefore, recognize and manifest a distinction between two fundamentally different classes of verbs which denote the two basic types of activities, respectively: transitive verbs (TVs) which require two arguments, A and P, in direct relationship with them; and intransitive verbs (TVs) which require only one argument, S, in direct relationship with them. Any other arguments (or NPs) in relationship with a verb are in an oblique relationship with it (e.g. dative, locative, instrumental, etc.).

Because the two participants in a transitive activity have different functions (i.e. 'doer' and 'nondoer') languages need to indicate the distinction between As and Ps by some grammatical means such as: (1) case inflections or particles, (2) word order, and (3) cross-referencing on the transitive verb. Cross-referencing, however, must always occur in
conjunction with (1) or (2), or take semantic features of the particular verbs and noun phrases into account, since cross-referencing alone can not distinguish the function of two 3rd person NPs, if they are not marked in some other way.

With respect to IVs, since S is the only participant in intransitive activities, there is no particular need to distinguish semantic As from semantic Ps, so languages tend to ignore the distinction. However, there are several possibilities for the manner in which Ss may be marked: (1) they may be marked in the same way as A (= nominative/accusative marking, see the section below 'Ergativity and Pivot'); (2) they may be marked in the same way as P (= absolutive/ergative marking, see 'Ergativity and Pivot'); (3) they may be marked like A if they are semantically agent-like and marked like P if they are semantically patient-like (= agent/patient marking); and (4) they may be distinguished completely from A and P. Nominative/accusative marking is perhaps the most common and well attested in the world's languages (e.g. Indo-European, Uto-Aztecan, Quechuan, and Japanese, to name but a few). Absolutive/ergative marking is found throughout the world in a number of different language groups such as Mayan, Eskimo, Basque, Georgian, some Australian languages, Hindi, and some Polynesian languages. Agent/patient marking is found in Siouan, Caddoan, and Iroquoian languages, as well as others, but the marking never seems to be 100 percent semantically consistent. Apparently, the only languages that distinguish S from both A and P are the Australian languages in the Ngura group and perhaps Motu (although the data isn't clear on the latter; cf. Dixon, 1979).

The choice of the term agent for the 'doer' participant in a transitive activity reflects the view held here that there is a central or
core meaning to the term: agents par excellence are human (or at least animate) and volitional, and they initiate and control activities. However, in natural languages the notion of agent is usually extended to include experiencing participants as well as inanimate and nonvolitional participants which do something to something else or cause some effect in something else. The term patient also has a central or core meaning: patients par excellence are inanimate, nonvolitional, noncontrolling, and noninitiating, and they normally receive, suffer, or are affected by the action of some agent. However, in natural languages the notion of patient is extended to include any 'nondoer' participant, animate or inanimate, in a transitive activity, as well as to things perceived or experienced (which usually are not affected by being experienced).

There is, then, a hierarchy of agency\textsuperscript{4}, or potentiality for being an agent, the inverse of which is a hierarchy of patienthood (see (1) below). On the far left are first and second persons, who, because they are inevitably human, are highest on the agency hierarchy (but lowest on the patienthood hierarchy). On the far right are inanimate objects which are lowest on the agency hierarchy (but highest on the patienthood hierarchy).

Agency Hierarchy

(1) \{1st person\} > proper human > human > animate > inanimate noun

Included within the agency hierarchy are two other (sub)hierarchies: (1) the animacy hierarchy (human > animate > inanimate), and (2) the person hierarchy (1st and 2nd person > 3rd person). These hierarchies indicate what is most likely to be an agent (or inversely, a patient), or at least what ought to be an agent, other things being equal. The hierarchies
are especially relevant in ergative languages, and perhaps in all languages.

The term S (of IV) has no particular central meaning since it is the only participant in an intransitive activity and may be either a 'doer' or a 'nondoer'. However, some languages, perhaps, assign meaning to S by aligning it with either A or P.

It should be noted that the terms A, P, and S are defined as fundamental (deep) semantic-syntactic functions as well as relations; these semantic-syntactic relations and functions along with the contrast between transitive and intransitive verbs are taken as axiomatic. This view contrasts markedly with, for example, Chomsky's constituent structure and dependency relations, and with relational grammar (as outlined by Johnson 1976, and Perlmutter and Postal 1974). In neither of these two schools is there a claim made that the transitive/intransitive contrast is fundamental. In the Chomskyan school there are no basic terms (e.g. agent, patient, subject, object, etc.), rather, these kinds of relations are secondarily defined by dependency relations. In relational grammar, however, Subject, Object, and Indirect Object, are defined as basic and universal terms having first order relationships with the verb.

In relational grammar, the term deep, logical, or semantic Object more or less coincides with P as used here, although the two terms are defined differently. The Indirect Object in relational grammar has a first order relationship with a verb, but in this paper it is viewed as basically having an oblique relationship with the verb. The deep, logical, or semantic Subject (with a capital S) in relational grammar includes the two terms, A and S (subject with small case s). The view held here is that Subject is not a fundamental term (like A, P, and S), although it is
a universal category which primarily has syntactic relevance. Thus, in
the languages of the world the S of intransitive verbs and A of transitive
verbs are usually treated alike in constructions like: (1) imperatives
where both A and S are second persons; (2) jussives (e.g. I order you to
X) where the P of the main verb is coreferential with an A or S of the
subordinate verb, and therefore one of them is usually deleted; (3) those
with verbs like 'can', 'try', 'begin', and 'finish', where the A (or S?)
of the main verb is coreferential with the A or S of the subordinate verb
and therefore is usually omitted under Equi-deletion; and (4) 'make do X'
causatives where the P of the main clause is coreferential with the A or
S or the subordinate verb, and therefore, one is usually deleted. A and
S are treated alike (i.e. as Subjects) in constructions like these pro-
bably because A is a 'doer' and S may be a 'doer.'

The universal category of Subject as outlined above is often con-
fused with language specific categories having to do with topic, theme,
and syntactic pivot, by mistakenly calling them 'surface' subject.

Ergativity and Pivot

A language is morphologically ergative if it treats the S of IVs and
the P of TVs in the same way grammatically, and the A of TVs in another
way. This may be done with case inflections or cross-referencing affixes
on the verb (e.g. in Mayan languages). The S of IVs and the P of TVs are
called the absolutive and the A of TVs is called the ergative. This abso-
lutive/ergative system (henceforth, simply 'ergative' system) contrasts
with a nominative/accusative system (henceforth, simply 'accusative'
system), where the S of IVs and A of TVs are treated alike grammatically,
as opposed to the P of TVs. Here the S of IVs and A of TVs are the nomin-
ative and the P of TVs is the accusative.
Absolutive/Ergative vs Nominative/Accusative

(2) \[
\begin{array}{c|c|c}
\text{ergative} & \text{absolutive} \\
\hline
\text{TV agent} & \text{IV subject} & \text{TV patient} \\
\end{array}
\]

nominative accusative

Note that whether the S of IVs falls with the A of TVs morphologically, or with the P of TVs, makes no difference with respect to maintaining the fundamental distinctions of A, P, and S, since IVs have only one basic argument.

In general, the absolutive is the unmarked category and the ergative the marked category in ergative languages, while in accusative languages the nominative is usually unmarked and the accusative is the marked category (cf. Dixon 1979, Silverstein 1977). Marking here has to do with things like: (1) which categories have null case inflections and/or cross-referencing affixes; (2) which categories are obligatorily included in each sentence; and (3) which categories feed syntactic processes.

Note that in ergative languages, since the ergative category is the marked one, what really is being given special status is the fundamental term agent. It is kept distinct from both P and S (= absolutive), and in many ergative languages the syntactic processes affecting the ergative (= agent) are highly restricted. For example, in many Mayan languages an ergative NP can not be questioned, relativized, or clefted. Smith-Stark (1976b) has called this the 'Inert Ergative Constraint', which really means that the fundamental relationship between a TV and its agent cannot be violated directly, at least without making compensatory modifications. The notion of transitive agent is marked as if it were sacrosanct in ergative languages.
One characteristic feature of most ergative languages is that they display split-ergative systems (cf. Dixon 1979, Silverstein 1977). In other words, along side of constructions with ergative marking, they also have constructions which are marked in a nominative/accusative way (cf. Table 3, which identifies languages in Mayan which have morphological split-ergativity). Split-ergativity is not random, rather it occurs in well defined areas: (1) In some ergative languages certain IVs may require more agent-like marking because of their semantic features (e.g. jump, run, speak). In other words, IVs whose Ss are semantic agents, may require their Ss to be treated like As of TVs, thus making this part of the system essentially nominative. (2) In some ergative languages some Ss of IVs may be treated like As of TVs because their semantic features make them higher on the agency hierarchy (e.g. 1st and 2nd persons). Thus, with these NPs the system becomes essentially nominative. (3) In some ergative languages, nonpast/incomplete/imperfective tenses or aspects may require nominative marking while past/complete/perfective require ergative marking. Dixon (1979) claims that nonpast/incomplete/imperfective presuppose a more agent-like perspective for both IVs and TVs, so the Ss of IVs tend to be treated like As of TVs, making this part of the system essentially nominative. And (4), in some ergative languages subordinate clauses may require accusative marking while main clauses require ergative marking (we shall discuss the motivation for this in the Conclusion).

In Mayan languages, only one language, Mot, has a split system based on (2), the semantic features of nouns; first and second person are treated in an accusative way, third person in an ergative way. There are a number of Mayan languages which have splits in their tense/aspect systems, (3),
with accusative structures in the incomplete or progressive aspects and ergative marking in other tense/aspect. There are also a number of languages which use accusative structures in subordinate clauses, (4), and ergative structures in main clauses. It should be noted that in all Mayan languages with split-ergative structures, ergative marking has been extended to intransitive verbs. This is what is called an **Extended Ergative System** or a **Marked Nominative System** (Dixon 1979).

A language has **syntactic ergativity** if there are syntactic processes which rely on ergative structures for their operation. In other words, ergative structures feed syntactic processes like coordination, subordination, relativization, etc., rather than accusative structures. It is possible for a language to be morphologically ergative but not syntactically ergative if syntactic processes rely on accusative structures rather than on ergative ones (cf. examples in Dixon 1979). The more ergative a language is, the more syntactic processes will operate on ergative structures.

As noted above, in syntactically ergative languages usually the ergative category is restricted syntactically. For example, in Dyirbal (Dixon 1972, 1979), ergative NPs cannot be coordinated, relativized, or subordinated in purposive clauses. Ergative NPs first must be converted into absolutes by an antipassive transformation (see the next section on 'Voice' for a discussion of 'antipassive'). In many Mayan languages, ergative NPs cannot be questioned, relativized, or clefted: they must first be converted into absolutes by antipassivization (or passivization, at times) in order to participate in these syntactic constructions. In accusative languages, the obverse is usually the case; accusative NPs are restricted syntactically. For them to participate in certain syntactic operations they must be converted into nominatives via
passivization (cf. Keenan 1976, and Keenan and Comrie 1976, for examples). What this means is that in ergative languages, the absolutive is the syntactic pivot, whereas in accusative languages the nominative is the syntactic pivot. ⁸

In the literature there has been a good deal of discussion over the notion of 'subject', in large part due to confusing the universal category of Subject as defined above and the language specific category of syntactic pivot.

Voice

Voice is an overt grammatical category basically pertaining to transitive verbs. The function of voice is to indicate the relationship the verb has with its arguments. The normal unmarked voice is the active voice which manifests the fundamental transitive relationship in which there are two arguments, agent and patient, in direct relationship with a transitive verb.

(3) Normal Active Voice: TV A P

A change in voice involves a disruption of the basic transitive relationship, along with overt morphological or syntactic marking of such a change. ⁹

One important function of voice changes is to allow the omission of one of the basic arguments of a TV in the semantic (deep) structure. This is a discursive device which makes possible the discussion of a transitive activity without mentioning one of the basic arguments when it is unknown or irrelevant, or when there is a desire on the part of the speaker to withhold such information. However, the omission of one of the arguments requires overt marking indicating that the basic transitive relationship
has been disrupted. When one of the arguments is omitted, the normally transitive verb must be converted into a derived intransitive verb and the remaining argument is inflected like any other S of IV. (Obviously, the techniques for doing this are language specific.)

Thus, the **absolutive passive** voice marks the omission of the agent. Since the A is omitted the verb is converted into a derived intransitive and the P becomes the S. Thus, other things being equal, in an ergative language the S will be inflected as an absolutive, and in an accusative language the S will be inflected as a nominative.

\[(4) \quad \text{ACTIVE} \quad \rightarrow \quad \text{ABSOLUTIVE PASSIVE} \]
\[
\text{VT} A \quad P \\
\text{VI} S \quad (= P)
\]

Note that this view of passive is not the same as having transformations delete agents (whole, or dummy elements like someone/something). The agent is simply not specified lexically or referentially, although a non-specific agent is implied in *absolviv- construction*. True medio-passive voices differ from passives only in that there is no implied agent. An activity which is normally viewed as transitive is marked by medio-passive voice indicating that it occurs without an agent (see Table 9).

The **absolutive antipassive** voice is the obverse of the passive (see Table 10). It indicates that the patient has been omitted from the discussion of a normally transitive activity. Since the patient is omitted, the normally transitive verb must be converted into a derived intransitive, and the A becomes S of the derived IV. Other things being equal, in an ergative language the S will be inflected as an absolutive, and in an accusative language having an antipassive the S will be inflected as nominative.
Another important function of voice changes is to make possible the rearrangement of the arguments in relationship with a transitive verb. Here none of the participants basic to a transitive activity are omitted from the discussion, rather they are put in different functions with respect to the verb, usually by promotional or demotional processes. These rearrangements are usually motivated by syntactic restrictions (e.g. on what may be the syntactic pivot), and by discourse requirements (e.g. what is topic, theme, or in focus). Thus, for example, in many accusative languages, an accusative NP is not accessible to certain syntactic processes, so passivization is required in order to promote the P (= accusative) to the nominative case; the A is demoted from the nominative case to an oblique case. Since the A is no longer in direct relationship with the TV, it must be converted into a derived IV with P becoming S like with the absolutive passive. This type of passivization is also usually required when the P is the topic or theme of the discourse.

As discussed above in 'Ergativity and Pivot', in a number of Mayan languages there is an Inert Ergative Constraint which restricts the participation of ergative NPs in syntactic processes such as questioning, relativization, and clefting, or more generally when they are in focus. In order for an ergative NP to participate in such processes the normally transitive verb is put into the focus antipassive voice converting it to a derived intransitive. The ergative NP is clefted and promoted to
absolutive. There is also evidence that it is promoted to a higher
clause becoming a predicate noun. The patient is either demoted to an
oblique case or apparently remains intact.\textsuperscript{14}

\begin{center}
\begin{tabular}{ccc}
\textbf{FOCUS} & \textbf{ACTIVE} & \textbf{ANTIPASSIVE} \\
VT & A & P \\
\hline
\end{tabular}
\end{center}

\text{\hspace{.5cm} ---\rightarrow \hspace{.5cm}}

\begin{center}
\begin{tabular}{ccc}
\textbf{VI (oblique)} & A \\
\end{tabular}
\end{center}

(Here the circled A indicates that the agent is clefted, and
perhaps is in a higher clause.)

The \textit{incorporating antipassive} is another voice occurring in a number
of Mayan languages and many other ergative languages (as well as in non-
ergative languages, e.g. Algonquian). The incorporating antipassive voice
makes it possible for a nondistinct, nonspecific patient of a certain
class to be incorporated into the verb complex. Here there is no parti-
cular patient specified, only a class of patients without specific refer-
ence (like in the English forms 'to bird-watch' and 'to deer-hunt').
Since the patient is incorporated into the verb complex, it is no longer
an argument in direct relationship with the verb, so the verb must be
converted into a derived intransitive, and the A becomes S.

\begin{center}
\begin{tabular}{ccc}
\textbf{INCORPORATING} & \textbf{ACTIVE} & \textbf{ANTIPASSIVE} \\
IV-P & A & S (= A) \\
\end{tabular}
\end{center}

There is another type of voice which occurs in a number of Mayan
languages that allows the rearrangement of the transitive relationship.
In what is called the \textit{referential voice} an NP that is normally in an
oblique relationship with a TV is promoted out of the oblique case and
is then treated as if it were one of the basic arguments in a transitive
predication (see Table \text{\textit{II}}). The kinds of oblique arguments that may be
promoted are datives, benefactives, malefactives, and possessors of Ps.
In those languages which have a referential voice, if one of these oblique arguments occurs in a transitive sentence it usually must be promoted. The TV remains transitive but has a referential voice suffix. The promoted argument is cross-referenced on the verb with absolutive marking as if it were the P. The P is not cross-referenced on the verb, nor does it receive oblique marking. Thus, it is demoted from having a direct relationship with the verb, but it is not demoted enough to receive oblique marking, rather it is left in limbo, or, using a term from relational grammar, it becomes a chômeur without status.

(9)  

<table>
<thead>
<tr>
<th>ACTIVE</th>
<th>REFERENTIAL VOICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TV A P oblique R</td>
<td>TV A R en chômeur P</td>
</tr>
</tbody>
</table>

(R = dative, benefactive, malefactive, and possessor of P)

Another voice related historically to the referential voice is the instrumental voice. The instrumental voice optionally allows an instrument NP to be promoted out of the oblique instrumental case; however, if it is promoted, it must also be put in focus. The instrumental voice is used when an instrument is questioned, relativized, or clefted.\(^{15}\) A TV in the instrumental voice remains transitive but has an instrumental voice suffix. Whether the instrument or the patient is cross-referenced on the verb in the instrumental voice depends on the language (see details in the section below on 'Voice and Ergativity in Selected Mayan Languages').
A General Characterization of Mayan Languages

In this section a brief and rather general characterization of Mayan languages is given. The discussion is not meant to be comprehensive, but rather is provided to give the reader a sense of what Mayan languages are like, especially with regard to verb structure and basic sentence structure.

Mayan languages are mildly synthetic and agglutination is the primary technique used in word formation. Prefixation is mostly restricted to person and tense/aspect inflections. Suffixation is the most common technique; almost all derivational affixes as well as many inflectional affixes are suffixes. Infixation occurs but is rare. Compounding is also a moderately important word formative process.

In general, Mayan languages are verb-first languages (see Table 4). The basic word order in the majority of the languages is Verb Patient Agent (VPA); however, VAP is the basic order in a number of languages, especially in the Kanjobalan and Mamean groups. AVP is recorded as the basic order in Chortí only, and is probably a rather recent innovation. All of the VPA languages, and some of the VAP languages have topicalization processes which front agents allowing them to occur before the verb; some of the languages also permit topicalized patients to occur before the verb as well. Thus, other possible (nonbasic) orders are AVP, APV, and occasionally, PVA. The latter is usually only permissible if the agent is higher on the animacy hierarchy than the patient (see the section above on 'Transitivity and the Notion of Subject.').
Hua and Tenejapa Tzotzil are reported to have both VAP and VPA as basic orders. VAP is the normal order when the agent and patient are equal on the animacy hierarchy and VPA is the normal order when the agent is higher on the hierarchy than the patient. Norman and Campbell (1978) suggest that this was the situation in Proto-Mayan (PM). In most languages PAV order is ungrammatical in normal active transitive sentences but does occur in some languages (e.g. Quichean) in focus antipassive constructions when the patient is topicalized while the agent is in focus.

It should be noted that what is meant by 'basic word order' is the order used in normal active declarative transitive sentences which are not subordinate and in which none of the major constituents (V, A, P) is topicalized or in focus. The basic orders are not necessarily the most frequent statistically, since in normal conversation one or the other of the major constituents is often the topic or theme, in which case the basic order is usually modified. The basic order is the one which strictly provides information about a predicate and its arguments without highlighting any one of them.

Mayan languages are morphologically ergative, since the pronominal affixes which cross-reference Ss of IVs are the same as those indicating the Ps of TVs, while, on the other hand, the As of TVs are cross-referenced with a different set of pronominal affixes. The affixes cross-referencing Ss of IVs and Ps of TVs are called absolutive (B) pronominal affixes (see Table 6), and are often referred to as 'Set B' in Mayan studies. The affixes cross-referencing As of TVs are called ergative (A) pronominal affixes (see Tables 7 and 8), and are often referred to as 'Set A' in Mayan studies. Compare the following examples from Tzotzil and Tzotzil. 16

20
(11) Tzt tal - em - ot 'you have come'
    como - perf- B2
(12) Tzt s - maj - oj - ot 'he has hit you'
    A3 - hit - perf - B2
(13) Tzt a - maj - oj - on 'you have hit me'
    A2 - hit - perf - B1
(14) Kek x - at - b'oe 'you went'
    asp - B2 - go
(15) Kek x - at - x - b'oq 'he called you'
    asp - B2 - A3- call
(16) Kek x - in - aa - b'oq 'you called me'
    asp - B1 - A2 - call

The ergative markers are always prefixes and usually have differing preconsonantal and prevocalic forms. Besides cross-referencing the agents of TVs, they also cross-reference possessors (e.g. Kek x-tz'i7 'his dog', aa-tz'i7 'your dog', r-aaq 'his pig', aaw-aaq 'your pig'). The absolutive markers are prefixes in some languages and suffixes in others, and sometimes occur as both prefixes and suffixes in the same language, usually in differing constructions. They also often function as independent pronouns, or at least as the formative bases on which the independent pronouns are built. Finally, they are used to indicate subjects in stative (nonverbal) sentences with, for example, predicate adjectives and predicate nouns (e.g. Chl winik on 'I am a man'). It should be noted that the third person singular absolutive marker is $\emptyset$ in Mayan languages (with nonnull morphological variants in Hua and Ch'ol only), and that inflection, especially tense/aspect inflection, is sometimes somewhat distinct with the third person singular null element.
Nouns in a direct relationship with a predicate (i.e. subjects of IVs or stative predicates, and agents and patients of TVs) are morphologically unmarked in Mayan languages (except for the cross-referencing on the verb). That is, there are no overt case marking affixes on nouns. However, nouns in an oblique relationship with a predicate are marked for case with either prepositions or relational nouns or a combination of both. Most Mayan languages have few prepositions but those that exist function like prepositions in European languages (e.g. Tzu *pa tinaamit* 'in town'). Relational nouns function like prepositions or case inflections but are formally nouns and are usually followed by their head noun and possessed by it (e.g. Tzu *w-umaal* 'by me', *aw-umaal* 'by you', *r-umaal jar aachi* 'by him the man = by the man', *ch-aaw-e* 'to you' < *ch* prep 'to, at' *aaw- E2 -e* 'to').

As the examples above indicate, possession in Mayan languages is normally indicated with an ergative prefix on the possessed noun, and if the possessor is third person, then it normally follows the possessed noun (e.g. Tzu *ruu-tz'ii7 jar aachi* 'his dog the man = the man's dog', *ruu-q'a7 jar iixoq* 'her hand the woman = the woman's hand'). Under possession some nouns undergo stem modification or omission of suffixes (cf. Tzu *tz'ii7* 'dog', *q'ab'-aaj* 'hand'). Possessor nouns may be omitted in context and fronted via topicalization processes.

Mayan languages have at least the following word classes, defined morphologically and syntactically: nouns, verbs, adjectives, and particles of various sorts. There are often other classes such as affect words, adverbs, numeral classifiers, and noun classifiers. Nouns are normally subcategorized depending on their behavior under possession. Verbs are always minimally sub-classified into transitive and intransitive,
and sometimes further into root transitives (RTV) and derived transitives (DTV); and rarely (Yucatecan only?) different kinds of intransitives are distinguished. TVs are always distinguished from IVs since they take both the ergative and absolutive person markers, while IVs only take the absolutive markers (or under special circumstances in some languages, only the ergative markers), never both sets at once.

The different classes of verbs are also usually distinguished by what modal suffixes they require. For example, in Quiché in the indicative mode RTVs require -o, DTVs -vj, and IVs -ik (cf. xub'an-o 'he did it', xukamisa-vj > xukamisaaj 'he killed it', xkam-ik 'he died').

There is also a special root class in Mayan called positional roots. These never occur in isolation but rather have special derived adjective and verb forms, as well as others. The roots get their name from the fact that they typically indicate the position, condition, state, or form that an object is in.

The structure of transitive verbs (see Table 5) in Mayan languages is typically of the form:

(17) aspect + absolutive + ergative + TV stem (+ mode)

Qui x-at-u-ch'ay-o 'he hit you'

or

(18) aspect + ergative + TV stem (+ mode) + absolutive

Lac t-inw-il-aj-ech 'I saw you'

Aspects minimally include completive and incompletion, and other common ones are progressive, potential or future (unrealized). What goes in the aspect slot, in some languages may actually be tense markers (e.g. past, remote past, present, future), or mode markers (e.g. imperative, optative). As noted above, the ergative person markers
are always prefixes while the absolutive markers may be prefixes or suffixes depending on the language. Mode suffixes include indicative, imperative, and subjunctive. In many languages, one or more verb classes may have zero marking in a given mode, and in some cases the mode suffix only appears when it would be in phrase final position. That is, it is omitted when anything follows in the same phrase or clause; this is especially true in the indicative mode. These suffixes are often called 'phrase final suffixes'. To a certain extent they have a dual function of indicating mode as well as phrase and clause boundaries.

Intransitive verbs conform to the structures above except that they always only have a single person marker, either a prefixed or suffixed absolutive marker; or, in some languages under certain conditions, an ergative prefix instead of an absolutive marker (see Tables 3 and 5).

Some languages distinguish between perfective and nonperfective verb structures, the ones above being the forms used in the nonperfective. Perfective verbs have no aspect prefix and a perfective marker occurs in place of a mode suffix. Cf.

(19) Tzu x - in - war - i 'I slept'
    compl-B1 - sleep - indic

(20) Tzu in - war - naq 'I have gone to sleep'
    B1 - sleep-perf

(21) Tzu x - at - nuu - ch'ey - Ø
    compl-B2 - A1 - hit - indic 'I hit you'

(22) Tzu at - n - ch'ey - oon 'I have hit you'
    B2 - A1 - hit - perf

In many languages, directional affixes, and at times, adverbial particles, may intervene between one or more of the position slots
outlined above. However, the details are too complex to go into here.

Stative predicates, such as predicate adjective and predicate noun constructions, rarely have a formal copula. Usually, the predicate is simply juxtaposed with an absolutive person marker (e.g. Tzu in aachi 'I am a man', ə ixog 'she is a woman').

Mayan languages always have a series of plain stops and affricates and a corresponding glottalized series (see Table 2). There are also series of fricatives, nasals, resonants and semivowels. Most languages have a five vowel system, and some have a corresponding set of long vowels. A few languages have a six vowel system (see Campbell, Fox, Kaufman, and McQuown, especially Kaufman 1966-69).

Some other typical features of Mayan languages are briefly mentioned. Usually, there is an existential/locative particle which is roughly equivalent to Spanish estar 'be (located)' and hay 'there is/are'. It is used to indicate the existence or location of something. It is also used in possessive sentences in the construction: exists my X, meaning 'I have X'. Plurality as an inflectional category on nouns is well developed only with human nouns, if at all. Noun and numeral classifiers are important in some of the languages. And finally, reflexive constructions are formed with transitive verbs having a possessed object noun meaning 'self', the possessor being cross-referenced with the ergative prefix on the verb, which also has the ə third person singular absolutive marker.

Voice and Ergativity in Selected Mayan Languages

In this section the voice systems of a number of Mayan languages are outlined, and split-ergative constructions are discussed, in those
languages which have them (the languages marked with an asterisk '*' in Table 1 are those under discussion in this section). The presentation will begin with Kek in the Greater Quichean group and proceed westward through other languages of the Eastern division, then move on to languages of the Western division, then northward to Yucatecan, and end with Huastec. At the beginning of each subsection the sources for each language are listed.¹⁷

Greater Quichean


Pcm: Smith-Stark (1976a, b, 1978a, b).


Cak: Brinton (1884), Larsen (field notes), Norman (1978), Townsend (1960).


Greater Quichean - Kekchi

TVs in the active voice are unmarked in the indicative. RTVs end in a consonant and DTVs end in a vowel. Some examples are given below and should be compared with forms in other voices discussed later on.
(23) x-at-in-sak' 'I hit you'
    T-B2-Al-hit

(24) x-∅-a-yok' li si7 'you cut the firewood'
    T-B3-A2-cut the firewood

(25) x-at-in-ch'iila 'I scolded you'
    T-B2-Al-scold

(26) x-∅-x-kamsi 'he killed it'
    T-B3-A3-kill

Simple passives with absolutive and rearranging functions are marked
with -e7 on RTVs, and vowel length plus an intransitive mode suffix (e.g.
-∅) on DTVs. An A may optionally be expressed in the sentence with a RN
(e.g. -b'aan 'by').

(27) x-at-sak'-e7 (in-b'aan) 'you were hit (by me)'
    T-B2-hit-Ps    A1-by

(28) x-at-ch'iila-a-k (in-b'aan) 'you were scolded (by me)'
    T-B2-scold-Ps-M    A1-by

There is also an 'impersonal' passive formed with -man which can
only have a 3rd person P and never allows an A to be expressed, although
it is implied.

(29) x-∅-yok'-man li si7
    T-B3-cut-Ps    the firewood

    'the firewood was cut/they cut the firewood'

Like most Mayan languages, Kek has past participial adjectives de-
derived from TVs, which function in passive-like constructions. These
participles form stative predicates requiring an absolutive S which
cross-references the P of the underlying TV. The participles emphasize the result of a transitive activity; that is, that the P is in the state or condition resulting from the activity. Agents may optionally be indicated by a RN. -b'il derives participles from RTVs, and -mb'il derives them from DTVs (e.g. sak'-b'il at 'you are/have been hit', ch'iila-mb'il at in-b'aan 'you are/have been scolded by me').

In the Chamil dialect, there is also a reduplicating suffix -C₁₀ which derives both adjective and IV passives from RTVs only (cf. x-in-aa-b'ak 'you tied me up', b'ak'-b'co-k in 'I am tied up', x-in-b'ak'-b'o 'I got tied up').

Antipassives are formed with -o on RTVs and -n on DTVs. These suffixes derive all three types of antipassives: absolute, focus, and incorporating. The P is demoted to the dative case marked with the RN -o 'to, of, for' in focus antipassive constructions. In incorporating antipassive sentences, a generic patient NP follows the verb without further marking.

**Absolutive**

(30) x-at-b'is-o-k

T-B2-measure-Ap-M

'you were measuring'

(< b'is 'measure' TV)

(31) x-in-ch'iila-n

T-B1-scold-Ap

'I was scolding'

**Incorporating**

(32) x-at-yok'-o-k si7

T-B2-cut-Ap-M firewood living

'you cut firewood (e.g. for a living)'

28
Focus

(33) laa7at x-at-sak'-o-k w-e
     'it was you who hit me'

(34) ani x-ŋ-sak'-o-k aw-e        'who hit you?'
     who T-B3-hit-Ap-M A2-to
     (cf. in active: ani x-ŋ-a-sak'    'who did you hit?)'
     who T-B3-A2-hit

(35) li wing li x-ŋ-kamsi-n r-e ...
     the man who T-B3-kill-Ap A3-to
     'the man who killed him...'

Morphological split-ergativity does not occur with finite verbs in Kek. However, there are passive infinitives of TVs formed in -b'al (e.g. sak'-b'al 'to hit'), which may take the ergative prefixes which then cross-reference the underlying P of these nominalizations (e.g. aa-sak'-b'al 'your being hit/for you to be hit'). The ergative prefixes, here are formally possessive prefixes that semantically mark patients. The passive infinitives are used in, for example, the progressive aspect marked with the stative predicate yoo- 'to be in the act of doing something', which requires an absolutive S that is coreferential with the Subject (A of TV or S of IV) of an infinitive embedded in a clause marked by the preposition chi 'at'.

(36) yoo-k in chi aa-sak'-b'al
     prog-M Bl at E2-hit-nom
     'I am hitting you'
It is difficult to decide if this is a true case of split-ergativity or whether the apparent split is due simply (and perhaps unimportantly) to the possessive nature of the ergative prefixes. However, since the construction does occur in the progressive aspect, I have indicated in Table 3 that it is a possible case of split-ergativity.18

Greater Quichean - Pocom

Pcm and Pch are essentially alike in terms of their voice categories and ergativity, so they will be discussed together. Differences will be noted when relevant to the discussion. The active voice is unmarked on RTVs, butDTVsv require a mood suffix in the indicative (i.e. Pch -\textit{v}j, and Pcm - \textit{v}h - \textit{e}j). IVs take the phrase final suffixes -\textit{ak} in Pch and -\textit{i} in Pcm.

(37) Pch x-at-in-ch'oy
     T-B2-A1-hit
     'I hit you'

(38) Pcm x-in-ru-q'at
     T-B1-A3-cut
     'he cut me'

(39) Pcm ih-\emptyset-ru-kaansa-ah
     T-B3-A3-kill-M
     'he killed it'

The simple passive is marked with -\textit{vr} on RTVs and -\textit{j} on DTVs. Agents may optionally be expressed with a RN (e.g. -u7\textit{uum} 'by').

(40) Pcm ih-\emptyset-q'at-ar-i (r-u7\textit{uum})
     T-B3-cut-Ps-M E3-by
     'it was cut (by him)'

(41) Pcm ih-\emptyset-kaans-j-i
     T-B3-kill-Ps-M
     'he was killed'
Pcm has another passive in -aam'j (e.g. ih-Ø-q'axp-aam'j-i 'it was split' < q'axpi 'to split'). Pch has a similar historically related form in -v\_17nj, which derives passives from TVs derived from positionals (e.g. Ø-Ø-oq'-o7nj-ik 'it was wrapped up' < oq'aa7 'to wrap up'). There are two other passives in Pch, one in -mv\_1j (e.g. x-in-ch'ey-nej-ik 'I was hit'). Brown states that As are not specified with this passive. The other passive, derived in -vb', has the meaning 'to be repeatedly Xed' (e.g. x-in-ch'ey-eb' ik 'I was repeatedly hit').

Both Pch and Pcm have past participles used in adjective passive constructions, marked with -ooj on RTVs (e.g. Pch ch'ey-ooj 'it is/has been hit'), and -v\_1maj on DTVs (e.g. Pcm kaans-amaj 'it is/has been killed').

Absolutive and focus antipassives are marked with -w on RTVs, and with -in in Pcm and -vn in Pch on DTVs. There apparently is no incorporating antipassive.

Absolutive

(42) Pcm x-in-k'at-w-i 'I burned/did some burning'

T-B3-burn-Ap-M

(cf. active: ih-Ø-nu-k'at 'I burned it')

T-B3-Al-burn

(43) Pcm ih-Ø-kaans-in-i 'he was killing'

T-B3-kill-Ap-M

In Pcm the focus antipassive has the functions listed above in the section 'Voice' and an additional one: it also occurs when the A is a negative indefinite pronoun. In both Pch and Pcm the focus antipassive is optional; and in Pcm it is used only with 3rd person As, especially
when ambiguity would occur if it were not used. Ps are optionally de-
moted to the dative case marked with the RN -eh 'to, for, of'. However, in Pcm when the P is not demoted, it is impossible to tell whether the A or P is marked on the verb, since with 3rd persons the S is always the Ø absolutive. Examples of the focus antipassive with clefted, questioned, and negative As are given below.

(44) Pch re7 tz'i7 x-Ø-hoy-w-ik r-e winaq
    the dog  T-B3-bite-Ap-M A3-to man
    'it was the dog that bit the man'

(45) Pch re7 hin x-in--b'uyu-n-ik r-eh
    the I  T-B1-quiet-Ap-M A3-to
    'I am the one who quieted him down'

(46) Pcm ha7 wach ih-Ø-tok-w-i  (r-eh)
    who  T-B3-hit-Ap-M B3-to
    'who hit him?'
    {cf. with 1st per: ha7 wach ih-Ø-nu-tok
    'who did I hit?'}

(47) Pcm qu ham' wach Ø-Ø-to7-w-i  r-eh
    nobody  T-B3-help-Ap-M A3-to
    'nobody helped him'

Both Pocom languages have an instrumental voice, marked with
-b'e in Pch and -b'e -7e in Pcm. The addition of this suffix results
in a DTV whether or not it is attached to a RTV or DTV. In both lan-
guages the instrumental voice is used in order to put the instrument
NP in focus (see section on 'Voice'). The instrument is promoted out
of an oblique case marked with a preposition or RN, and is clefted.
In Pcm the P is optionally demoted to an oblique case marked with a RN. Patient demotion apparently does not occur in Pch, but since the examples available for Pch are only in the 3rd person it is impossible to tell whether the P or the instrument is marked on the verb. The examples below are all from Pcm.

(48) hila7 x-∅-a-q'at-7i-eh w-ihchin (N.B. -7e+Vh
what T-B3-A2-cut-I-M Al-to > -7ich)
'what did you cut me with?'

(49) hila7 x-in-a-q'at-7i-eh (cf. with (48))
what T-B1-A2-cut-I-M
'what did you cut me with?'

(50) ma7 ha7 ih-∅-nu-muh-7i-eh r-o
the water T-B3-Al-wet-I-M A3-to
'with the water I wet it'

Notice in (48) and (50) that the instrument is promoted to absolutive and marked on the verb as B3, and the P is demoted to the dative case. In (49) the instrument is promoted out of an oblique case but the P is not demoted. In Pcm, the instrumental voice suffix may also be used optionally in a sentence with an instrument in an oblique case, without instrument promotion or patient demotion.

(51) hin \{ih-∅-nu-sir \} ma7 xuut pech r-iij ak'ach
\{ih-∅-nu-sir-7i-eh\}
I T-B3-Al-paint-I-M the jug with A3-back feather
'I painted the water jug with a feather'
And it can be used in sentences where the P is demoted, but the instrument is not promoted!

(52) ih-∅-nu-mek-7i-eh r-eh ma7 chie7 pech ma7 maachit
T-B3-A1-pull-I-M A3-to the tree with the machete
'I pulled the tree with the machete'

These last two examples seem to indicate that the instrumental voice in Pcm is in a state of flux and is perhaps breaking down, as it has in Cak (see the next section on 'Quichean Proper').

Both Pcm and Pch have accusative marking in the incompletive aspect marked with the proclitic na in Pch and nV in Pcm. And Pch has accusative marking with the progressive particle k'ahchiʔ. In both cases the ergative prefixes have been extended to IVs in these aspects because the IVs are really action nominalizations derived with nominalizing infinitive suffixes. They, therefore, require the ergative possessive prefixes. These nominalizations are subordinate to the aspect particles, which Brown and Smith-Stark claim are higher predicates.

(53) Pcm nu ru-wir-iik 'he sleeps'
T A3-sleep-infin

(54) Pcm nu ∅-ru-q'at-om 'he cuts it'
T B3-A3-cut-infin

Greater Quichean - Quichean Proper

Qui, Tzu, and Cak¹⁹ are discussed together, because in general their voice systems are similar. Differences will be noted as they occur.
The active voice in the indicative on RTVs is unmarked in Tzu and Cak but is marked with the phrase final suffix \(-o\) in Qui. DTVs have the indicative mode suffix \(-v\) in Qui and Tzu, and \(-j\) in Cak. IVs are marked with \(-ik\) in Qui, \(-i\) in Tzu, and unmarked in Cak. Examples of the active voice are given below.

(55) Qui x-in-a-ch'ay-o 'you hit me'
Tzu x-in-aa-ch'ey
Cak x-in-a-ch'ay
T-B1-A2-hit (-M)

(56) Qui x-in-a-tzuku-uj 'you looked for me'
Tzu x-in-a-kaano-oj
Cak x-in-a-kano-j
T-B1-A2-search-M

The simple passive with both absolutive and rearranging functions is marked with \(-x\) on DTVs, and with the infix \(-i\) in Tzu on RTVs, which has become vowel length \(-v\) in Qui, and in Cak has either become zero or vowel ablaut. In the Milpas Altas dialect of Cak the suffix \(-e7\) may optionally mark passive on RTVs, especially where zero marking would otherwise result.

(57) Qui x-in-ch'ay-ik 'I was hit'
Tzu x-in-ch'ejy-i
Cak x-i-ch'ay(-e7)
T-B1-hit(Ps)-(M)

(58) Qui x-in-tzuku-x-ik 'I was looked for'
Tzu x-in-kano-x-i
Cak x-i-kano-x
T-B1-search-Ps(-M)

35
Agents may be expressed with RNS (e.g. Tzu \textit{xinch'e}j\textit{y} aw-\textit{umaal} 'I was hit by you'), but in Qui expressed As are restricted to 3rd person.

All three languages have a 'completive' passive in \textit{-(V)taj} used on both RTVs and DTVs. This passive emphasizes the result of the activity and/or its termination. Agents may be optionally expressed in any person.

(59) Qui \textit{x-in-tzuku-taj aw-umaal} 'I was finished being looked for by you'

T-B1-search-Ps A2-by

All three languages have a passive in \textit{-Vr} which is restricted in use to only a handful of RTVs and seems to be identical in function to the simple passive discussed above (e.g. Cak \textit{x-g-k'ang-\textit{ar}} 'it was taken). Dayley (1978) has called this the 'archaic' passive because of its restriction in use and because it is cognate with the simple passive \textit{(-Vr)} in Pocom.

Tzu has a nonproductive medio-passive formed with the infixes \textit{-7-} or \textit{-j-}, which occurs only on a dozen or so roots (e.g. \textit{x-g-k'i7s-i} 'it stopped' < RTV \textit{k'is} 'stop'; \textit{x-g-tzajg-i} 'it fell down' < RTV \textit{tzag} 'lose').

Adjectival passives based on past participles occur in all three languages. They are formed with \textit{-oon} (> \textit{-on} in Cak) on RTVs and \textit{-Vn} (> \textit{-n} in Cak) on DTVs. Agents may optionally be expressed with these forms (e.g. Tzu \textit{at Ch'ey-oon w-umaal} 'you are/have been hit by me', \textit{in kaano-on f-umaal} 'I am/have been looked for by him'). Another adjectival passive occurs, in at least Qui and Tzu, that is related to the \textit{-(V)taj} passive. It is built on \textit{-(V)tal}. This adjectival passive emphasizes the result of the activity and deemphasizes the activity itself (e.g. Qui \textit{in kuna-tal-ik} 'I am cured').
The absolutive antipassive is formed with -oon (> -on in Cak) on RTVs and -Vn (> -n in Cak) on DTVs.

(60) Tzu x-at-ch'ey-oon-i 'you were hitting/did
    T-B2-hit-Ap-M some hitting'
    x-at-kaano-on-i 'you were searching'
    T-B2-search-Ap-M

It should be noted that the absolutive voice is not used with all TVs. There are a few which apparently, because of their semantic features, cannot be put into absolutive voice. For example, in Qui the TVs meaning 'take out' and 'cry over' do not have absolutive antipassive forms. There are also a few derived IVs that are formally like absolutive antipassives; however, they have been lexicalized so that their Ss refer to the P of the underlying TV not the A. For example, in Qui x-Ø-raqi-n-ik (< DTV raqi 'to break apart') means 'it broke apart' not the expected 'he broke (something) apart'. It is also noteworthy, that in Qui (but not in Tzu or Cak), the P may be expressed with the dative relational noun in absolu-

tive antipassive sentences.

(61) ka-Ø-log'o-n  lee in-tat     ch w-eh
    T-B3-love-Ap  the my-father  to me-to
    'my father loves (to) me'
    cf. active:  k-in-u-log'o-oj  lee in-tat
                 T-B1-A3-love-M  the my-father
    'my father loves me'

Mondloch and Hruska (1975) say this construction is used to disambiguate possible ambiguities. Norman and Larsen (1979) say that absolutive
sentences with patients expressed, have a slightly different meaning than their active counterparts. For example, the active form x-in-u-ch'ay-o means 'he hit me', whereas the absolutive sentence x-2-ch'ay-on ch q-eh means something like 'he hit at me'. In the latter case, the P in the dative case is in some sense only marginally a P, since it isn't necessarily affected by the activity. The absolutive voice with Ps demoted to the dative case may, then, be a means for expressing the distinction between true Ps and marginal ones in Quich (in the same way that prepositions apparently function in English; i.e. 'I hit him' vs. 'I hit at him').

Focus antipassives are marked with -ow (> -o in Cak) on RTVs, and -vn (> -n in Cak) on DTVs. Verb agreement in the focus antipassive is rather interesting in Quichean. Normally, the absolutive person marker on the verb cross-references whatever is higher on the person hierarchy (cf. the section above on 'Transitivity and the Notion of Subject')., whether it is the A or the P. Compare the Tzu examples below.

(62) Tzu Inin x-in-ch'ey-ow-i jar aachi
     I T-Bl-hit-Ap-M the man
     'I was the one who hit the man'

(63) Tzu Jar aachi x-in-ch'ey-ow-i
     the man T-Bl-hit-Ap-M
     'the man was the one who hit me'

There is no problem with ambiguity in these cases because the A is always clefted and occurs immediately before the verb. The P may either occur after the verb or be fronted via topicalization, in which case the P occurs first in the sentence. Sentences in the focus antipassive with a topicalized P are the only instances in Quichean with
P A V word order. Compare the following examples.

(64) Tzu jar aachi x-w-ch'ey-ow-i jar iixoq
    the man T-B3-hit-Ap-M the woman
    'the man was the one who hit the woman'

(65) Tzu jar iixoq x-w-ch'ey-ow-i jar aachi
    the woman T-B3-hit-Ap-M the man
    'the woman was the one who hit the man'

(66) Tzu jar iixoq jar aachi x-w-ch'ey-ow-i
    the woman the man T-B3-hit-Ap-M
    'the woman, the man was the one who hit her'

In Qui one of the arguments, A or P, must be 3rd person in focus antipassive sentences. However, in Tzu the non-3rd persons may occur, and when they do the verb always agrees with the A and the P occurs in the genitive case marked with the RN -vxiin 'of, for'. Also, in Tzu the person hierarchy can be overridden by using -vxiin to mark the P.

(67) Tzu Inin x-in-ch'ey-o aw-xiin
    I T-B1-hit-Ap A2-of
    'I was the one who hit you'

(68) Tzu Je7ee7 x-ee-ch'ey-o w-xiin
    they T-B3p-hit-Ap Al-of
    'they were the ones who hit me'

Tzu has an incorporating antipassive marked with 'short' forms of the focus antipassive, -o for RTVs and -n for DTVs, which normally occur when indefinite NPs follow. The A is always clefted in the incorporating antipassive, like in the focus antipassive, and a generic P immediately follows the verb.
(69) Tzu Je7ee7 x-ee-ch'ey-o winaq
they T-Br3-hit-AP people
'they are the ones who hit people'

Instrumental voice is marked with -b'e in Quichean. This voice behaves differently in the three languages, however. In Qui and Tzu the instrument must be put in focus by clefting in the instrumental voice. In Qui the instrument is promoted to absolutive and the P is demoted to the genitive case marked by the RN -ee(ch).

(70) Qui ch'iich' x-Ø-in-sok-b'e-j aw-eech
machete T-B3-Al-wound-I-M A2-of
'it was a machete that I wounded you with/
I used a machete to wound you'

In Tzu the instrument is promoted out of the instrumental case but it is not advanced as far as the absolutive; the P remains in the absolutive.

(71) Tzu machat x-in-r-choy-b'e-ej
machete T-B1-B3-cut-I-M
'it was a machete that he cut me with'

In Cak when -b'e is used, the instrument is not promoted at all, although the instrument along with its RN case marker may be fronted (i.e. topicalized).

(72) Cak t-ik'in jun machat x-i-ru-sok-b'e-j
A3-with a machete T-B1-A3-wound-I-M
'with a machete he wounded me'
(73) Cak x-∅-u-choy(b'e-j) r-ik'in jun machāt
     T-B3-A3-cut(-I-M) A3-with a machete
 'he cut it with a machete'

Thus, in Cak -b'e no longer has voice changing power; the basic trans-
sitive relationship is not disrupted when it occurs. Rather it is used
simply to optionally indicate that there is an instrument in the sentence.

The instrumental voice may be used in conjunction with the passive.
Examples are given in (74). These examples also illustrate the different
degrees of advancement of the instrument NP in Qui, Tzu, and Cak, res-
pectively. In Qui the instrument is clearly advanced to the absolutive
since it is cross-referenced on the verb while the P is in an oblique
case. In Tzu the instrument is extracted out of the oblique case but
is not advanced to absolutive, since the P is cross-referenced on the
verb with the absolutive prefix. In Cak the instrument is fronted (i.e.
topicalized) but is neither extracted from an oblique case nor advanced.

(74) Qui ch'iich' x-∅-sok-b'e-x aw-eech
     machete T-B3-wound-I-Ps A2-of
 'a machete was used to wound you'

Tzu machat x-in-choy-b'e-x-i
     machete T-B1-cut-I-Ps-M
 'it was a machete that I was cut with'

Cak r-ik'in jun machāt x-i-sok-b'e-x
     A3-with a machete T-B1-wound-I-Ps
 'with a machete I was wounded'

It is interesting to note that the three Quichean languages taken
together reflect the rather complex situation in Pcm with the
instrumental voice.

There are no morphological accusative structures in Quichean proper, although there are passive infinitives of TVs that may take the ergative possessive prefixes which reference semantic Ps. These nominalizations do not function in any paradigmatic way in the tense/aspect systems of Quichean, as they did, for example, in Kek and Pocon.

Mamean


Ixl: Ayres (1979), Elliot and Elliot (1966).

Mamean - Mam

Relative to other Mayan languages, Mam is rather distinct in two ways. First, some of the pronominal affixes have conjunct forms, and the second and third persons are merged to a certain degree (cf. Tables 6, 7, and 8, and England for details). Second, active transitive verbs almost always require directional (D) particles within the verb complex. Some active TVs may occur without Ds but it is uncommon. The Ds are similar in function and meaning to directional particles used with many verbs in English (e.g. up, down, in, out, away, at, etc.), and like English, many of the particles are lexically determined. That is, all Ds do not occur with any TV, and usually one or two particular Ds always occur with a given TV. In many cases, the Ds do not add any significant meaning to the TV, they are simply (almost) obligatory

42
elements. All of the Ds are derived from IVs of motion; they are usually slightly modified, short forms of the IV root (e.g. xi 'away' < xi7 'go', ok 'in' < ook 'enter'). Whenever there is a directional element in the TV complex, the verb must always have the suffix -7n (~ -7-...-n ~ -VVn). Compare the examples below.

(75) ma ʒ-t-tsuy
  T B3-A3-grab
  ma ʒ-tzaj t-tzy-u7n
  T B3-D A3-grab-suf
  ma chin t-tzeq'a-ya
  T B1 A..hit-..2
  ma chin-ok t-tzeq'a-n-a
  T B1-D A..hit-suf-..2
  ma ʒ-ky-tx'aj
  T B3-A3p-wash
  o ʒ-kub' n-b'y-o7n n-man
  T B3-D Al-hit-suf Al-father

Mam has a number of passives, but they are not well distinguished functionally in the literature (cf. England 1979, with England et al, 1979). The simple passive seems to be marked with -eet (~ -at ~ -t).

There are conflicting reports on whether or not As may be expressed, and in any event, no examples are given.

(76) ma ʒ-tzy-eet
  T B3-grab-Ps
  ma cyi-txj-eet
  T B3p-wash-Ps
Other passives occurring in Mam are: (a) an unmarked 'impersonal' passive used only with 3rd person Ps, and which does not allow As to be expressed; (b) one in -b'aj which requires a directional; (c) one in -i -l used on RTVs only, which emphasizes the result of the activity, and allows an expressed A, but if one is not expressed, then England says that it implies there was no A; (d) one in -(VV)njtz which allows an expressed A, but like the preceding passive, if one is not expressed, then it implies there was no A; (e) participial passives marked by -7n (~ -7...-n . -Vvn).

(77) ma ƍ-tzuy

T B3-grab

ƍ tz-ul qu'a-b'aj

T B3-D leave-Ps

ma ƍ-ku7x yuup-ju (w-u7n-a) 'it want out/it was put out

T B3-D put out-Ps (Al-by)

fire

ma ƍ-yuup-anjtz (t-u7n

T B3-put out-Ps (Al-by)

'it was put out by itself

(by him)' SI

ƍ txaqo-7n

'T it is/has been lit' SI

B3 light-Pp

All antipassives in Mam are marked with the suffix -(V)jn.

The absolutive voice has two functions, one for omitting Ps, the other allowing a P to be expressed in the sentence, usually with the topic RN -i71 'about'. Its primary function, apparently, is to allow the expression of a transitive activity without a directional particle.
(78) ma cyin-txj-on 'I washed' TS
T B1 - wash-Ap
ma cyi-txj-on t-17j mes 'they washed the table' TS
T B3p-wash-AP A3-RN table

Apparently, there is an incorporating antipassive function which allows generic nouns to occur in the sentence without further marking.

(79) ma cyi-txj-on mes 'they washed tables' TS
T B3p-wash-AP table
ma chi-tzy-uun xiinaq cheej
T B3p-grab-AP man horse
'the men caught horses' SI

There is also a focus antipassive in which the A is clefted. If an overt 3rd person P occurs in the sentence, it is marked with the dative RN -e(o) 'to'. However, the focus antipassive verb has the 3rd person absolutive marker cross-referencing the P anyway; that is, even though the overt patient NP is in the dative case. For example, in (80) the overt patient NP is in the dative case, the verb has the Ø 3rd person absolutive marker cross-referencing the P anyway, and the A is not cross-referenced on the verb at all. In (81), there is no overt patient NP, but the verb is inflected with cyi-, 3rd person plural absolutive, cross-referencing the P. Thus, in Mam, a verb in the focus antipassive does not agree with the A, but does agree with the P, even if it is overtly marked by a RN in the dative case.

(80) na7yan e Ø-kub' b'y-on t-e n-man
I T B3-D hit-Ap A3-to A1-father
'I was the one who hit my father' TS

45
(81) na7yan ma cyi-kub' b'y-on
    I T B3p-D hit-Ap
    'I was the one who hit them' TS

(82) aal tzun ma 6-kub' paat-an t-e jaa-7
    who T B3-D burn-Ap A3-to house
    'who burned the house down?' TS

(83) Xwan ma 6-kub' paat-an t-e jaa-7
    Juan T B3-D burn-Ap A3-to house
    'John was the one who burned the house down' TS

In certain kinds of subordinate clauses, and after affect words
and some adverbs (which may also be subordinating), the ergative pre-
fixes are also used to mark Ps in these subordinate clauses, at least
when there is a directional (no examples have been found without
directional). Cf.

(84) aj t-kan-eet ... 'when it was found...' 
    when A3-find-Ps
    na7xtzan t-ex q-laq'o-7n 'we still haven't bought it'
    still not A3-D Alp-buy-suf

Since the directionals are ultimately derived from IVs, and since
patient markers are always prefixed to the directionals, it is difficult
to say whether the ergative prefixes have been only extended to Ss of
IVs or to Ps of TVs as well. In any case, these constructions are
examples of extended ergativity, and perhaps of accusative marking.
It should be noted that if the ergative prefixes are extended to both
Ss of IVs and Ps, then this is not a case of accusative marking, since
all three arguments in these constructions are marked with ergative
prefixes.
Mamean - Aguacatec

Agt and Ixl will not be discussed in the same detail as Mam. The various voice categories will simply be noted and examples will be given.

Examples of TVs in the active voice in Agt are given below. Note that there is a distinction between RTVs and DTVs.

(85)   ja kxh n-b'iy  
       T  B2 Al-hit

ja 7n-qatz  
T  B3 Al-tear

ja 7n-kyima-aj  
T  B3 Al-kill-M

The simple passive is marked with -lij (-chij -wij) on RTVs and -ij on DTVs. Agents may optionally be expressed with the RN -a?n 'by'. RTVs also have another passive in -xij which seems to emphasize the fact that there was an agent present in the activity (more than with the simple passive), regardless of whether or not it is expressed in the sentence. Participial passives are marked with -ij on RTVs and -7n on DTVs.

(86)   ja ø-b'iy-lij  (w-a?n)  
       T  B3-hit-Ps  Al-by

ja ø-b'iy-xij  (t-a?n)  
T  B3-hit-Ps  A3-by

b'iy-ij-chin  
hit-Pp-B1

'I hit you'

'I tore it'

'I killed it'

'he was hit (by me)'

'he was hit (by him)'

'I am/have been hit'

47
ja ʊ-kyims-ij  'he was killed'
T B3-kill-Ps

kyims-a7n-ʊ  'he is/has been killed'
kill-Pp-B3

There is also a nonproductive medio-passive occurring with a few RTVs, which is indicated by lengthening the root vowel; historically (perhaps synchronically ??) the lengthened vowel comes from an infixed -h- (e.g. ja ʊ-gaatx 'it tore' < gatx RTV 'tear').

The absolutive antipassive voice is marked with -oon on RTVs and -Vn on DTVs. The root vowel of most RTVs is lengthened in this construction. Patients may be expressed in the absolutive with either of the RNs tz-eetz 'to' or tz-e7 'with, about'; however, there is a slight meaning change from that of the active voice.

(87) ja chin-b'iy-oon  'I fought/ was hitting'
T B1-hit-Ap

ja chin-b'iy-oon tz-aw- eetz to-A2- to  you'
T B1-hit-Ap

ja chin-kyimsa-an  'I killed'
T B1-kill-Ap

The focus antipassive is also marked with -oon/-Vn but there is no lengthening of a root transitive vowel. In this voice the A is clefted and the verb agrees with the P if it is non-3rd person, and it may agree with either the A or the P if the P is 3rd person. Note that focus antipassives require the dependent tense/aspect markers (e.g. m - n dependent proximate past = ja main proximate past). This probably indicates that the clefted A is in a higher clause and is a predicate noun.
(88) in n-xh-b'iy-oan  'I was the one who hit you'
    B1 T-B2-hit-Ap
 in m-ŋ-kyimsa-an  'I was the one who killed him'
    B1 T-B3-kill-Ap
 ~ in m-in-kyimsa-an
    B1 T-B1-kill-Ap

Extended ergative accusative structures occur in Agt in dependent
time adverbial clauses in the indefinite past tense (with null marking).
The verb forms, however, in these cases are nominalizations and therefore
the ergative prefixes cross-referencing underlying Ss of IVs are formally
possessors. Cf.

(89) ye a-b'iy-oal in...  'when you hit me...'
    art A2-hit-nom B1
 ye a-kyim-e7n...  'when you died...'
    art A2-die-nom

Mamean - Ixil

Examples of TVs in the active voice in Ixil are given in (90). Note
that there is a distinction between RTVs and DTVs.

(90) kat a-q'os in  'you hit me'
    T A2-hit B1
 kat t-echb'u ŋ  'he ate it'
    T A3-eat B3

DTVs have a passive in -l. RTVs have passives in -ax (- -vx),
-p, and -l - -ch, which Ayres does not distinguish further. Agents
may be expressed with the RN -a7n 'by' with at least some of these
passives. Participial passives are marked with -el on RTVs and -mal on
DTVs.
(91) kat q'os-ax-i ø (t-a7n) 'it was hit (by him)'
    T hit-Ps-M B3 (A3-by)
kat q'os-p-i ø  'it was hit'
T hit-Ps-M B3
kat q'os-l-i ø  'it was hit'
T hit-Ps-M B3
q'os-el ø  'it is/has been hit'
hit-Pp B3
kat echb'u-l-i ø  'it was eaten'
T eat-Ps-M B3
echb'u-mal ø  'it is/has been eaten'
eat-Pp B3

Absolutive and focus antipassives are marked with -on on RTVs and
-n on DTVs. Patients may be expressed optionally in the absolutive with
either of the RNS -ø 'to' or -i7 'for, to'. It is not clear how abso-
lutives with expressed Ps differ from the active voice.

(92) kat q'os-on axh (s w-i7) 'you hit (me)'
kat echb'u-n o7  'we ate'
    T eat-Ap Blp

With focus antipassives the A is clefted and the verb agrees with
the P.

(93) in kat q'os-on axh  'I was the one who hit you'
    B1 T hit-Ap B2

Unlike Nam and Agt, Ix1 has an instrumental voice in b'e. In the
instrumental voice the instrument is promoted out of an oblique case
marked with a RN like -ucy' or -a7n, and it is clefted to the front of
the sentence. Depending on the speaker, the P may either be demoted to an oblique case marked with the RN _i7, or it may remain in the absolute. As in all Mayan languages with _b'e, the passive and instrumental voices may occur together. Compare the following sentences.

(94)  n-in-tzok'  Ø tze7 t-ucy'  machit
T-Al-cut   B3 tree A3-with machete
'I am cutting the tree with a machete'

(95)  machit n-in-tzok'-b'e  Ø tze7
machete T-Al-cut-I   B3 tree
'with a machete I cut the tree'

(96)  u  machit kat tzok'-ox-b'e  Ø  u  tze7  (s  w-a7n)
the machete T  cut-Ps-I   B3 the tree   (Al-by)
'with a machete the tree was cut (by me)'

(97)  ma7l  u  tze7  Ø-i-q'os-b'e  Ø  u  naj  t-i7  u  ixoj
one the stick T-B3-cut-I   B3 the man A3-to the woman
'with a machete the man hit (to) the woman'

(98)  uula  Ø-a-k'oni-b'e  in
sling-shot T-A2-shoot-I   Bl
'with a sling-shot you shot me'

The ergative prefixes are extended in use to cross-reference Ss of IVs in Ix1 in the progressive aspects, in certain dependent clauses without tense/aspect markers, with certain beginning adverbs without tense/aspect markers, with the question words for 'how' and 'why', and in one type of imperative in the Chajul dialect. In all of these cases the IV must occur with the suffix _e7, which is also the incomplete transitive phrase final suffix. It is noteworthy that focus anti-passives do not take ergative prefixes like other IVs in the
progressive aspects; rather they are inflected with the absolutive markers. An example of extended ergative constructions is given below in the progressive aspect marked with in.

(99) in w-il-c7 $\theta$

$T$ Al-see-M B3

in w-ul-e7

$T$ Al-arrive-M

Greater Kanjobalan

Chj: Hopkins (1967), Lengyel (notes), Maxwell (1976, m.s.), Williams and Williams (1966).


Greater Kanjobalan - Chujoan

In Chj, in the indicative RTVs take the phrase final suffix -a7, DTVs take -e^j, and IVs take -i. Examples of TVs in the active indicative are given in (100) and (101).

(100) tz-in-y-il-a7

$T$-Bl1-E3-see-M

tz-oii-s-mak'-a7

$T$-Blp-E3-hit-M

(101) 7ix-6-s-mak' waj Xun 7ix Malin

'Mary hit John'

$T$-B3-A3-hit he John she Mary

Chuj has two passives, one in -aj (- -chaj) and another in -ax (- -max). Maxwell says these are essentially the same. Lengyel claims

52
that -aj places emphasis on the fact that there is an agent involved while -ax does not. Compare the examples below with the RN y-uj 'by', on account of'.

(102) 7ix-Ø-mak'-chaj waj Xun (y-uj 7ix Malin)
T-B3-hit-Ps he John A3-by she Mary
'John was hit (by Mary)'

(103) no wakx 7ix-Ø-mak'-chaj (> xma7chj) Y-uj Petul
the cow T-B3-hit-Ps A3-by Peter
'the cow was hit by Peter'

(104) no wakx 7ix-Ø-mak'-ax (> xma7x) y-uj Petul
the cow T-B3-hit-Ps A3-on account of Peter
'the cow was hit on account of Peter'

There is also a participial passive adjective formed with -b'il (e.g. mak'-b'il 'he is/has been hit').

Chj has an incorporating antipassive in -w that allows indefinite NPs to be incorporated into the verb complex.

(105) 7ix-ach-mak'-w-i 7anima 'you hit people'
T-B2-hit-Ap-M people

There is an absolutive antipassive in -waj which optionally allows Ps to be expressed obliquely in a prepositional phrase with t'a 'to'.

Note the difference in meaning between active sentences and absolu-
tive sentences with expressed Ps.

(106) 7ix-Ø-mak'-waj 7ix Malin (t'a waj Xun)
T-B3-hit-Ap she Mary to he John
'Mary did some hitting (to John)'

A focus antipassive occurs in -an. It is not clear how verb
agreement works in this voice in Chj, since the only available examples
are in the 3rd person.

(107) ha 7ix Malin 7ix-Ø-mak'-an waj Xun
cleft she Mary T-B3-hit-Ap he John

'It was Mary who hit John'

Chj has accusative marking in the progressive aspect marked with wan+. When wan+ is used with IVs their Ss are cross-referenced with the ergative prefixes. In addition, when wan+ is used with TVs, they are (pseudo)intransitivized with the focus antipassive suffix -an. They, therefore, require the IV phrase final suffix -i. Nevertheless, their As are cross-referenced with the ergative prefixes and their Ps with the absolutive prefixes. In other words, wan+ makes IVs behave somewhat like TVs, and TVs somewhat like IVs. Cf.

(108) wan k-olu7maj-i 'we are getting dirty'
    T  Alp-get dirty-M
wan Ø-k-aw-an-i 'we are planting it'
    T  B3-Alp-plant-Ap-M

Toj is briefly mentioned. In the indicative TVs take the phrase final suffix -a and IVs -i(y). Like Chj, Toj has a passive in -j which seems to place emphasis on the fact that there is an agent involved, and another in -x which does not emphasize agent involvement. There is a participial passive in -ub'al. Cf.

(109) x-Ø-j-mak'-a 'I hit him'
    T-B3-Al-hit-M
Ø-mak'-j-iy-on (y-uj) 'I was hit (by him)'
    T-hit-Ps-M-B1
Ø-mak'-x-iy-on (y-uj) 'I was hit (by him)'
    T-hit-Ps-M-B1
mak'-ub'al Ø  'he is/has been hit'

hit-Pp  B3

There is an absolutive antipassive in -wan - wun (e.g. mil-wan-
'kill' IV < mil- TV and te7-wun 'wound' IV < te7- TV). From the examples
in Furbee-Losee, there do not seem to be focus or incorporating anti-
passives. However, there is an intransitivizing suffix -Vn which
apparently is cognate with focus antipassives in nearby languages
(e.g. nul-in 'suck, chew' < nul- 'suck on, eat' TV). There is also
an intransitivizing infix -j- which is probably cognate with the
Quichean passive and Agt medio-passive (cf. Toj majk'- 'hit' IV <
mak' TV, nijk- 'tremble' < nik- 'stir' TV, kujch- 'be carried' < kuch-
'carry on the back' TV). From the available examples it is difficult
to tell the precise functions of -Vn and -j-.

There apparently are no accusative structures in Toj.

Greater Kanjobalan - Kanjobalan Proper

From this group Jac will be discussed in some detail. In the active
indicative RTVs have the phrase final suffix -a, and DTVs -a; IVs have
-(y)₁ in the indicative (n.b. the enclitic an occurs finally in a clause
with a 1st person). In the irrerealis (including future) RTVs have the
phrase final suffix -a7 (≈ -o7 - u7) and DTVs -7, while IVs have -oj.
Examples of TVs in the active voice are provided in (110).

(110) ch-in ha-maq'an  'you hit me'
   T-B1  A2-hit  1st

(x)-Ø-s-maq' naj ix   'he hit her'
   T-B3-A3-hit  he she
x-Ø-a-maq'-a  'you hit it'
T-B3-A2-hit-M

ch-Ø-aw-il-a  'you see it'
T-B3-A2-see-M

ch-ach-w-il- an  'I see you'
T-B2-Al-see 1st

x-Ø-y-il-a  'he saw it'
T-B3-A3-see-M

(n.b. x- ~ (x)k= = complete; ch- = incomplete)

Jac has several passives which Craig distinguishes in the following ways: (a) -ot: can not occur in the future; usually does not occur with an expressed A, and when it does only with a 3rd person A; however, As are not necessarily presupposed; (b) -lax: same as -ot except that it can occur in the future and presupposes an A; (c) -lo: normally occurs with an expressed A in any person; usually used in negative or restrictive senses; is not 100 percent productive; (d) -cha: may occur with an expressed A in any person; implies the P is helpless; is not 100 percent productive; and (e) -b'il: participial adjective passive, which may take an expressed A.

(lil)  x-Ø-maq'-ot ix (y-u  naj)  'she was hit (by him)'
T-B3-hit-Ps she (A3-by he)

x-in-maq'-ot an  'I was hit'
T-B1-hit-Ps 1st

ch-in-il-lax-oj (y-u naj)  'I'll be seen (by him)'
T-B1-see-Ps-M

matxa x-Ø-7il-la (w-u) an)  'it was not seen yet not yet T-B3-see-Ps (Al-by 1st)  (by me)'

56
ch-ach-kil-cha (w-u an) 'you are helped (by me)'
T-B2-help-Ps (Al-by 1st)
tz'un-b'il Ø (w-u an) 'it is/has been planted
plant-Pp B3 (Al-by 1st) (by me)'

Jac has an incorporating antipassive in -w and an absolutive antipassive in -wa.

(112) ch-ach-il-w-i anima 'you people-watch'
T-B2-see-Ap-M people
ch-ach-il-wa-yi 'you watch'
T-B2-see-Ap-M

There is a focus antipassive in -n which is normally used only with 3rd person As. The absolutive prefix on the verb cross-references the P not the clefted A. This indicates that the clefting function of the focus antipassive raises the A into a higher clause making it a predicate noun.

(113) ha7 naj x-Ø-maq'-n-i ix 'it is he who hit her'
cleft he T-B3-hit-Ap-M she
mak k-ach ʔil-n-i 'who saw you?'
who T-B2 see-Ap-M

There is another construction involving the suffix -n, which combines focus and incorporating functions. The RN -et 'to, with' along with an ergative prefix marking its head (e.g. w-et 'to/with me') may be incorporated into the verb structure between the absolutive marker and the verb stem. TVs are always intransitivized with -n in this construction. In addition, there is a clefted conjunct argument occurring outside the verb structure.
Craig notes that informants consistently translate the head of the incorporated RN as the subject in Spanish, not the clefted conjunct, in sentences such as these. This is the only case that I know of in Mayan languages where incorporating constructions of this sort occur. They deserve further study.

Accusative structures occur in Jac in what are called 'tenseless' embedded clauses. Tenseless embedded clauses are those which do not have either of the two 'tenses': 'incompletive' marked by ch- and a number of allomorphs, and 'completive' marked by x- ~ xx- (and other allomorphs). Tenseless clauses include, among others, those with the subordinating aspect particles: lañan continuative, kat post-sequentive, and lañwi presequentive. In tenseless embedded clauses, IVs require the ergative prefixes cross-referencing their Ss. TVs in tenseless embedded clauses become 'pseudo-intransitives' marked with the focus antipassive suffix -n, and with intransitive mode and/or phrase final suffixes. However, these pseudo-intransitives require ergative prefixes cross-referencing their As, and absolutive prefixes cross-referencing their Ps. Thus, in tenseless clauses TVs behave somewhat like TVs in that they take ergative prefixes, and TVs behave somewhat like IVs in that they require IV mode and phrase final suffixes. This is similar to the situation in Chj in the progressive aspect. Cf.
(115) lan an ha-way-i 'you are sleeping'
    prog A2-sleep-M

(116) x-Ø-ul naj kat Ø-y-il-n-i naj ix
    T-B3-arrive he seq B3-A3-see-Ap-M he she
    'he came and then he saw her'

(117) x-Ø-y-il naj him ha-maq'-n-i
    T-B3-A3-see he B1 A2-hit-Ap-M
    'he saw you hit me'

Greater Kanjobalan - A Note on Motozintlec (= Mocho)

There are no available sources on Mot but Larsen and Norman (1979) cite examples from Kaufman's field notes showing that Mot exhibits split-ergativity according to semantic features of NPs. Specifically, the ergative prefixes are used to cross-reference 1st and 2nd person As of TVs and Ss of IVs. On the other hand, in the 3rd person, As of TVs are cross-referenced with the ergative prefixes, and the Ss of IVs in the 3rd person are cross-referenced with the absolutive Ø marker. Ps of TVs are always cross-referenced with the absolutive markers. Thus, in the 1st and 2nd persons Mot exhibits marked nominative (= extended ergative) structures; while in the 3rd person Mot exhibits ergative marking. Compare the examples in (118) and (119).

(118) ii-muq-u-Ø 'I burned it'
    Al-burn-M-B3

(119) ii-maaq-i 'I went up'
    Al-go up-M
(119) x-muq-u-∅
    A3-burn-M-B3
    'he burned it'

maaq-i-∅
    'he went up'
go up-M-B3

Mot is apparently the only Mayan language which exhibits split-ergativity governed by the semantic features of NFs.

Tzeltalan


Tzeltalan - Tzotzil

In Tzt the absolutive markers cross-referencing Ps of TVs can be both prefixes and suffixes with differing forms in each position (cf. Table 6). In the active indicative, TVs are unmarked. IVs are also unmarked in the indicative. Subjects of IVs are prefixes in the nonperfective indicative, but suffixes in the perfective, and Ss of stative predicates are also suffixes. Examples of TVs in the active voice are given below.

(120) ch-a-mil-un
    T-A2-kill-B1
    ch-i-s-maj
    T-B1-A3-hit
    'you kill me'
    'he hits me'
i-ʃ-k-ak' 'I gave it'
T-B3-Al-give
ch-i-s-kolta-un 'he helps me'
T-B1-A3-help-B1
(121) i-ʃ-s-maj Maryan li nloʔ e
T-B3-A3-hit Mariano the Chamulan dem
'the Chamulan hit Mariano'

Tzt has a simple passive in -at (~ -ot in San Bartolo). In some
dialects (Zinacantán only ?), there is another passive in -e, which appar-
tently is used only on RTVs. There is a present passive participle in -v1
and a past passive participle in -b'11. In passive sentences, As may
optionally be expressed with the RN -uʔun 'by'.

(122) n-i-maj-at 'I was hit'
T-B1-hit-Ps
ch-imaj-e (y-uʔun leʔe) 'I am hit (by that one)'
T-B1-hit-Ps (A3-by that)
pak'-a1-ʃ 'it is stuck on'
stick on-Prp-B3
pak'-b'11-ʃ 'it has been stuck on'
stick on-Pp-B3
(123) i-ʃ-maj-e y-uʔun u10ʔ li Maryan e
T-B3-hit-Ps A3-by Chamulan the Mariano dem
'Mariano was hit by the Chamulan'

Tzt has an absolutive antipassive in -wan which can be used on
virtually any TV.

(124) n-i-mil-wan 'I killed'
T-B1-kill-Ap

61
There are two other absolutive antipassives which are not productive: -(o)maj (e.g. tzuk'-omaj 'wash' IV < tzuk' 'wash' TV), and -vwaj (e.g. nau-waj 'sew' IV < nau 'sew' TV).

There is a focus antipassive in -on which is used only with 3rd persons, and is rare. Since it is used only with 3rd persons it is impossible to say whether the verb agrees with the A or P. The A is clefted to the front of the sentence.

(125) ulo7 i-Ø-maj-on li Maryan

Chamulan T-B3-hit-Ap the Mariano

'the Chamulan was the one who hit Mariano'

Tzt has a referential and instrumental voice in -b'e. In the referential voice, dative, benefactive, and ablative or malefactive NPs, and possessors of Ps, are obligatorily promoted from an oblique case to the absolutive. The P is not demoted to an oblique case, but is no longer cross-referenced on the verb with an absolutive marker. However, if the P is plural, it is marked on the verb with the plural suffix -ik.

(126) dative: ch-a-k-al-b'-ot 'I'll tell it to you'

T-B2-Al-tell-R-B2

ch-au-ak'-b'e-Ø 'you'll give it to him'

T-A2-give-R-B3

benefactive: ch-a-j-mi1-b'e-ik 'I'll kill them for you'

T-B2-Al-kill-R-plr

malefactive: ch-i-s-po1-b'-un 'she'll snatch it from me'

T-B1-A3-snatch-R-B1

62
possessor
of P: n-a-tik-b'-un j-wakax
T-A2-enclose-R-B1 Al-cow
'you enclosed my cattle'
1a-∅-j-mil-b'e
T-B3-Al-kill-R
'I killed it of/for him'

Note that the referential voice may be used in conjunction with the
the passive voice.

(127) ch-i-7ak'-b'-at jun tzeb'
T-B1-give-R-Ps a girl
'I am being given a girl'

In the instrumental voice, an instrument NP may optionally be promo-
ted out of an oblique case marked with the preposition ta 'by, with, at, to'. However, it is not clear whether the instrument or the P is cross-
referenced on the verb in the instrumental voice, because the only avail-
able examples are in the 3rd person where the absolutive marker is null.

Compare the two examples in (128); in the first one, the instrument has
not been promoted; in the second, it has.

(128) 1a ∅-s-mil ta machita
T-B3-A3-kill with machete
'she killed him with a machete'
ta-∅-s-paj-b'e akuxa ti ka7e
T-B3-A3-prick-I needle the horse
'he pricked the horse with a needle'

It is noteworthy, that the instrumental voice in Tzt does not require
clefting of the instrument as it does, for example, in Quichean.
Tzeltalan - Tzeltal

Voice categories in Tzl are similar to those of Tzt. Examples of TVs in the active voice in Tzl are provided in (129). Note that the absolutive markers are always suffixed in Tzl.

(129)  
la j-pas-Ø  
'Ve did/made it'

  T A1-do-B3

ya j-mil-at  
'I kill you'

ya s-maj-Ø  
'she hits him'

  T A3-hit-B3

In (130) a number of sentential examples of TVs (from Tenejapa Tzl) in the active voice are provided. These sentences illustrate how word order differs depending on the status of the A and P on the animacy hierarchy: if A equals P the order is V A P; if A is higher than P the order is V P A.

(130)  
la s-mil-Ø  Jpetul te Jwan  
A=P

  T A3-kill-B3 Peter the John

'Peter killed John'

la s-t'om-Ø  ta ti7el tz'17 te baka  
A=P

  T A3-bite-B3 with teeth dog the cow

'the dog bit the cow (with teeth)'

la s-mil-Ø  baka te Jpetul e  
A>P

  T A3-kill-B3 cow the Peter dem

'Peter killed the cow'

ya s-maj-Ø  y-inam ta Jpetule  
A>P

  T A3-hit-B3 A3-wife agt Peter

'Peter hits his wife'
Tzl has a passive in -ot (e.g. maj-ot 'he was hit' with null tense and person marking). Kaufman (1963) states that As may be expressed in passive sentences with either the RN -u?un 'by' or the preposition to 'by, with, to, at'; or they may be unmarked. Kaufman gave no examples; however, Smith has presented a number of examples of passive sentences with unmarked As, some of which are given here in (131). Note that in passive sentences, like with active sentences, word order differs depending on the status of the A and P. But the order in passive sentences is the converse of that in active sentences. In passive sentences, if the P is equal to the A, then the order is V P A; if the P is higher than the A, then the order is V A P.

(131)  ḧ-mil-ot-Ø  Jpetul te Jwan  A=P
       T-kill-Ps-B3 Peter the John
       'Peter was killed by John'
       ḧ-mil-ot=Ø  baka te Jpetul  P>A
       T-kill-Ps-B3 cow the Peter
       'Peter was killed by the cow'
       ḧ-maj-ot-Ø  y-inam te Jpetul  P>A
       T-hit-Ps-B3 A3-wife the Peter
       'Peter was hit by his wife'

There is a past passive participle in -b'il (e.g. pas-b'il-Ø 'it has been done').

Tzl has a fully productive absolutive antipassive in -awan (e.g. mil-awan 'kill' IV < mil TV), and there are two other fairly productive ones: -omaj (e.g. tz'is-omaj 'sew' IV < tz'is TV), and -(V)wej (e.g. k'aj-wej 'harvest IV < k'aj TV). No evidence of incorporating or focus antipassives have been found in Tzl.
Like Tzt, Tzl has a referential voice in -b'a, which obligatorily promotes dative and benefactive NPs, and possessors of Ps, to the absolutive. No examples of instrumental promotion have been found.

(132) la y-ak'-b'e-∅  'he gave it to him'
   T A3-give-R-B3

ya s-pas-b'-on  'he does it for me'
   T A3-do-R-B1

la s-mil-b'e-∅  s-tz'i7  Jwan te  Jpetul e
   T A3-kill-R-B3  A3-dog  John the Peter dem

'Peter killed John's dog'

The referential and passive voices may be used together.

(133) ∅-ak'-b'o-t-∅  jun  Jwan te  Xmal e
   T-give-R-Ps-B3  a  John the Maria dem

'Maria was given paper by John'

No split-ergative constructions have been reported for either Tzt or Tzl.

Cholan


Ch1: Aulie and Aulie (1978), Warkentin and Whittaker (1970),
     Whittaker and Warkentin (1965).

Cholan - Chorti

In Chr, TVs are not inflected for tense or aspect; these categories are distinguished in context and by adverbials. RTVs end in a consonant and in the indicative they take the mode suffix -i (~ -e).DTV end in a vowel and have a null indicative marker. Absolutive markers cross-
referencing Ps are always suffixed. In the indicative, most IVs require a mode suffix (e.g. \(-\text{vy}\), or \(-\text{a} \sim -\text{i} \sim -\text{o} \sim \emptyset\)). IVs distinguish incomple- 
tive and completive aspects by the way person markers are attached. In 
the incompletive, a distinct set of person markers (neither absolutive 
or ergative) is prefixed (e.g. \(\text{in-kan-o}'I \text{ learn}'\)). In the completive, 
absolutive person markers are suffixed (e.g. \(\text{kan-v-e7n}'I \text{ learned}'\)). 
Note that often there are phonological modifications when affixes are 
attached to stems and when they are juxtaposed. Examples of TVs in the 
active indicative are provided in (134).

\[
\begin{align*}
(134) & \quad \text{in-xur-i-\(\emptyset\)} & \quad 'I \text{ cut it}' \\
& \quad \text{Al-cut-M-B3} & \\
& \quad \text{u-xur-i-\(\emptyset\)} & \quad 'he cut(s) it' \\
& \quad \text{A3-cut-M-B3} & \\
& \quad \text{u-chamse-\(\emptyset\)} & \quad 'he kills/killed it' \\
& \quad \text{A3-kill-B3} & \\
& \quad \text{inw-ajk'u-\(\emptyset\)} & \quad 'I \text{ hit it}' \\
& \quad \text{Al-hit-B3} & \\
& \quad \text{inw-ajk'-e7t} & \quad 'I \text{ hit you}' \\
& \quad \text{Al-hit-B2} & \\
\end{align*}
\]

The simple passive on RTVs is marked with an infix \(-\text{j}\), and on DTVs 
with \(-\text{n}\). Only 3rd person As may be expressed in the simple passive. 
There is another passive in \(-\text{te}\), but it is not clear how it is distin-
guished from the simple passive. All TVs have a past passive participle in \(-\text{b'ir}\). An A in any person may be expressed with the participial 
passives.

\[
\begin{align*}
(135) & \quad \text{in-xujr-a} & \quad 'I \text{ am cut}' \\
& \quad \text{B1-cut(Ps)-M} & \\
\end{align*}
\]
xujr-e7n u-men 'I was cut by him'
cut(Ps)-M-B1 A3-by
chamso-n-a-∅ 'it was killed'
kill-Ps-M-B3
ajk'u-n-a-∅ 'it was hit'
hit-Ps-M-B3
a-man-tz'-a 'it is bought'
B3-buy-Ps-M
xur-b'ir-∅ a-men 'it has been cut by you'
cut-Pp-B3
ajk'u-b'ir-∅ 'it is/has been hit'
hit-Pp-B3

There are three suffixes which apparently derive medio-passives: -k', -p, and -t. It is not certain how productive these suffixes are. The -t suffix is probably cognate with the -ot passive in Tzeltalan.

(136) a-b'ut'-k'-a 'it fills'
B3-fill-Mp-M
cf. u-b'ut'-i 'he fills it'

a-muk-p-a 'it gets covered'
B3-cover-Mp-M
cf. u-muk-i 'he covers it'

a-tz'ak-t-a 'it accumulates'
B3-add to-Mp-M
cf. u-tz'ak-i 'he adds to it'

Absolute antipassives are formed with -on on RTVs and -(w)an on DTVs.

(137) a-xur-on 'he cuts'
B3-cut-Ap

68
a-chams-an
'B3-kill-Ap
RTVs have another absolutive antipassive in -m, which emphasizes the fact that there is a P involved in the activity, even though it is not specified.

(138) a-xur-m-a
'B3-cut-Ap-M
There do not appear to be focus or incorporating antipassives in Chr.

Chr does not have an overtly marked referential voice (as, for example, in Tzeltalan and Ch1), but dative NPs are usually advanced to the absolutive. Verbs with advanced datives may be in the passive as well. In (139) dative advancement has not occurred; the dative NP is marked by the preposition ta 'to' and the RN -b'a. In (140) dative advancement and passive occur together.

(139) e Jwan uy-sb'ta'f in-te7 karta ta ni-b'a
the John A3-send-B3 one letter to Al-dative
'John sent a letter to me'

(140) uy-ajk'u-7on e chicha
A3-give-B1p the rum
'he gave me the rum'

(141) ajk'u-n-o7on e chicha
give-Ps-B1p the rum
'we were given the rum'

In the incompletive, Chr displays certain accusative-like features.
The 1st person prefixes used to cross-reference Ss of IVs are identical with the ergative prefixes (cf. Tables 6, 7, and 8). The 2nd person singular IV prefix may be an extension from the 2nd person plural erga-
tive. The 3rd person singular is not null as it is in other Mayan languages, although it is not identical with the 3rd person ergative either. It seems that there is a tendency towards accusative marking in the incompleteive, especially with 1st person. 30

Cholan - Chol

In Chl, there is an important distinction between the completeive aspect (marked with tza7 and allomorphs) and the incompleteive aspects (e.g. mi present, woli progressive, and mu7 - mux incompleteive). In the completeive indicative, RTVs have the mode suffix _V1_ (γ) which is a reduplication of the root vowel; and DTVs have null marking. In the incompleteive aspects, RTVs have null marking, and DTVs take _-n_. IVs in the completeive (indicative) take the suffix _-i(γ) (~ _-e(γ)_), and their Ss are cross-referenced with the absolutive suffixes (e.g. tza7 ach-iy-on 'I entered'). In the incompleteive, IVs take the suffix _-el_, and their Ss are cross-referenced with the ergative prefixes (e.g. mi k-och-el 'I enter'). Thus, Chl displays split-ergativity since it has ergative structures in the completeive, and accusative structures in the incompleteive. The accusative structures in the incompleteive are examples of extended ergative or marked nominative structures. Examples of TVs in the active indicative are presented below. 24

(142)  

<table>
<thead>
<tr>
<th>mi-j-k'el-Ø</th>
<th>'I see it'</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-Al-see-B3</td>
<td></td>
</tr>
<tr>
<td>mi-j-k'el-et</td>
<td>'I see you'</td>
</tr>
<tr>
<td>T-Al-see-B2</td>
<td></td>
</tr>
<tr>
<td>tza-j-k'el-e-Ø</td>
<td>'I saw it'</td>
</tr>
<tr>
<td>T-Al-see-M-B3</td>
<td></td>
</tr>
</tbody>
</table>
tza-j-k'el-ey-et
T-Al-see-M-B2
'I saw you'

mi-k-taj-∅
T-Al-meet-B3
'I meet him'

tza-k-taj-a-∅
T-Al-meet-M-B3
'I met him'

mi-7-taj-on
T-A3-meet-B1
'he meets me'

Tz-i7-taj-ay-on
T-A3-meet-M-B1
'he met me'

mi-k-cha7le-n-∅
T-Al-do-M-B3
'I do it'

tza-k-cha7le-∅
T-Al-do-B3
'I did it'

Passives of RTVs are formed with the infix -j-, unless the root ends in s, x, or j. If the root ends in one of these consonants, then the passive is formed with -t in the incompletive and -l in the completive.

Passives of DTVs are formed with -nt.

(143)  mi-j-k'ejl-cl
T-Al-see(Ps)-M
'I am seen'

tza7 k'ejl-iy-on
T see(Ps)-M-B1
'I was seen'

mi-k-taj-t-āl
T-Al-meet-Ps-M
'I am (being) met'

tza7 taj-l-ey-on
T meet-Ps-M-B1
'I was met'

71
woli-7-cha7le-nt-el 'it is being done'
T A3-do-Ps-M
tza7 cha7le-nt-i-∅ 'it was done'
T do-Ps-M-B3

Agents may be expressed with the preposition ti 'by, with, to' or cha7an 'by, because of'.

(144) tza7 päs-b'e-nt-i-∅ jun Jwan ti Alponso
T show-R-Ps-M-B3 paper John by Alphonse

'John was shown paper by Alphonse'

Past passive participles are formed with -b'il (e.g. k'el-b'il-on 'I am/have been seen', kolta-b'il-∅ 'he is/has been saved').

There are no antipassive constructions in Chl. Aside from Hua, Chl is the only Mayan language which does not have an antipassive voice of any kind.

Chl has a referential voice in -b'e in which dative, benefactive, and malefactive NPs, and possessors of Ps, are advanced to absolutive. Ps are removed from the absolutive but are not demoted to an oblique case.

(145) mi-j-k'ajti-b'e-n-et 'I am asking you it/it of you'
T-Al-ask-R-M-B2
woli-j-k'el-b'e-n-∅ 'I am watching for him'
T-Al-see-R-M-B3
tz-i7-taj-b'e(y)-on 'he met him for me'
T-Al-meet-R-B3
m-i7-b'ejlū-b'e-n-∅ i-wa7el 'she drags his food'
T-A3-drag-R-M-B3 A3-food

As the example in (144) illustrates, the referential voice may be used in conjunction with the passive.
At the beginning of this discussion of Chl, it was noted that in the incompleteive Chl has accusative marking since the ergative prefixes are used to cross-reference the Ss of IVs as well as the As of TVs. It is noteworthy that the mode suffix on IVs in the incompleteive is -Vl. This suffix is used throughout Mayan languages as a nominalizing suffix forming verbal nouns as well as deriving nouns in other types of constructions (e.g. Chl yäx Adj. 'clear of water', yäx-el n. 'greeness'; Tzu rax 'green', rax-aal 'greenness'). It seems likely, then, that IVs in the incompleteive aspects are (or at least originally were) nominalizations requiring the possessive ergative prefixes to cross-reference their Ss. The incompleteive aspect particles may be (have been) higher predicates subordinating verbs following them.

Yucatecan


Yuc is the only language discussed from this group of Mayan languages.

In Yuc there is a major distinction between RTVs and DTVs. Active IVs are also distinguished from inactive IVs. Active IVs (AIVs) are those which have a semantic A as subject, and inactive IVs (IIVs) are those which have a semantic P as subject. The distinction between AIVs and IIVs results in a further sub-classification of DTVs. Inflectionally, TVs derived from AIVs are distinguished from TVs derived from IIVs, under certain conditions. In addition, there is an important distinction between verbs in the (indicative) incompleteive aspect, and verbs in all other aspects and/or modes. In the incompleteive, verbs display accusa-
tive marking, since both As of TVs and Ss of IVs are cross-referenced with ergative prefixes, and Ps of TVs are cross-referenced with absolutive suffixes. In the completive aspect, and subjunctive and perfective modes, verbs have ergative marking since As of TVs are cross-referenced with ergative prefixes, and Ss of IVs and Ps of TVs are cross-referenced with absolutive suffixes. This kind of split-ergativity is an example of extended ergative or marked nominative morphological marking.

In the incompative, TVs have the mode suffix -\text{ik}, and IVs have -\text{V}_l, while AIVs are unmarked. In the completive, TVs take the mode suffix -\text{aj}, and IVs have the phrase final suffix -\text{i}. Examples of TVs in the active indicative are given in (146).

\begin{verbatim}
(146) RTV  k-in-tz'\text{on-ik-}\emptyset  'I shoot it'
        t-in-tz'\text{on-aj-}\emptyset  'I shot it'
        T-Al-shoot-M-B3
        k-in-jok'-\text{ik-}\emptyset  'I break it'
        t-in-jok'-\text{aj-}\emptyset  'I broke it'
        T-Al-break-M-B3
        DTV  k-u-m\text{list-ik-}\emptyset  'he sweeps it'
        < AIV t-u-m\text{list-aj-}\emptyset  'he swept it'
            T-A3-sweep-M-B3
        DTV  k-in-k\text{\ings-ik-ech}  'I kill you'
        < IIV t-in-k\text{\ings-aj-ech}  'I killed you'
            T-Al-kill-M-B2
\end{verbatim}

The passive voice on RTVs is marked with the infix $\text{\=2}\text{\=7}_1$ - which causes the preceding vowel to have high tone. On DTVs the passive is marked with the suffix $-\text{\=7}a$ in the incompletive and $-\text{\=7}a$ plus $-\text{b}'$ ($-\text{\=7}ab'$) in the completive. Agents may be expressed with the preposition $c(i)$ plus the RN
u-mêen plus an independent pronoun.

(147) k-in-tz'ôlon-ol  'I am (being) shot'
T-Al-shoot(Ps)-M

tz'ôlon-en  'I was shot'
shoot(Ps)-Bl

k-u-jèek'-el  t-u-mêen tèen 'it is broken by me'
T-A3-break(Ps)-M by I
jèek'-Ø  t-u-mêen tèen 'it was broken by me'
break(Ps)-B3 by I

k-u-mîst-â7a-l  'it is (being) swept'
T-A3-sweep-Ps-M

mîst-â7ab'-i-Ø  'it was swept'
sweep-Ps-M-B3

k-in-kîins-â7a-l  t-u-mêen leti7
T-Al-kill-Ps-M by he
'I am (being) killed by him'

kîins-â7ab'-en  t-u-mêen leti7

kill-Ps-Bl by he
'I was killed by him'

Some RTVs have a medio-passive voice which is marked by vowel length and high tone. Here the P is subject but no A is implied.

(148) k-u-jèek'-el  'it breaks'
T-A3-break(Mp)-M

jèek'-i-Ø  'it broke'
break(Mp)-M-B3

Both RTVs and DTVs have 'passive infinitives' in -b'il (e.g. kon-b’il 'to be sold' and kîins-b’il 'to be killed'). These forms in -b’il
are cognate with past passive participles in many other Mayan languages (cf. Table 9). It is not clear whether the forms in -b'íl in Yuc can serve as past passive participles as well as infinitives.

Bickerton states that Yuc also has 'agentless' passives in -p and -k', but no examples are given. These suffixes seem to be cognate with the corresponding medio-passive suffixes in Chr.

The absolutive antipassive of RTVs is formed by lengthening the root vowel, along with low tone, in the incomplete. In the complete, the suffix -naj is added to the stem with a lengthened vowel and low tone. The absolutive antipassive of DTVs from IIVs is formed with the suffix -aj in the incomplete, and -aj plus -naj in the complete. DTVs from AIVs do not have an absolutive antipassive per se, since the active intransitive verb itself fills this semantic slot. However, in the complete AIVs require the -naj absolutive suffix anyway. Let's look at the examples in (149).

(149) RTV  k-in-tz'óon
         T-Al-shoot(Ap)
         tz'óon-naj-en
         shoot(Ap)-Ap-Bl
         k-in-jèek'
         T-Al-break(Ap)
         jèek'-naj-en
         break(Ap)-Ap-Bl

DVT
< IIV  k-in-kfins-aj
         T-Al-kill-Ap

'I shoot'
'I shot'
'I break (something)'
'I broke (something)'
'I kill'
kíns-aj-naj-en 'I killed'
kill-Ap-Ap-B1
<k-AIV k-u-mís 'he sweeps'
T-A3-sweep
mís-naj-i-∅ 'he swept'
sweep-Ap-M-B3

There is an incorporating antipassive in Yuc, in which indefinite generic patient NPs may be incorporated within the verb, at least within RTVs (forming IVs). No examples of an incorporating antipassive have been found with DTVs. In the incompleteive the incorporated P occurs immediately following the verb root with no further marking. In the completive, the incorporated P occurs immediately following the verb root, and then is followed by the antipassive suffix -naj. Cf.

(150) k-in-p'o7-nòok 'I wash clothes'
T-Al-wash-clothes

cf. active: k-in-p'o7-ik nòok 'I wash the clothes'
p'o7-nòok-naj-en 'I washed clothes'
wash-clothes-Ap-B1

cf. active: t-in-p'o7-aj nòok 'I washed the clothes'

Patient incorporation in Yuc is complete. It is the only instance in a Mayan language (that I know of), where the incorporated generic P is inserted between the verb stem and following suffixes.

There is a focus antipassive used in Yuc when As are questioned and clefted. In this construction, the agent NP occurs at the beginning of the sentence but is not cross-referenced on the verb. In fact, the verb apparently has no tense, aspect or mode marking, with the exception of the phrase final suffix -q. The P is cross-referenced on the verb with
an absolutive suffix. Compare the examples in (151).

(151)  
mas puch-en?  
who hit-B1  
't who hit me?'  

òen puch-ech  
I hit-B2  
'I was the one who hit you'

max puch-e-∅?  
who hit-M-B3  
'who hit him?'

tèen puch-e-∅  
I hit-M-B3  
'I was the one who hit him'

max miist-e-∅? naj  
who sweep-M-B3 house  
'who swept the house?'

tèen miist-e-∅ naj  
I sweep-M-B3 house  
'I was the one who swept the house'

max il-ech?  
who see-B2  
'who saw you?'

cf. active: max t-aw-il-aj-∅?  
'who did you see?'

These forms appear to be intransitive and in the completive aspect, since there is null mode marking and only Ps are cross-referenced. However, the phrase final suffix -∅ is usually used on TVs when no other mode suffix occurs (e.g. in the subjunctive), and the phrase final suffix -∅ is normally used on IVs in the completive. In other words, forms in the focus antipassive display features of both transitive and intransitive verbs. The focus antipassive needs further study in Yuc.

The historical origins of the passive and medio-passive markers in Yuc are noteworthy. High tone and vowel length marking the medio-passive on RTVs is cognate with the infix -∅ occurring as a passive or medio-passive marker in a number of other Mayan languages (cf. Table 9). Yuc
syllables of the form CVVC are reflexes of PM *CVhC or *CVjC (cf. Campbell 1977, Fox 1978, Kauman 1966-69, and McClaran 1977). Evidence from Quichean shows that the PM infix was originally *-h- not *-j-. There is evidence that the infixed ʃV₁妻 marking the passive on RTVs in Yuc was originally *ab'. In Classical Yuc the passive on both RTVs and DTVs was -ab' (cf. McQuown 1967); and in Modern Yuc the passive is marked with -V₁b' on RTVs which end in ʃ (e.g. tz’a7 'give' > tz’a7-ab' 'be given', chi7 'bite' > chi7-ib' 'be bitten'). Also, in contemporary Mopán, closely related to Yuc, the passive on RTVs is -V₁b' and on DTVs it is -ab' (Will Norman, p.c.). The suffix -a7a marking the passive on DTVs in Modern Yuc reflects the earlier -ab'; the *b' has become ʃ, and an echo vowel has been added (ʃ ~ b' is a common alternation in many Mayan languages). The form of the passive in the completeive of DTVs (i.e. -a7ab') has probably been created analogically on the basis of -a7a-1 incompleteive passive, and -a7a-k subjunctive passive, originally from *ab'-al and *ab'-ak, respectively. Thus, it seems likely that RTVs originally had passives of the form *CVC-ab' which became *CVC-V₁ seating. This suffix became an infix resulting in passives of the form *CVVC. Syllables of this form regularly become CVVC in Yuc with the addition of an echo vowel.

As was discussed at the beginning of this section on Yuc, the ergative prefixes are used to cross-reference Ss of IVs in the incompleteive aspect, resulting in accusative marking. It is noteworthy that the ergative prefixes cooccur with the (IV mode) suffix -V₁ in the incompleteive. It was pointed out in the previous section on Ch'ol, that -V₂ is a pan-Mayan nominalizing suffix. Therefore, accusative marking in Yuc probably can be attributed to the use of possessive ergative prefixes on verbal
nouns in the incompleteive aspect. These verbal nouns were (are?) probably subordinate to adverbial particles used in the incompleteive, some of which have prefixal forms (e.g. k- < ki present, which is kaj in Lac; t- < tan durative; tz'- < tz'o7ok terminative).

Huastec


There are no good grammatical descriptions of Hua, so what is said here is rather brief and tentative. TVs always end in a thematic suffix (Tm), which is usually of the form -Vy, -Vw, or -v(7). In the incompleteive this suffix is then followed by -al; in the completeive no suffix follows. The ergative and absolutive person markers are proclitics. When the marker of the P of a TV is nonnull (i.e. when it is not 3rd person), then collapsed pronominals are used (e.g. tin A2/A3 -> B1, ti A3 -> B2, tu A2/A3 -> Blp or A1 -> B2). Some examples of TVs in the active voice are presented in (152).

(152)  
6-u  nuj-uw-al  'I sell it'
B3-A1 sell-Tm-M
6-in  al-iy-al  'he looks for it'
B3-A3 search-Tm-M
6-in  kwee7-ey-al  'he robs it'
B3-A3 rob-Tm-M
6-u  t'aj-a-al  'I do it'
B3-A1 do-Tm-M
6-u  t'aj-a7  'I did it'
B3-A1 do-Tm

80
tin kw'i-y-a-al 'he scolds me'
A3->B1 scold-Tm-M

A3->B1 scold-Tm

tu kwath-a-al 'I hit you' present
A1->B2 hit-Tm-M

tin kwath-a7 'he hit me' past
A3->B1 hit-Tm

Like Tz'il, Hua has differing word order depending on the status of the A and P on the animacy hierarchy. If the A equals the P then the order is V A P; if the A is higher than the P then the order is V P A.

(153)  Ø-in kwath-a7 an uxum an inik A = P
B3-A3 hit-Tm the woman the man
'the woman hit the man'

Ø-in k'at'-uw an pik'o7 an olom A = P
B3-A3 bite-Tm the dog the pig
'the dog bit the pig'

Ø-in k'oaj-ow oox i ajan an inik A > P
B3-A3 pick-Tm three corn the man
'the man picked three ears of corn'

Hua has a passive formed in -aab'; and -b'il plus -aab' form past passive participles.

(154)  Ø juj-uw-aab' 'it is (being) sold'
B3 sell-Tm-Ps

in al-iy-aab' 'I am (being) looked for'
B1 search-Tm-Ps
t'aj-b'il-aab'  'made-done'
do-Pp-Ps
kwe7-b'il-aab'  'robbed'
rob-Pp-Ps

There is another passive suffix -n, which apparently forms both passives and medio-passives.

(155) in kw'i-iy-n-al   'I am scolded'
B1 scold-Ps-M
ti t'aj-a-n   'te hiciste'
B2 do-Tm-Ps   (Sp. 'you do it to yourself/
                it was done to you')
Ø kWajl-a-n   'it fell'
B3 knock down-Tm-Mp
              < kWajl-a-al  'knock over/down' TV
Ø k'ib'-a-n   'it got lost'
B3 lose-Tm-Mp
              < k'ib'-a-al  'lose' TV

This passive or medio-passive suffix is probably cognate with the anti-
passive suffix -(V)n found in most other Mayan languages (cf. Table 10).
However, there does not appear to be an antipassive of any kind in Hua.

There is a referential voice formed in -ch(i), which has the reduplic-
cated plural form -chinch.

(156) Ø u nuj-ch-al   'I sell it to him'
B3 Al sell-R-M
Ø u nuj-chinch-al  'I sell it to them/
B3 Al sell-R-M    I sell them to him'
Ø in kwe7-ch-al  'he robs him of it'
B3 A3 rob-R-M

Ø in t'aj-ch-al  'he does it to him'
B3 A3 do-R-M

The referential voice may be used in conjunction with either of the two passives.

(157) in nuj-chinch-aab'  'I am (being) sold them'
B1 sell-R-Ps
in kwe7-chi-n-al  'I am (being) robbed of it'
B1 rob-R-Ps-M

Kaufman (p.c.) has noted that there is a suffix -na7 which derives instrumental verbs (e.g. t'aj-na7 'do with' < t'aj-a-al 'do', and koto-x-na7 'cut with' < kot oy-al 'cut'). It is not clear whether this suffix is productive or whether it is a voice changing suffix, as, for example, -b'e in some other Mayan languages (cf. Table 11).

Conclusion

In this section a number of generalizations about ergativity and voice are stated, especially with respect to Mayan languages. In addition, some speculations are made about the voice system of Proto-Mayan (PM). The speculations are meant to be working hypotheses pointing to areas of further research, and are not intended as actual reconstructions. And since the primary purpose of this paper is not historical, the argumentation justifying the hypotheses is neither rigorous nor comprehensive. The motivation for presenting the hypotheses is to stimulate further investigation on the part of Mayanists to provide evidence supporting or disproving them. It should be noted that what is said about PM primarily
pertains to the non-Huastec part of the family (i.e. to the stage of PM after Huastec split off), since Huastec is not only the most divergent Mayan language, it is also one of the least documented.

Ergative languages differ in one important respect from other types of languages (e.g. accusative languages): the agent of a transitive verb is a marked (ergative) category, and has a rather sacrosanct position. Therefore, it is often the case that the (ergatively marked) transitive agent is not accessible to certain kinds of grammatical processes like relativization, questioning, and clefting (and, in some non-Mayan languages, coordination). In order for a transitive agent to undergo processes such as these, it must be modified so that it does not formally look like a transitive agent; that is, it must not have ergative marking. The focus antipassive construction, therefore, plays an important role in many ergative languages. Its primary function is to make unequivocally explicit who the agent is without marking it (sacrosanctly) ergative. It is used in those grammatical processes which are not accessible to ergatively marked agents. It may also be used when there is a possibility of ambiguity as to which NP in a transitive sentence is the agent and which is the patient.

In ergative languages like Mayan languages, which do not have morphological case marking on agent and patient NPs, but rather only cross-reference these roles with affixes on the verb, the possibility of ambiguity can arise when there are two 3rd person NPs in the sentence, and when certain kinds of grammatical processes (e.g. topicalization) disrupt the normal word order. Thus, for example, if the normal order is V Λ P (or V P Λ, for that matter), when one of these NPs is fronted via topicalization the result is NP V NP, and therefore, it may not be clear
which NP, the agent or the patient, has been fronted. The focus antipassive may be used to disambiguate in this situation because it explicitly indicates that the NP immediately preceding the verb is the agent.

In all Mayan languages which have a focus antipassive the agent is always clefted to the left of the verb and the verb becomes a derived intransitive. However, generally speaking, there are three different ways in treating patients and verb agreement: (1) In Kek and Pocom, and sometimes in Tzu, the patient is demoted to an oblique case marked with a RN, and the verb agrees with the agent because it is the only argument left in direct relationship with the verb. Pcm optionally allows the patient to not be marked with an RN. In Tzu, the patient is demoted and the verb agrees with the agent when there are two non-3rd person arguments in the sentence, or, in order to override person hierarchy agreement (discussed next). (2) In Qui and Cak, and usually in Tzu, the verb agrees with whatever argument (agent or patient) is highest on the person hierarchy, and the patient is not demoted. Aquacatec also optionally allows the verb to agree with a non-3rd person agent if the patient is 3rd person. (3) In other Mayan languages (i.e. in Mamean, Kanjobalan, Tzeltalan, Yucatecan) with a focus antipassive, the verb agrees with the patient.

In those cases where the verb agrees with the patient, it seems likely that the process of focus antipassivization removes the agent from having a direct relationship with the transitive verb. The removal of the agent then leaves the patient as the only argument of the verb, so the verb agrees with it and also becomes intransitive. This situation is illustrated in (158).
In those cases where the verb agrees with the argument highest on the person hierarchy the situation is essentially like that in (158), if the patient is equal to, or higher than, the agent. But if the agent is higher on the person hierarchy than the patient, its higher status overrules verb agreement with the patient, and forces the verb to agree with the agent even though it has been removed from a direct relationship with the verb.

Since Kek and Pocom are the only languages in which the verb in the focus antipassive regularly agrees with the agent, and since in Quichean proper, and optionally in Aguacatec, verb agreement with the agent is restricted to special circumstances (i.e. those governed by the person hierarchy), it seems likely that verb agreement with the agent in the focus antipassive is an innovation, given that in all other Mayan languages the verb agrees with the patient.

The suffix -(V)n occurs as a marker of the focus antipassive in all Mayan languages, except Yuc (cf. Table 10). In Greater Quichean, another suffix, -ow (~ -o -w), marks the focus antipassive on RTVs only. The use of this suffix in Quichean in the focus antipassive is viewed here as an innovation. As a first approximation then, the following hypothesis is made with regard to the focus antipassive in PM.

Hypothesis 1: PM had a focus antipassive voice marked with the suffix *(V)n, in which the verb agreed with the patient, and the agent was clefted to the left of the verb and removed from having a direct relationship with it (as in (158)). The focus
antipassive was used when agents were questioned, relativized, and clefted.

In Chl, Tzeltalan, Greater Kanjobalan, Kek, and Pocom, there is a suffix -\(w\) (with various modifications) used to mark the absolutive antipassive (cf. Table 10). Other languages with an absolutive antipassive mark it with -\((V)n\), at least under certain conditions. In this paper, the view is that the use of -\((V)n\) as an absolutive marker has come about by extending its usage from the focus antipassive construction. And conversely, the use of the -\(ow\) suffix in Greater Quichean to mark the focus antipassive on RTVs has come about through extension from the absolutive passive. At present, I have no good explanation of precisely how this postulated flip-flop in functions of -\((V)n\) and -\(ow\) came about in Quichean.

**Hypothesis 2**: PM had an absolutive antipassive voice marked with *-\(w\), which allowed the omission of the patient in the expression of a transitive activity (as in (5)).

Several languages scattered throughout the Mayan family have an incorporating antipassive (cf. Table 10). In Greater Kanjobalan this voice is marked with -\(w\), and in Kek and Tzu it is marked with -\(o\) (\(<*-ow\) on RTVs.

**Hypothesis 3**: In PM the suffix *-\(w\) was also used in incorporating antipassive constructions in which a nonspecific generic patient was incorporated into the verb complex (as in (6)). The verb became a derived intransitive and agreed with the agent.

The incorporating antipassive is marked with -\((V)n\) in Mam and Yuc, and it is used on DTVs in Kek and Tzu. The use of -\((V)n\) in these languages to mark the incorporating antipassive is viewed here as having come about.
through extending its use from the focus antipassive.

Hua and Pch have passives in -\(\mathbf{V'b}'\), and Classical Yuc had one in -\(\mathbf{ab}'\) (cf. Table 9). Mam has one in -\(b'\mathbf{aj}'\), which historically is a compound formation from -\(b'\) plus -\(\mathbf{aj}'\). Kek has a passive in -\(e7\) which historically comes from *-\(eb'\). The suffix -\(e7\) (<*-eb'\>) is also found in Quichean marking intransitive verbs from positional roots, and in Milpas Altas Cak marking passives of RTVs. Adjetival passives in -\(b'\mathbf{il}\) are found in Hua, Yucatecan, Cholan, Tzeltalan, Kanjobalan, and Kek. Historically, this suffix is a compound formation built on -\(b'\) plus an adjetival/nominal suffix -\(\mathbf{il}\). Whenever there is a passive construction based on the suffix -\(b'\), unrestricted expression of agents in an oblique case is allowed.

Hypothesis 4: PM had a passive in *-\(\mathbf{V'b}'\), which allowed unrestricted expression of agents in an oblique case (as in (6)). PM also had a past passive participle in *-\(b'-\mathbf{il}\), which also allowed unrestricted expression of agents in an oblique case.

In Yuc, Tzl, Agt, and Tzu, there are medio-passives marked with an infixed -\(\mathbf{i}'\) or -\(\mathbf{V}'\). And in Cholan and Quichean there are passives in -\(\mathbf{i}'\) and -\(\mathbf{V}'\). Historically, these infixes can be shown to be a reflex of PM *-\(h'-\). In those languages where the infix is used as a passive marker, there are usually restrictions on the expression of agents (e.g. only 3rd person agents may be expressed in Chr and Qui), and in general, the infix passive seems to de-emphasize the role of the agent.

Hypothesis 5: PM had a medio-passive, or perhaps a passive de-emphasizing the role of the agent, marked with the infix *-\(h'-\) (cf. (4) and discussion of medio-passive).

Passives in -\(t\) are also fairly widespread in a number of Mayan languages. Cholan has one in -\(t\); Tzeltalan and Jac have one in -\(ot\) (- -\(at\));
and Quichean has one in -taj which can be further analyzed into -t plus -aj historically. Will Norman (p.c.) says that there is a cognate of this suffix in Hua which is used as an intransitive verb marker on positional roots.

Hypothesis 6: PM may have had a passive in *-t.

A recurrent element in Mayan languages in passive constructions, and in some cases in antipassive constructions, is the suffix -v_1 -v_n, from PM *-v_j. This suffix occurs in many languages as a generalized intransitive marker (and in a few languages as a transitive marker as well). It probably did not have a specific passive (or antipassive) function in PM, but has been restricted in use in a number of languages.

In Tzeltalan and Chol the suffix -b'e marks the referential voice, and in several Eastern Mayan languages, as well as in Tzt, -b'e functions as the instrumental voice marker. There is internal evidence in Quichean, which indicates that there was once a referential voice in these languages as well (e.g. Tzu tzijo-b'e-xik 'to talk to' < tzijo-xik 'to speak about').

Hypothesis 7: PM had a referential voice marked with *-b'e, which (obligatorily ?) promoted datives, benefactives, malefactive, and possessors of patients, to the absolutive. *-b'e also (optionally ?) promoted instruments to absolutive, especially when they were in focus. The patient removed from the absolutive was not put in an oblique case, but rather left en chômage without further marking.

Accusative morphological marking occurs in a number of Mayan languages (cf. Table 3). All accusative structures in Mayan are instances of extended ergativity: that is, in every case the ergative prefixes have been extended to IVs marking their subjects. Accusative structures in
Yuc, Chl, Agt, Pocom, and Kek, have clearly come about through the use of nominalizations in certain kinds of constructions, especially in embedded clauses which may include the incompleteive and progressive aspects. In all of these cases the possessive ergative prefixes occur on nominalizations of IVs cross-referencing their underlying subjects. The use of nominalizations in the incompleteive and progressive aspects in Kek and Pocom are clearly cases of (aspect marking) adverbial particles acting as higher predicates with embedded clauses containing the nominalizations. The use of nominalizations in the incompleteive aspects in Yuc and Chl may have come about in a similar way, but it isn't clear that this is the case synchronically.

Chj, Jac, Mam, and Ixl have accusative marking in several kinds of subordinate clauses (which do not seem to be embedded nominalizations). In all of these cases, it appears that the verbs of the subordinate clauses are not marked with the normal tense/aspect inflections. They are either tenseless/aspectless, or have subordinating adverbial particles which may denote time and aspect, among other things.

Mot appears to be the only Mayan language with accusative marking governed by the semantic features of NPs (i.e. non-3rd persons are marked accusatively and 3rd persons are marked ergatively).

Hypothesis 8: (A) PM was morphologically ergative in that agents of transitive verbs were cross-referenced on the verb with a set of ergative prefixes, and the subjects of intransitive verbs and the patients of transitive verbs were cross-referenced with another set of absolutive affixes. Possessors of NPs were also cross-referenced with ergative prefixes. (B) PM may have had marked accusative structures (= extended ergative structures) in certain
kinds of tenseless/aspectless subordinate clauses, especially with embedded nominalizations in which the underlying subjects of intransitive verbs were cross-referenced with ergative possessive prefixes.

An important question remains unanswered in the study of ergativity. In many ergative languages accusative structures are found in the incomplete/nonpast tenses and aspects. Why should there be such a tendency? Even if accusative marking can be attributed to nominalizations in many of these cases, why should there be more of a tendency towards nominalizations in these aspects and tenses than in the completive or past?

Footnotes

1 The family of Mayan languages contains about 30 languages which, with the one exception of Huastec, are all spoken in a more or less contiguous area stretching from the states of Chiapas, Tabasco, Yucatán, Campeche, and Quintana Roo, in southeastern Mexico, into Guatemala, Belize, and Honduras. Huastec is separated from the rest of the family and is spoken over 1000 miles to the north in the states of San Luis Potosí and Vera Cruz, Mexico. The separation of Huastec dates very early in the development of the Mayan family, and it is generally thought that the territory occupied today by Mayans has not changed a great deal for a long time, although there have been internal movements of particular language groups within the contiguous area (see Kaufman 1966-69, 1974).

The abbreviations used for grammatical terms in this article are as follows:

A agent; and ergative (Set A) person prefix (e.g. A1 = 1st person singular ergative, A2p = 2nd person plural ergative).
AIV active intransitive verb.

Ap antipassive suffix.

B absolutive (Set B) person affix (e.g. B1 = 1st person singular absolutive, B2p = 2nd person plural absolutive).

C consonant.

Cx consonant determined by the x-consonant of the preceding root or stem.

D directional particle.

DTV derived transitive verb.

I instrumental voice suffix.

IIIV inactive intransitive verb.

IV intransitive verb.

M mode suffix and/or phrase final suffix.

Mp medio-passive voice suffix.

P patient.

Pp past participle deriving suffix.

Pr preposition.

Ps passive voice suffix.

R referential voice suffix.

RN relational noun.

RTV root transitive verb; i.e. a nonderived transitive verb.

S subject.

T tense/aspect prefix or proclitic.

Tm thematic suffix.

TV transitive verb.

V vowel (lexically determined).

Vx vowel harmonizing with the x-vowel of the preceding root or stem.
The discussion of Huastec, however, is wanting, because, unfortunately, there is not sufficient literature on this important language to give a full account of voice.

Some of the discussion of transitivity here, parallels that found in Dayley (1978) and Dixon (1979), and has been influenced by earlier works, especially Fillmore (1968), Chafe (1970), and Smith-Stark (1976a).

See M. Silverstein's (1977) important article on ergativity and the agency hierarchy for a more detailed and rigorous discussion. Note that the diagram of the agency hierarchy given here omits many details. For example, plural forms are always higher on the hierarchy than singular forms.

See Dixon (1979) for a detailed discussion of 'Subject' in this sense.


Dixon (1979) makes the stronger claim that no language is 100 percent morphologically ergative, that all ergative languages have constructions displaying accusative morphology. This seems to be an overstatement, since, for example, there are a number of Mayan languages which do not appear to have any signs of accusative morphology (cf. Table 3).

The term pivot is defined and discussed in detail in Dixon (1979). The notion of pivot is a good one, although I wish there was a better term.

The overt nature of a voice change is meant to exclude such sentences as:
We ate.

John drinks.

The stick broke.

from being viewed as having undergone a voice change (e.g. antipassivization and medio-passivization). The normally transitive verbs 'eat', 'drink' and 'break', are viewed here as having lexical intransitive forms. This view contrasts, for example, with Heath (1976) and Postal (1976).

Cf. Friedn (1975) and Langacher and Munro (1975) for insightful discussions of nonderived passives similar to the one outlined here.

Antipassive voices most commonly occur in ergative languages but also occur in some accusative languages as well (e.g. Uto-Aztecan). The term antipassive was first coined by Silverstein ca. 1968 (cf. 1977). There are a number of different kinds of antipassives with differing functions (cf. Heath 1976). In this paper we will only be concerned with absolutive, incorporating, and focus or agent antipassives because they are the only ones occurring in Mayan languages (cf. Table 10). For other discussions of antipassives see Dayley (1978), Dixon (1972, 1979), Postal (1977), and Smith-Stark (1976b, 1978). The term absolutive (or verbo absoluto in Spanish) describing the absolutive antipassive in Mayan languages goes back at least to the colonial period (cf. Ximénez 1701-03).

The two different functions of voice changes, i.e. the omission of arguments and rearrangement of arguments, could be viewed as semantic and syntactic voice changes, respectively. Although, I prefer to view them both as ultimately semantically motivated.

For a relationship between focus and relativization see Schacter (1973).
This is a simplified generalization; the details of the focus antipassive differ in different Mayan languages, especially with respect to which argument, A or P, is referenced on the verb with an absolutive person marker. See the section on 'Voice and Ergativity in Selected Mayan Languages' for details in particular languages. I prefer the term 'focus antipassive' because it describes the function of this voice. Perhaps a more accurate, albeit cumbersome, name would be the agent focus antipassive.

Compare the uses or functions of the instrumental voice with those of the focus antipassive and agents. See Norman (1978) for a discussion of the instrumental voice in Mayan.

Abbreviations of the language from which examples come sometimes precede the examples in order to avoid confusion.

Data on the various languages discussed here varies a great deal in both quantity and quality, and so exemplification of voice and ergativity will differ accordingly in the presentation of each language. It would be nice if, for example, the same verb(s) could be used to exemplify each of the voices in a given language. However, in the literature it is seldom the case that examples of each voice are available for the same verb. I present examples of forms cited by authors of the primary sources, and use the same verb(s) when they are available. Usually examples are presented with dashes indicating morpheme boundaries and minimal glosses are given for each morpheme (cf. the list of abbreviations in footnote 1). A certain amount of morphological analysis is left up to the reader in order for the discussion not to become too redundant. Phonological rules are not discussed unless absolutely necessary, but the reader should be aware that there are rules lengthening, shortening, and deleting vowels.
under certain conditions (e.g. in stressed or unstressed syllables, because of certain affixes, etc.). Also many affixes are cited with one of the segments being V or C. Without a subscript number a V or C is lexically determined; with a subscript number a $V_x$ or $C_x$ is determined by the $x$-V or C of the root or stem. For example, $-V_1$ means the vowel of the suffix is lexically determined; $-V_11$ means the vowel of the suffix is the same as the first vowel of the preceding root or stem. Orthographies of all of the languages have been regularized to conform with Kaufman (1975).

Many of the cases of split-ergativity in Mayan can be traced historically (if not synchronically) to nominalizations like these where ergative possessive prefixes are used to mark underlying Ss of IVs (see the Conclusion). Kek seems to show an incipient stage of this development.

Data on Tzu is from the San Juan dialect; the Qui data is from the Nahualá dialect; and the Cak data is from Magdalena Milpas Altas and Comalapa. In general, dialectical differences within Mayan languages are not discussed unless they are crucially relevant to voice and ergativity.

There is actually another class of DTVs marked with the indicative suffix $-V_7$. These will not be discussed since they behave essentially like DTVs in $-V_1$ (see Dayley 1978 and 1982, for details).

The absolutive affixes in Table 6 given for Mam as well as for Agt are underlying or basic forms. These two languages have complex morphophonemics in the absolutives. See England (1975) and Larsen (1978) for detailed discussions. The data on Mam is from Northern Mam, either from Todos Santos (TS) or from San Ildefonso Ixtahuacán (SI). Northern Mam differs substantially from Western and Southern Mam, not discussed here.

Norman and Campbell (1978) mistakenly called passives in $-ot$, 'antipassives'. This error was probably due to the fact that Tzl allows
unmarked As in passive sentences, an abnormality in Mayan languages. This makes them look somewhat like focus antipassives which, in some languages, allow A and P to both be unmarked even though the verb is formally intransitive. That these are not focus antipassives is clear since the As are not clefted, and according to Kaufman the As can be overtly marked in an oblique case with ta or -u7un, like normal passives. Verbs in -ot are not absolutive antipassives either; this voice is marked by -awan (cf. (132)). Active and passive sentences are clearly distinguished in Tzl by the suffix -ot and by the contrasts in word order in the two different voices. It is noteworthy, however, that the preposition ta may be used to mark As in active sentences as well as passives (cf. the last example in (130) where ta is used to mark the A of an active sentence).

23 The 1st person singular is not particularly criterial, since in many Mayan languages the 1st person singular absolutive and ergative affixes are similar.

24 N.B. mi plus the 3rd person ergative prefix i- becomes mi7 and woli plus i- becomes woli7; mi and woli plus the 2nd person singular ergative prefix a- become ma7 and wola7, respectively. The 1st person ergative prefix k- becomes j- before k.

25 Throughout this paper the perfective, subjunctive, and imperative modes have been ignored since they are not really relevant to a discussion of voice and ergativity. It should be noted, however, that most Mayan languages have at least three of the following four modes: indicative (= nonperfective), perfective, subjunctive, and imperative. In Yuc the perfective is formed with the suffix -maj on TVs and with -a7an on the completive stem of IVs. The subjunctive is marked with the phrase final suffix -a on TVs, and with the suffix -V,k - -ak on the completive stem
of IVs.

The focus antipassive is not well documented in Yuc. The examples cited here are from Blair (1964:118) and Tozzer (1921:94). This construction needs further study to see, for example, if it is used in agent-relativization and other aspects and modes.

Pocon reflects PM *h as h and PM *j as j. The infixed -h- occurring in Pocon cognate with the (medio) passive in other languages, is used to mark intransitive verbs from positional roots (e.g. yohk-ik 'to lie down' < yok- 'lying down'). Also Qui reflects syllabic internal PM *-j- as -j- (e.g. najt 'far' < PM *najt), but reflects syllabic internal PM *-h- as -v- (e.g. b'aalam 'tiger' < *b'ahlam).

N.B. the -et passive suffix in Mam is not cognate with other passives in -t. The t in the Mam suffix reflects PM *r.

N.B. -vn (-vj) here from PM *-vn is distinguished from -vn of the (focus) antipassive which clearly comes from PM *-vn, since the *n is reflected as n in all languages.

Since this article was written there have been two others commenting on the section herein on Chorti. One is John Fought's (1982) criticisms of Dayley's treatment of Chorti verb morphology, and the second is Dayley's (1982) response to Fought's comments. For a detailed discussion of the verb morphology of Chorti, as well as other Cholon languages, the reader might also consult Kaufman and Norman's (1982) excellent article.
Mayan Language Family
(Kaufman 1974, 1975)

<table>
<thead>
<tr>
<th>I</th>
<th>Huastec complex</th>
<th>Chicomuceltec (Chi)</th>
<th>Huastec (Hua)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Yucatec (Maya) complex</td>
<td>Yucatec (Yuc)*</td>
<td>Itzá (Itz)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lacandón (Lac)*</td>
<td>Mopan (Mop)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>III</th>
<th>Western division</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A Greater Cholan branch</td>
</tr>
<tr>
<td></td>
<td>1 Cholan group</td>
</tr>
<tr>
<td></td>
<td>Chol (Ch1)*</td>
</tr>
<tr>
<td></td>
<td>Chortí (Chr)*</td>
</tr>
<tr>
<td></td>
<td>Chontal (Chn)</td>
</tr>
<tr>
<td></td>
<td>2 Tzeltalan group</td>
</tr>
<tr>
<td></td>
<td>Tzeltal (Tzl)*</td>
</tr>
<tr>
<td></td>
<td>Tzotzil (Tzt)*</td>
</tr>
<tr>
<td></td>
<td>B Greater Kanjobal branch</td>
</tr>
<tr>
<td></td>
<td>1 Chujoan group</td>
</tr>
<tr>
<td></td>
<td>Tojolabal (Toj)*</td>
</tr>
<tr>
<td></td>
<td>Chuj (Chj)*</td>
</tr>
<tr>
<td></td>
<td>2 Kanjobal group</td>
</tr>
<tr>
<td></td>
<td>Kanjobal (Kan)</td>
</tr>
<tr>
<td></td>
<td>Acatec (Act)</td>
</tr>
<tr>
<td></td>
<td>Jacalteco (Jac)*</td>
</tr>
<tr>
<td></td>
<td>b Motozintlec (Mot)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IV</th>
<th>Eastern division</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A Greater Mamian branch</td>
</tr>
<tr>
<td></td>
<td>1 Mamian group</td>
</tr>
<tr>
<td></td>
<td>Mam*</td>
</tr>
<tr>
<td></td>
<td>Teco (Tec)</td>
</tr>
<tr>
<td></td>
<td>2 Ixilan group</td>
</tr>
<tr>
<td></td>
<td>Aguacatec (Agt)*</td>
</tr>
<tr>
<td></td>
<td>Ixil (Ix1)*</td>
</tr>
<tr>
<td></td>
<td>B Greater Quichean branch</td>
</tr>
<tr>
<td></td>
<td>1 Uspantec (Usp)</td>
</tr>
<tr>
<td></td>
<td>Quiché (Qui)*</td>
</tr>
<tr>
<td></td>
<td>Achí (Ach)</td>
</tr>
<tr>
<td></td>
<td>Sacapultec (Sac)</td>
</tr>
<tr>
<td></td>
<td>Sipacapa (Sip)</td>
</tr>
<tr>
<td></td>
<td>Tzutujil (Tzu)*</td>
</tr>
<tr>
<td></td>
<td>Cakchiquel (Cak)*</td>
</tr>
<tr>
<td></td>
<td>3 Pocom complex</td>
</tr>
<tr>
<td></td>
<td>Poconam (Pcm)*</td>
</tr>
<tr>
<td></td>
<td>Pocomchí (Pch)*</td>
</tr>
<tr>
<td></td>
<td>4. Kekchi (Kek)*</td>
</tr>
</tbody>
</table>

*These languages are discussed in detail in this paper. Abbreviations for languages are enclosed in parentheses.

Table 1

99
Orthographies for Mayan Languages

(Kaufman 1975)

<table>
<thead>
<tr>
<th>p</th>
<th>t</th>
<th>tz</th>
<th>ty</th>
<th>ch</th>
<th>tx</th>
<th>cy</th>
<th>ky</th>
<th>k</th>
<th>kw</th>
<th>q</th>
</tr>
</thead>
<tbody>
<tr>
<td>p'</td>
<td>t'</td>
<td>tz'</td>
<td>ty'</td>
<td>ch'</td>
<td>tx'</td>
<td>cy'</td>
<td>ky'</td>
<td>k'</td>
<td>kw'</td>
<td>q'</td>
</tr>
<tr>
<td>b'</td>
<td>d'</td>
<td>f</td>
<td>th</td>
<td>s</td>
<td>xh</td>
<td>x</td>
<td>sh</td>
<td>j</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>n</td>
<td>ŋ</td>
<td>ŋ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>l</td>
<td>r</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Loans: b, d, g, rr

<table>
<thead>
<tr>
<th>i</th>
<th>ñ</th>
<th>u</th>
<th>uu</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>a</td>
<td>o</td>
<td>ee</td>
</tr>
</tbody>
</table>

Explanations

Symbols have their normal phonetic values except for the following:

- C' = [¢] glottalized ejective consonant.
- VV = [Vo] long vowel.
- 7 = [?] glottal stop.
- ŋ = [ŋ] velar nasal.
- Cy = [ç] palatal consonant.
- Cw = [cw] labial consonant.
- j = [ʃ] uvular fricative.
- th = [θ] theta.
- tz = [ǯ] split apico-alveolar affricate.
- tx = [ç] retroflex palatal affricate.
- cy = [c] lamino-palatal affricate.
- ch = [c] lamino-palatal affricate; except in Todos Santos Mam and Chajul Ixil where it is [¢], a groove apico-alveolar affricate contrasting with Cy and tz.
- xh = [x] lamino-palatal fricative; except in Todos Santos Mam where it is groove apico-alveolar.
- x = [ç] retroflex fricative in Mam, Agt, Ixl, Jac, Kan, and Aet where it contrasts with xh.
- = [ʃ] lamino-palatal fricative in all other languages where there is no contrast between xh and x.
- sh = [ʃ] lamino-palatal fricative occurring only in Todos Santo Mam, where it contrasts with xh and x.
- ā = [衙] or [ə] depending on the language.

Table 2
Mayan Languages Exhibiting Nominative/Accusative Structures
(Split-Ergativity)

<table>
<thead>
<tr>
<th>Lang</th>
<th>Nom/Acc</th>
<th>Incomp</th>
<th>Subord</th>
<th>Prog</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hua</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yuc</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lac</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chr</td>
<td>(x)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chl</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Tzl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tzt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chj</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Jac</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Mot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Mam</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Agt</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Ixl</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Qui</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tzu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pcm</td>
<td>x</td>
<td></td>
<td></td>
<td>=</td>
<td>x</td>
</tr>
<tr>
<td>Pch</td>
<td>x</td>
<td></td>
<td></td>
<td>=</td>
<td>x</td>
</tr>
<tr>
<td>Kek</td>
<td>x</td>
<td></td>
<td></td>
<td>=</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 3

101
Basic Word Order in Mayan Languages

**V O S:**
(V P A)
Yuc, Lac, Mop
Chl
Tzl, Tzt
Toj, Chj

**V S O:**
(V A P)
Chj, Jac
Mam, Agt, Ixl
Cak

**S V O:**
(A V P)
Chr

**V O S/V S O:**
(V P A/V A P)
Hua, Tenejapa Tzl

N.B. All of the V O S (V P A) languages commonly allow subjects (agents) to occur before the verb under topicalization; some of the V S O (V A P) languages do as well. Some languages also allow topicalized objects (patients) to occur before the verb. Thus, other commonly occurring (nonbasic) word orders are S V O and S O V, and occasionally, O V S. In Hua and Tenejapa Tzl, V S O is the basic order when agent and patient are equal on the animacy hierarchy, and V O S when the agent is superior to the patient. In Chj and Cak, there are dialects with basic V S O, and others with basic V O S.

Table 4
Verb Structures in Mayan

<table>
<thead>
<tr>
<th>Transitive Verb (TV)</th>
<th>1) asp + erg + TVstem (+ mod) + abs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yuc, Lac</td>
</tr>
<tr>
<td></td>
<td>Tzl, Tzt</td>
</tr>
<tr>
<td></td>
<td>Chr, Chl</td>
</tr>
<tr>
<td></td>
<td>Toj</td>
</tr>
<tr>
<td></td>
<td>Ixl</td>
</tr>
<tr>
<td>2) asp + abs + erg + TVstem (+ mod)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hua</td>
</tr>
<tr>
<td></td>
<td>Tzt</td>
</tr>
<tr>
<td></td>
<td>Chj, Jac</td>
</tr>
<tr>
<td></td>
<td>Mam, Agt</td>
</tr>
<tr>
<td></td>
<td>Qui, Cak, Tzu</td>
</tr>
<tr>
<td></td>
<td>Pcm, Pch</td>
</tr>
<tr>
<td></td>
<td>Kok</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intransitive Verb (IV)</th>
<th>1) asp + IVstem (+ mod) + abs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yuc, Lac</td>
</tr>
<tr>
<td></td>
<td>Tzl</td>
</tr>
<tr>
<td></td>
<td>Chr, Chl</td>
</tr>
<tr>
<td></td>
<td>Toj</td>
</tr>
<tr>
<td></td>
<td>Ixl</td>
</tr>
<tr>
<td>2) asp + abs + IVstem (+ mod)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hua</td>
</tr>
<tr>
<td></td>
<td>Tzt</td>
</tr>
<tr>
<td></td>
<td>Chj, Jac</td>
</tr>
<tr>
<td></td>
<td>Mam, Agt</td>
</tr>
<tr>
<td></td>
<td>Qui, Cak, Tzu</td>
</tr>
<tr>
<td></td>
<td>Pcm, Pch</td>
</tr>
<tr>
<td></td>
<td>Kok</td>
</tr>
<tr>
<td>3) asp + erg + IVstem (+ mod)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yuc, Lac</td>
</tr>
<tr>
<td></td>
<td>Chr, Chl</td>
</tr>
<tr>
<td></td>
<td>Jac</td>
</tr>
<tr>
<td></td>
<td>Mam, Ixl</td>
</tr>
<tr>
<td></td>
<td>Pcm, Pch</td>
</tr>
</tbody>
</table>

N.B. 'Directional' particles may occur between some of the morpheme slots indicated above in some of the languages.

Table 5

103
## Absolutive Affixes (Set B)

<table>
<thead>
<tr>
<th>Lang</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>Plexcl</th>
<th>Plincl</th>
<th>P2</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hua</td>
<td>in</td>
<td>it</td>
<td>ū/u</td>
<td>u</td>
<td>ix</td>
<td>u..chik</td>
<td>chik</td>
</tr>
<tr>
<td>Yuc</td>
<td>on</td>
<td>ech</td>
<td>ū</td>
<td>-o7on</td>
<td>-e7ex</td>
<td>-07ob'/ū</td>
<td></td>
</tr>
<tr>
<td>Lac</td>
<td>en</td>
<td>ech</td>
<td>ū</td>
<td>-oon</td>
<td>-eeex</td>
<td>-i79/ū</td>
<td></td>
</tr>
<tr>
<td>Itz</td>
<td>en</td>
<td>ech</td>
<td>ū</td>
<td>-o7on</td>
<td>-e7ex</td>
<td>-oo7</td>
<td></td>
</tr>
<tr>
<td>Chr</td>
<td>in-</td>
<td>i-</td>
<td>a-</td>
<td>ka-</td>
<td>ix-</td>
<td>a..ob'</td>
<td></td>
</tr>
<tr>
<td>Chl</td>
<td>en</td>
<td>et</td>
<td>ū</td>
<td>-onlan</td>
<td>-onla</td>
<td>-etla</td>
<td>-ob'</td>
</tr>
<tr>
<td>Tzl</td>
<td>on</td>
<td>at</td>
<td>ū</td>
<td>-otikotik</td>
<td>-otik</td>
<td>-ex</td>
<td>-ik/la(j)</td>
</tr>
<tr>
<td>Tzt</td>
<td>i-</td>
<td>ot</td>
<td>ū</td>
<td>i..kutik</td>
<td>-unkutik</td>
<td>-a..ik</td>
<td>-ik</td>
</tr>
<tr>
<td>Toj</td>
<td>on</td>
<td>a-</td>
<td>ū</td>
<td>-otik</td>
<td>-ex</td>
<td>-e7</td>
<td></td>
</tr>
<tr>
<td>Chj</td>
<td>(h)in</td>
<td>(h)ach</td>
<td>ū</td>
<td>(h)on</td>
<td>(h)ex</td>
<td>(h)eb'</td>
<td></td>
</tr>
<tr>
<td>Jac</td>
<td>(h)in</td>
<td>(h)ach</td>
<td>ū</td>
<td>(h)on</td>
<td>(h)ex</td>
<td>(h)eb'</td>
<td></td>
</tr>
<tr>
<td>Mam</td>
<td>(ch)in..(y)a</td>
<td>ū</td>
<td>(q)o..(y)a</td>
<td>(q)o7-</td>
<td>i..(y)a</td>
<td>(ch)i-</td>
<td>-e7ya</td>
</tr>
<tr>
<td>Ix1</td>
<td>in</td>
<td>axh</td>
<td>ū</td>
<td>o7</td>
<td>ex</td>
<td>ū</td>
<td></td>
</tr>
<tr>
<td>Qui</td>
<td>in-</td>
<td>at</td>
<td>ū</td>
<td>oj-</td>
<td>ix-</td>
<td>e7-/e-</td>
<td></td>
</tr>
<tr>
<td>Cak</td>
<td>in-/i-</td>
<td>at</td>
<td>ū</td>
<td>oj-</td>
<td>ix-</td>
<td>e7-/o-</td>
<td></td>
</tr>
<tr>
<td>Tzu</td>
<td>in-</td>
<td>at</td>
<td>ū</td>
<td>oq-</td>
<td>ix-</td>
<td>e7-/o-</td>
<td></td>
</tr>
<tr>
<td>Pcm</td>
<td>(h)in</td>
<td>ti-(h)at</td>
<td>ū</td>
<td>aj/ah</td>
<td>ti..ta</td>
<td>i-/ū</td>
<td></td>
</tr>
<tr>
<td>Pch</td>
<td>in-</td>
<td>at-</td>
<td>ū</td>
<td>oj-</td>
<td>at..taq</td>
<td>i..tago</td>
<td>-e7/eb'</td>
</tr>
<tr>
<td>Kok</td>
<td>in</td>
<td>at</td>
<td>ū</td>
<td>ch/o/oo</td>
<td>ex</td>
<td>e7/eb'</td>
<td></td>
</tr>
</tbody>
</table>

Forms marked with a preceding dash are suffixes; with a following dash they are prefixes; with no dash they are both prefixes and suffixes.

When no 1st person plural inclusive form occurs, then there is no distinction between inclusive and exclusive.

Table 6
Preconsonantal Ergative Prefixes (Set A)

<table>
<thead>
<tr>
<th>Lang</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>Flexcl</th>
<th>Plincl</th>
<th>P2</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hua</td>
<td>u</td>
<td>a</td>
<td>in</td>
<td>i</td>
<td>i/a..chik</td>
<td>in..chik</td>
<td></td>
</tr>
<tr>
<td>Yuc</td>
<td>in</td>
<td>a</td>
<td>u</td>
<td>k</td>
<td>a..e7ex</td>
<td>u..o7ob'</td>
<td></td>
</tr>
<tr>
<td>Lac</td>
<td>in</td>
<td>a</td>
<td>u</td>
<td>i..o7</td>
<td>ėk..eex</td>
<td>a..eex</td>
<td>u..o7</td>
</tr>
<tr>
<td>Itz</td>
<td>in</td>
<td>a</td>
<td>u</td>
<td>k</td>
<td>a..e7ex</td>
<td>u..oo7</td>
<td></td>
</tr>
<tr>
<td>Chr</td>
<td>in/ni</td>
<td>a</td>
<td>u</td>
<td>ka</td>
<td>i</td>
<td>u..ob'</td>
<td></td>
</tr>
<tr>
<td>Chl</td>
<td>k/j</td>
<td>a</td>
<td>i</td>
<td>k..lojon</td>
<td>lak</td>
<td>lą7</td>
<td>i..ob'</td>
</tr>
<tr>
<td>Tzl</td>
<td>j</td>
<td>a</td>
<td>s</td>
<td>j..kotik</td>
<td>j..tik</td>
<td>a..ik</td>
<td>s..ik</td>
</tr>
<tr>
<td>Tzt</td>
<td>j</td>
<td>a</td>
<td>s</td>
<td>j..kutik</td>
<td>j..tik</td>
<td>a..ik</td>
<td>s..ik</td>
</tr>
<tr>
<td>Toj</td>
<td>j</td>
<td>ha</td>
<td>s</td>
<td>j..tikon</td>
<td>j..tik</td>
<td>ha..ex/ik</td>
<td>s..e7</td>
</tr>
<tr>
<td>Chj</td>
<td>hin</td>
<td>ha</td>
<td>s</td>
<td>ko/ki</td>
<td>he</td>
<td>s..heb'</td>
<td></td>
</tr>
<tr>
<td>Jac</td>
<td>hin</td>
<td>ha</td>
<td>s</td>
<td>ko</td>
<td>he</td>
<td>s..heb'</td>
<td></td>
</tr>
<tr>
<td>Mam</td>
<td>n..(y)a</td>
<td>t..(y)a</td>
<td>t</td>
<td>q..(y)a</td>
<td>q</td>
<td>ky..(y)a</td>
<td>ky</td>
</tr>
<tr>
<td>Agt</td>
<td>in</td>
<td>a</td>
<td>s/xh/ x/ᵩ</td>
<td>qa</td>
<td>i</td>
<td>chi</td>
<td></td>
</tr>
<tr>
<td>Ixl</td>
<td>in/un</td>
<td>a</td>
<td>i</td>
<td>ku/qu</td>
<td>e</td>
<td>i</td>
<td></td>
</tr>
<tr>
<td>Qui</td>
<td>in/nu</td>
<td>a(a)</td>
<td>u(u)</td>
<td>qa</td>
<td>i(i)</td>
<td>ki</td>
<td></td>
</tr>
<tr>
<td>Cak</td>
<td>in/nu</td>
<td>a</td>
<td>ru/u</td>
<td>qa</td>
<td>i</td>
<td>ki</td>
<td></td>
</tr>
<tr>
<td>Tzu</td>
<td>in/n(uu)</td>
<td>a(a)</td>
<td>r(uu)/ uu/ᵩ</td>
<td>qa(a)</td>
<td>e(e)</td>
<td>ki/kee</td>
<td></td>
</tr>
<tr>
<td>Pcm</td>
<td>nu</td>
<td>(h)a</td>
<td>ru</td>
<td>qa</td>
<td>(h)a..ta</td>
<td>ki</td>
<td></td>
</tr>
<tr>
<td>Pch</td>
<td>in/ni/n</td>
<td>a</td>
<td>ri/ri</td>
<td>qa</td>
<td>a..taq</td>
<td>ki</td>
<td></td>
</tr>
<tr>
<td>Kek</td>
<td>in</td>
<td>a(a)</td>
<td>(i)x</td>
<td>qa</td>
<td>e(e)</td>
<td>x..eb’</td>
<td>eʔx</td>
</tr>
</tbody>
</table>

Forms separated by two dots are discontinuous morphemes, the first being a prefix and the second a suffix or enclitic.

Table 7
# Provocalic Ergative Prefixes (Set A)

<table>
<thead>
<tr>
<th>Lang</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>Plexcl</th>
<th>Plincl</th>
<th>P2</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hua</td>
<td>u</td>
<td>a</td>
<td>in</td>
<td>i</td>
<td>a..chik</td>
<td>in..chik</td>
<td></td>
</tr>
<tr>
<td>Yuc</td>
<td>inw</td>
<td>aw</td>
<td>uy</td>
<td>k</td>
<td>aw..e7ex</td>
<td>uy..e7ob'</td>
<td></td>
</tr>
<tr>
<td>Lac</td>
<td>(in)w</td>
<td>aw</td>
<td>(u)y</td>
<td>(in)w..07</td>
<td>k..eox</td>
<td>aw..eox</td>
<td>(u)y..07</td>
</tr>
<tr>
<td>Itz</td>
<td>(in)w</td>
<td>aw</td>
<td>(u)y</td>
<td>k</td>
<td>aw..e7ex</td>
<td>(u)y..007</td>
<td></td>
</tr>
<tr>
<td>Chr</td>
<td>inw/niw</td>
<td>aw</td>
<td>uw/uy</td>
<td>kaw</td>
<td>iw/iy</td>
<td>i</td>
<td>uw/uy..ob'</td>
</tr>
<tr>
<td>Chl</td>
<td>k</td>
<td>aw</td>
<td>(i)y</td>
<td>k..lojon</td>
<td>la7w</td>
<td>(i)y..ob'</td>
<td></td>
</tr>
<tr>
<td>Tzl</td>
<td>k</td>
<td>aw</td>
<td>y</td>
<td>k..kotik</td>
<td>k..tik</td>
<td>aw..ik</td>
<td>y..ik</td>
</tr>
<tr>
<td>Tzt</td>
<td>k</td>
<td>aw</td>
<td>y</td>
<td>k..kutik</td>
<td>k..tik</td>
<td>aw..ik</td>
<td>y..ik</td>
</tr>
<tr>
<td>Toj</td>
<td>k</td>
<td>(ha)w</td>
<td>y</td>
<td>k..tikon</td>
<td>k..tik</td>
<td>(ha)w..{ex ik}y..c7</td>
<td></td>
</tr>
<tr>
<td>Chj</td>
<td>w</td>
<td>h</td>
<td>y</td>
<td>k</td>
<td>hey</td>
<td>y..heb'</td>
<td></td>
</tr>
<tr>
<td>Jac</td>
<td>w</td>
<td>haw</td>
<td>y</td>
<td>j</td>
<td>hey</td>
<td>y..heb'</td>
<td></td>
</tr>
<tr>
<td>Mam</td>
<td>w..(y)a</td>
<td>t..(y)a</td>
<td>t</td>
<td>q..(y)a</td>
<td>q</td>
<td>ky..(y)a</td>
<td>ky</td>
</tr>
<tr>
<td>Agt</td>
<td>w</td>
<td>aw</td>
<td>t</td>
<td>q</td>
<td>it</td>
<td>ky</td>
<td></td>
</tr>
<tr>
<td>Ixl</td>
<td>w</td>
<td>aw/a</td>
<td>t</td>
<td>q</td>
<td>et</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>Qui</td>
<td>w/inw</td>
<td>(a)aw</td>
<td>r</td>
<td>q</td>
<td>(i)iw</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>Cak</td>
<td>w/inw</td>
<td>aw</td>
<td>r</td>
<td>q</td>
<td>iw</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>Tzu</td>
<td>w/(i)nw</td>
<td>(a)aw</td>
<td>r</td>
<td>q</td>
<td>(e)cw</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>Pcm</td>
<td>w</td>
<td>(h)aw</td>
<td>r</td>
<td>q</td>
<td>(h)aw..ta</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>Pch</td>
<td>w</td>
<td>aw</td>
<td>r</td>
<td>q</td>
<td>aw..taq</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>Kek</td>
<td>w</td>
<td>(a)aw</td>
<td>(i)r</td>
<td>q</td>
<td>(e)er</td>
<td>r..eb'</td>
<td>e7r</td>
</tr>
</tbody>
</table>

Table 8

106
<table>
<thead>
<tr>
<th>Lang</th>
<th>Simple Pass</th>
<th>Other Pass</th>
<th>Medio-Pass</th>
<th>Adj Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hua</td>
<td>-aab'</td>
<td>-n</td>
<td>-n</td>
<td>-b'il-aab'</td>
</tr>
<tr>
<td>Yuc</td>
<td>-7v/-a7</td>
<td>-p -k'</td>
<td>-7v-</td>
<td>-b'il</td>
</tr>
<tr>
<td>Lac</td>
<td></td>
<td></td>
<td>-b'il</td>
<td></td>
</tr>
<tr>
<td>Chr</td>
<td>-j/-n</td>
<td>-t -tz'</td>
<td>-t -k' -p</td>
<td>-b'ir</td>
</tr>
<tr>
<td>Chl</td>
<td>-j--t/-nt</td>
<td></td>
<td>-b'il</td>
<td></td>
</tr>
<tr>
<td>Tzl</td>
<td>-ot</td>
<td>-j-</td>
<td>-b'il</td>
<td></td>
</tr>
<tr>
<td>Tzt</td>
<td>-at</td>
<td>-e</td>
<td>-b'il</td>
<td></td>
</tr>
<tr>
<td>Toj</td>
<td>-j</td>
<td>-x</td>
<td>-ub'al</td>
<td></td>
</tr>
<tr>
<td>Chj</td>
<td>-(ch)aj</td>
<td>-ax</td>
<td>-b'il</td>
<td></td>
</tr>
<tr>
<td>Jac</td>
<td>-ot</td>
<td>-lax -cha -lo</td>
<td>-b'il</td>
<td></td>
</tr>
<tr>
<td>Mam</td>
<td>-eet -j--l</td>
<td>-njwtz -b'aj</td>
<td>φ</td>
<td>-(V)7n</td>
</tr>
<tr>
<td>Ixl</td>
<td>-1- ch/-l</td>
<td>-p -Vx</td>
<td>-el/-mal</td>
<td></td>
</tr>
<tr>
<td>Agt</td>
<td>-lij/-ij</td>
<td>-xij</td>
<td>-ij/-7n</td>
<td></td>
</tr>
<tr>
<td>Qui</td>
<td>-V/-x</td>
<td>-(V)taj -Vr</td>
<td>-oon/-Vn</td>
<td></td>
</tr>
<tr>
<td>Cak</td>
<td>-V/-x</td>
<td>-(V)taj -Vr -e7</td>
<td>-on/-n</td>
<td></td>
</tr>
<tr>
<td>Tzu</td>
<td>-j/-x</td>
<td>-(V)taj -Vr</td>
<td>-7--j-</td>
<td>-oon/-Vn</td>
</tr>
<tr>
<td>Pcm</td>
<td>-Vr/-j</td>
<td>-aam'j</td>
<td>-ooj/-Vmaj</td>
<td></td>
</tr>
<tr>
<td>Pch</td>
<td>-Vr/-j</td>
<td>-mVj -Vb'</td>
<td>-ooj/-Vmaj</td>
<td></td>
</tr>
<tr>
<td>Kek</td>
<td>-e7/-V</td>
<td>-man</td>
<td>-(m)b'il -C10</td>
<td></td>
</tr>
</tbody>
</table>

Table 9

107
### Antipassive Voices in Mayan

<table>
<thead>
<tr>
<th>Lang</th>
<th>Absolutive</th>
<th>Incorporating</th>
<th>Focus</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hua</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yuc</td>
<td>-V-(-naj)/-aj(-naj)</td>
<td>Ø(-naj)</td>
<td>Ø + NO {agt}</td>
<td></td>
</tr>
<tr>
<td>Lac</td>
<td>Ø + IV mode suff</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chr</td>
<td>-on/-wan</td>
<td></td>
<td>-m</td>
<td></td>
</tr>
<tr>
<td>Chl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tzl</td>
<td>-awan</td>
<td></td>
<td>-(o)maj</td>
<td>-(V)wej</td>
</tr>
<tr>
<td>Tzt</td>
<td>-wan</td>
<td></td>
<td>-(V)waj</td>
<td></td>
</tr>
<tr>
<td>Toj</td>
<td>-wan ~ -wun</td>
<td></td>
<td>-(V)waj</td>
<td></td>
</tr>
<tr>
<td>Chj</td>
<td>-waj</td>
<td>-w</td>
<td>-an</td>
<td></td>
</tr>
<tr>
<td>Jac</td>
<td>-wa</td>
<td>-w</td>
<td>-n</td>
<td>-l</td>
</tr>
<tr>
<td>Mii</td>
<td>-(VV)n</td>
<td>-(VV)n</td>
<td>-(VV)n</td>
<td></td>
</tr>
<tr>
<td>Agt</td>
<td>-V-oon/-Vn</td>
<td>-oon/-Vn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ix1</td>
<td>-on/-n</td>
<td>-on/-n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qui</td>
<td>-oon/-Vn</td>
<td>-on/-Vn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tzu</td>
<td>-oon/-Vn</td>
<td>-o/-n</td>
<td>-ow/-Vn</td>
<td></td>
</tr>
<tr>
<td>Cak</td>
<td>-on/-n</td>
<td>-o/-n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pcm</td>
<td>-w/-in</td>
<td>-w/-in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pch</td>
<td>-w/-Vn</td>
<td>-w/-Vn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kek</td>
<td>-o/-n</td>
<td>-o/-n</td>
<td>-o/-n</td>
<td></td>
</tr>
</tbody>
</table>

**Table 10**
### Referential and Instrumental Voices

<table>
<thead>
<tr>
<th>Lang</th>
<th>Referential</th>
<th>Instrumental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hua</td>
<td>-ch</td>
<td>-chinch</td>
</tr>
<tr>
<td>Yuc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lac</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Itz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chr</td>
<td>-♂</td>
<td></td>
</tr>
<tr>
<td>Chl</td>
<td>-b'e</td>
<td></td>
</tr>
<tr>
<td>Tz1</td>
<td>-b'e</td>
<td></td>
</tr>
<tr>
<td>Tzt</td>
<td>-b'c</td>
<td></td>
</tr>
<tr>
<td>Toj</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chj</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jac</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ham</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ixl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qui</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cak</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tzu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pcm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kek</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11
References


ALEJANDRE, MARCELO. 1890. Cartilla huasteca con su gramática, dicionario, y varias reglas para aprender el idioma. Mexico: Tipografía de la Secretaría de Fomento.


____________. 1957. A grammar of modern Yucatec. MCMCA 41.


AYRES, GLENN. 1977. Es el antipasivo siempre una voz? Unpub. paper presented at Taller Maya II, Chiapas, Mexico.


BURGESS, DORA AND DAVID FOX. 1966. Quiché. LOG, pp. 49-86.


   By Norman McQuown. Austin: Univ. of Texas Press.


JACOBSSEN, WILLIAM. 1969. The analog of the passive transformation in ergative-type languages. Paper read at the 44th annual meeting, LSA, San Francisco, California.


KEENAN, EDWARD. 1976. Towards a universal definition of 'subject'. In Li (1976), 303-333.


___________ and WILLIAM M. NORMAN. 1979. Correlates of ergativity in Mayan grammar. Ergativity, ed. by Frans Planck.


PECK, DUDLEY and DOROTHY, and EDWARD SYWALKA. 1966. Mam. LOG, 166-95.


PERLMUTTER, DAVID and PAUL POSTAL. 1974. Notes on relational grammar. LSA Summer Linguistics Institute, Amherst.


ULRICH, MATTHEW, and ROSEMARY ULRICH. 1966. Mopan Maya. LOG, 251-70.


WHITTAKER, ARABELLE and VIOLA WARKENTIN. 1965. Chol texts on the supernatural. Norman: SIL.


XIMÉNEZ, FRAY FRANCISCO. Ca. 1701-1703. Arte de las tres lenguas Cakchiquel, Quiché y Tzutuhil. MCMCA 36.
Aguacatec Syntax from a Functional Perspective

Thomas W. Larsen

University of California, Berkeley

This paper reports the findings of a preliminary investigation of Aguacatec syntax from a functional perspective. The investigation is preliminary in that it is based on a restricted body of data, mostly from a single text. The perspective is said to be functional in that an attempt is made to explain certain syntactic phenomena found in the text in terms of certain pragmatic notions which determine the structure of narrative discourse in the language. In particular, we will be looking at some of the "packaging phenomena" or "statuses" that nouns may have in a discourse, as discussed by Chafe (1976). These packaging phenomena include such notions as "given" vs. "new", "contrastiveness", "definite" vs. "indefinite", "subject", "topic", and "point of view". There will also be some attention given to the notions of "transitivity", "ergativity", and "grounding". [1]

Most of the data comes from a story about Lu? Tzuu7 or Pedro Tecomate (tzuu7 'tecomate' is a type of hourglass shaped gourd; Pedro Tecomate is more generally known as Pedro Riales or Pedro Urdemales in Spanish) and a group of a?pyaa7 'travelling salesmen' (pyaa7 'journey' < Spanish viaje; a7 'one characterized by') [2]. A copy of this text is appended to the end of this paper. Hypotheses formed on the basis of this text were checked against, and in some cases revised on the basis of, data from a number of other texts. Three of these are tape recorded texts entitled "Lu? Tzuu7 nin Paalee7" 'Pedro Tecomate and the Priest'
(hereafter called LTzP), "Ye Q'ankyooq" 'The Thunder' (hereafter YQ'), and an untitled story about a naawloon (hereafter N; a naawloon is a person who can make his spirit, or naawl, leave his body and go out to do good and/or evil; naawl < Aztec naawalli 'mask, disguise; shaman, wizard, sorcerer; specter, totem, animal double, alter-ego', Andrews 1975:455). These three texts were recorded and transcribed by Aguacatec speaking students at the Proyecto Lingüístico Francisco Marroquín in Guatemala. Four others are texts published by the Summer Institute of Linguistics: "Yi Aj Cabinl Tu Yi Umul" 'The beekeeper and the rabbit' (L. McAthurn 1973:12-16; hereafter AU), "Aj Ranch Tc'u'l Txuc" 'The rancher inside the animal' (L. McAthurn 1973:17-21; hereafter AtTx), "Yi Jale'n Quipisio' E' Tx'i'n' 'When the dogs received their duty' (L. McAthurn 1973:31-4; hereafter JKyxTx'), and "El Hombre Y El Zopilote" (Shaw 1972:279-82; hereafter HZ). Copies of these additional texts have not been included here, but examples will be cited from them when appropriate. The spelling of the examples cited from the SIL texts has been changed to conform to that used in this paper.

Before looking at the text material, however, we will consider some of the basic typological and grammatical characteristics of Aguacatec which are prerequisite to understanding the points to be made later.

**Brief Grammatical Sketch of Aguacatec**

Aguacatec is a Mayan language spoken by perhaps some 14,000 people in the municipio of Aguacatán, Department of Huehuetenango, Guatemala. The language is divided into two main dialects. One, the larger and more prestigious, is called Chalchitec and is spoken in the eastern part
of the municipio. The text appended to the end of this paper as well as $N$ and, possibly, the four SIL texts are in this dialect. The other dialect is called Aguacatec and is spoken in the western part of the municipio. LTzp and YQ are in this dialect. The differences between these two dialects are relatively slight, being confined largely to some minor phonological and lexical differences. Native speakers, however, seem to feel that Chalchitec and the Aguacatec dialect are two separate (though mutually intelligible) languages. In this paper the simple term "Aguacatec" will be used exclusively to refer to the entire language, not to the Aguacatec dialect.

The Aguacatec forms cited here will be spelled according to a practical orthography developed by the Proyecto Lingüístico Francisco Marroquín. The symbols used in this phonemic orthography can be understood by referring to Table 1.

As in other Mayan languages, nouns are not marked for case; however, verbs agree with their subjects and objects according to a split-ergative verb agreement system; and, hence, Aguacatec is usually classified as an "ergative language". In most cases, an intransitive "subject" (or $S$, employing the terminology of Dixon 1979 [3]) and a transitive "direct object" (or $O$) are crossreferenced on their respective verbs by means of a set of "absolutive" prefixes while a transitive "subject" (or $A$) is crossreferenced by means of a set of "ergative" prefixes, as seen in:

(1) ța  
    kxh-ụ71
    proximate past 2sAbs-ARRIVE HERE
    'you arrived'

122
plain stops

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>t</td>
<td>ky</td>
</tr>
<tr>
<td>k</td>
<td>q</td>
<td>7</td>
</tr>
</tbody>
</table>

glottalized stops

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>b'</td>
<td>t'</td>
<td>ky'</td>
</tr>
<tr>
<td>k'</td>
<td>q'</td>
<td></td>
</tr>
</tbody>
</table>

voiceless affricates

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>tz</td>
<td>ch</td>
<td>tx</td>
</tr>
</tbody>
</table>

glottalized affricates

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>tz'</td>
<td>ch'</td>
<td>tx'</td>
</tr>
</tbody>
</table>

voiceless fricatives

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>xh</td>
<td>x</td>
</tr>
<tr>
<td>j</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

nasals

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>n</td>
</tr>
</tbody>
</table>

l laterals

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td></td>
</tr>
</tbody>
</table>

trills

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td></td>
</tr>
</tbody>
</table>

glides

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>w</td>
<td>y</td>
<td></td>
</tr>
</tbody>
</table>

short vowels

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>u</td>
</tr>
</tbody>
</table>

long vowels

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ii</td>
<td>uu</td>
</tr>
<tr>
<td>ee</td>
<td>oo</td>
</tr>
<tr>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

Table 1

(2) /a/ /e/-uu1
prox.past /sAbs-ARRIVE HERE
'he arrived'
Tense and aspect are generally indicated by a preposed particle such as the "proximate past" particle ɛɛ in the above examples. The absence of such particles is interpreted to mean the "indefinite past" tense, which indicates that the action took place sometime before today. This is the tense that is generally used in narrating stories such as the one we are concerned with here. The details of how tense/aspect is indicated and of the various forms of the ergative and absolutive prefixes that are used in different phonological and morphological environments are rather complex and will not be fully dealt with here (see Larsen 1978 for further details). Arguments of stative, equational, and existential predicates are also crossreferenced by absolutive morphemes, whereas the ergative prefixes are used to crossreference noun possessors, as seen in

(5) ooj
'avocado'

(6) aw-ooj
2sErg-AVOCADO
'your avocado'

(7) t-ooj
3sErg-AVOCADO
'his avocado'

In certain types of subordinate clauses 5's are not crossreferenced by the customary absolutive morphemes, however, but rather by the same
ergative prefixes that crossreference A's, thus giving the appearance of a nominative/accusative verb agreement system in those environments. For this reason the verb agreement is said to be "split ergative" (Dixon 1979:79-98). One environment in which this occurs is in time adverbial clauses in the indefinite past tense:

(8) ye aw-uul-e7n,
    THE 2sErg-ARRIVE HERE-nominalizer
    niin tzun na       chin-waan
    AND THEN incomplete 1sAbs-EAT
    'when you arrived, I was eating'

(9) ye t-uul-e7n,...
    THE 3sErg-ARRIVE HERE-nom.
    'when he arrived,...'

(10) ye t-il-oool          axh,...
    THE 3sErg-SEE-active infinitive 2sPro(Abs)
    'when he saw you,...'

(11) ye aw-il-oool   ꖇ,...
    THE 2sErg-SEE-act.inf 3sAbs
    'when you saw him,...'

Here it can be seen that the verb forms found in the adverbial clauses of (8-11) are actually derived verbal nouns and, thus, the ergative prefixes crossreferencing their respective S's and A's may be viewed as being formally noun possessors. O's are still marked by the absolutive marker ꖇ in the third person singular (as in 11) or by one of the absolutive independent pronouns in the other person/numbers (as in 10).

As mentioned previously, arguments of non-verbal sentences in Agua- catec are absolutes. Depending on the type of predicate involved, the absolutive marker may be either preposed in the form of an independent pronoun or postposed. The following illustrate some of the possible
(12) axh  wunaq
    2sPro(Abs) PERSON
    'you are a person'

(13) at-ixh  tzne7j
    EXIST-2sAbs HERE
    'you are here'

(14) txik-1-kixh
    VERTICAL-stative-2sAbs
    'you are standing'

(15) kyim-naq-(k)ixh
    DIE-past participle-2sAbs
    'you are dead'

Other case relations, such as dative, instrument, locative, etc. are usually indicated by "relational nouns". Relational nouns function in much the same way as prepositions do in English; however, unlike prepositions, relational nouns are formally possessed nouns where the possessive prefix refers to the "object" of the relational noun phrase. Many, though not all, relational nouns are preceded by one of the prepositions tz ("s) or t.

(16) ja  e-w-aq'
    prox.past 3sAbs-1sErg-GIVE prep-3pErg-POSSESSION
    yaaj
    MAN
    'I gave it to the men'

(17) ja  chin-b'een tz-aw-uuch'
    prox.past 1sAbs-GO prep-2sErg-COMPANION
    'I went with you'
(18) at-\(\emptyset\) t-\(\emptyset\)-k'u7l
EXIST-3sABS prep-3sErg-BELLY
'It is inside of it'

Many of the relational nouns that begin with the preposition tz lose
that preposition before a third person singular ergative prefix:

(19) na chin-b'een t-uuch'
incompletive 1sAbs-GO 3sErg-COMPANION
'I go with him'

Aguacatec also has a number of other prepositions besides tz and t.
Most of these are merely reduced forms of relational nouns, e.g.

(20) wi 'on (top of)' < tz-\(\emptyset\)-wi7
     prep-3sErg-HEAD
     'on top of it'

(21) tu 'with' < (tz-)t-uuch'
     3sErg-COMPANION
     'with him'

(22) tzi 'at the edge of' < (tz-)s-tzi17 'at its edge'
     3sErg-MOUTH

Two prepositions that do not seem to be related to any relational nouns
are

(23) len 'until, up to' (perhaps < leen 'leaving' (directional; see below) ?)

(24) le 'into' (perhaps also < leen ?)

The basic unmarked word order in Aguacatec is VSO, as in

(25) qa \(\emptyset\)-x-tx'aj xna7n b'u7y
prox.past 3sAbs-3sErg-WASH WOMAN RAG
'the woman washed the rag'
The only exception to this is with reflexives and reciprocals, where the verb is immediately followed by the reflexive/reciprocal relational noun, which is in turn followed by the subject, thus giving the appearance of having VOS order:

(26) ja  y'a-b'iy  t-ilb'  yaaj
    prox.past 3sAbs-3sErg-HIT 3sErg-REFL MAN
    'the man hit himself'

Various marked word orders are also possible, such as when a constituent that ordinarily comes after the verb is placed before the verb. One case in which this happens is in cleft sentences. For example, in the sentence

(27) ja  y'a-x-tx'aq  xna7n b'u7y  jaalu7
    prox.past 3sAbs-3sErg-WASH WOMAN RAG NOW/TODAY
    'the woman washed the rag today'

the time adverb may be clefted, thus deriving

(28) jaalu7  n-y'a-x-tx'aq  xna7n b'u7y
    'it was today that the woman washed the rag'

Note here that in the cleft sentence (28), the proximate past particle ja has been replaced by the prefix m- (→ n in (28) by a regular phonological rule). Proximate past is marked by ja in main clauses and by m- in subordinate clauses. This suggests that the constituent structure of (27) is something roughly like
whereas the constituent structure of (28) is something like

Similarly, with the O clefted, the following sentence is derived:

It happens, however, that the A of a transitive sentence like (27) cannot be clefted in this same simple way. In order to cleft an A in Agua-catec, a special verb form must be used, as seen in

Several things should be noted about the verb stem -tx'a'oon in (32). First of all, in spite of the fact that (32) gives the impression of being a transitive sentence like (27), (28), and (31), the verb -tx'a'oon is intransitive being marked only for absolutive agreement. It also has other morphological characteristics of intransitive verbs;
for example, it takes the suffix -'(q) in the future/potential tense/aspect rather than the suffix -e7 used with transitive verbs. Also the suffix -oon is used to derive other clearly intransitive verb forms. Second of all, though it is impossible to tell which of the third person singular NPs it is that the verb agrees with in (32), it can be seen from sentences like

(33) in n-kxh-b'i y-oon
1sPro(Abs) prox.past-2sAbs-HIT-suffix
'I was the one who hit you'

that it is not the clefted A that the verb agrees with but rather the underlying O. However, when the underlying O is 3s, the intransitive verb may optionally agree with the underlying A, as seen in

(34) in \{m- \delta-b'i y-oon
\{n-chin-\}
1sPro(Abs) prox.past-{3sAbs-}HIT-suffix
\{1sAbs-\}
'I was the one who hit him'

This, plus the fact that clefted A's may appear as one of the absolutive independent pronouns (as seen in 33 and 34; these pronouns may never be "ergative NPs" [4]), demonstrates that the clefted constituents in (32-34) are "absolutive NPs". Thus it can be seen that the rule that forms cleft sentences is sensitive to the ergative/absolutive "relations" in that ergative NPs cannot be clefted; or, using Dixon's terminology, A's cannot be clefted unless they are first put into "derived S function".

The construction illustrated by (32-34) is usually referred to by Mayanists as the "focus antipassive" or "agentive antipassive" construction. The focus antipassive is not a true antipassive construction if
antipassive is taken, as it often is, to mean an intransitive sentence in which the agent is the absolutive S and the patient is optionally present in an oblique case. Aguacatec does have such a "true" antipassive, called the "absolutive antipassive", to be discussed below. Since the verb form of the focus antipassive is nearly identical to that of the absolutive antipassive, the term "antipassive" has come to be used for this construction also. This makes a certain amount of sense anyway in that the so-called focus antipassive in Aguacatec seems to have a function similar to that of antipassives found in other languages like Dyirbal; namely, it is used to convert ergative NPs into absolutes, thus making them accessible to certain rules which apply only to absolute NPs.

The characteristics that we have just seen for cleft sentences are also found with relative clauses and WH-questions; as can be seen in

(35) 'a .dictionary
    $w-\text{i}$  xna7n (ye)
    \text{prox.past 3sAbs-1sErg-SEE WOMAN (THE)}
    m-$w-u7l$
    \text{prox.past-3sAbs-ARRIVE HERE}
'I saw the woman who arrived'.

(36) 'a .dictionary
    $w-\text{i}$  b'u7y (ye)
    \text{prox.past 3sAbs-1sErg-SEE RAG (THE)}
    n-$w-x-\text{tx'aj}$  xna7n
    \text{prox.past-3sAbs-3sErg-WASH WOMAN}
'I saw the rag that the woman washed'.

(37) 'a  dictionary
    $w-\text{i}$  xna7n (ye)
    \text{prox.past 3sAbs-1sErg-SEE WOMAN (THE)}
    n-$w-\text{tx'aj-oon}$  b'u7y
    \text{prox.past-3sAbs-WASH-suff RAG}
'I saw the woman who washed the rag'
(38) na7 m-β-u71
WHO prox.past-3sAbs-ARRIVE HERE
'who arrived?'

(39) na7 m-β-β-b'i y  yaaj
WHO prox.past-3sAbs-3sErg-HIT MAN
'who did the man hit?'

(40) na7 m-β-b'i y-oon  yaaj
WHO prox.past-3sAbs-HIT-antipass MAN
'who hit the man?'

It can be seen from this that in addition to morphological ergativity, Aguacatec can also be said to have some ergative syntax since the syntactic rules of clefting, relative clause formation, and WH-question may apply to absolutive NPs but not to ergative NPs. (Using the terminology of Dixon 1979, one could say that these rules operate on an "S/O pivot" rather than an "S/A pivot".)

In addition to the focus antipassive construction, Aguacatec also has a true antipassive, which is usually called the "absolutive antipassive". For example, corresponding to

(41) ja  ʃ-a-tzok'  sii7
prox.past 3sAbs-2sErg-CUT FIREWOOD
'you cut firewood'

there is

(42) ja  kxh-tzook'-oon  t-etz  sii7
prox.past 2sAbs-CUT-antipass 3sErg-POSS FIREWOOD
'you cut (on) the firewood'

Here the ergative A in (41) becomes an absolutive S in (32), and the absolutive O in (41) becomes an oblique patient in (42) introduced by the dative relational noun tz- -etz (many speakers, at least some of
the time, use the relational noun tz- e7 ('on, at, with, about' rather than the dative). This oblique patient is optional so that the absolutive antipassive may be used to form an intransitive sentence with agent S and no patient in the same way that the passive can be used to form an intransitive sentence with patient S and no agent. The verb form in (42) is nearly the same as that used in the focus antipassive, the only difference being that the root vowel of most non-derived transitive verb roots lengthens in the absolutive antipassive (there are a few transitive roots that do not do this for some reason). With derived transitive verb stems the focus and absolutive antipassive verb forms are always indistinguishable. When the patient is present in the absolutive antipassive construction, as it is in (43), the meaning is usually slightly different from the meaning of the regular transitive sentence.

This difference usually seems to involve the fact that the result of the action on the patient has not been fully realized, as suggested by the use of the preposition on in the translation of (42). With a few transitive verbs, the absolutive antipassive form without patient may take on a reflexive meaning, as in

(43) \( \text{a} \)  ꩠ-tx'aa-j-oon  
\( \text{prox.past 3sAbs-WASH-antipass} \)  
'he washed (things/himself)'

The passive in Aguacatec has two forms with transitive verb roots, as shown in

(44) \( \text{a} \)  ꩠ-b'i-y-1-i\( j \)  yaa-j aw-a7\( n \)  
\( \text{prox.past 3sAbs-HIT-pass-intr.suff MAN 2sErg-AGENT} \)  
'the man was hit by you'
(45) ꝏ-b'iy-x-ij yaj aw-a7n
   'the man was hit by you'

The difference between the two forms seems to be that the form in -l (or -ch with roots containing an ı) emphasizes the result of the action on the patient (or perhaps takes the patient's "point of view"), whereas the form in -x puts more emphasis on the fact that the action was caused by an agent regardless of whether or not one is explicitly mentioned. Both forms may optionally take an oblique agent phrase as shown in (44-5). The relational noun -aaq'un (lit. 'work') may be used in place of -a7n in the agent phrase. With derived transitive verb stems, there is only one passive form:

(46) ꝏ-k'aay-ıj aw-a7n
    prox.past 3sAbs-SELL-intr.suff 2sErg-AGENT
    'it was sold by you'

from

(47) ꝏ-a-k'aay
    prox.past 3sAbs-2sErg-SELL
    'you sold it'

(the stem is derived from the noun k'a7î 'sale').

Another common syntactic construction in Aguacatec takes the form of a main clause, usually with a finite verb, followed by a complement clause with a nominalized verb form. This complement clause is introduced by tan, a reduced form of the relational noun t-a7n (3sErg-AGENT) which is always used when this relational noun is followed by its "object" NP. All such nominalized clauses introduced by tan will here be referred to as "purpose clauses". Some examples of purpose clauses
are

(48) ja chin-a-chaq
prox.past 1sAbs-2sErg-ORDER
at x-tx'aj-l-e7n b'u7y
Comp 3sErg-WASH-passive-infinite RAG
'You ordered me to wash the rag'

(49) ja n-o7k tan wa-a7n
prox.past 1sAbs-Enter Comp EAT-infin
'I began to eat'

(50) chin-b'een tan 2-loq'-ch-e7n txiikun
1sAbs-GO Comp 3sErg-BUY-pass-infin BEAN
'I will go to buy beans'

(51) ja q-iky' junt tiir
prox.past 1pAbs-PASS ANOTHER TIME
at q-oopoon-e7n
Comp 1pAbs-ARRIVE THERE-infin
jalen Antigua
until Antigua
'We passed on again in order to reach Antigua'
(H. McArthur 1979:109)

(52) 6 qa-k'wuch-e7 t-etz
3sAbs-1pErg-PRAY-potential 3sErg-POSS
qa-taaj tan xh-ch'eey-aal
1pErg-FATHER Comp 3sErg-HELP-active infinitive
o7
1pPro(Abs)
'we will pray (it) to God so that he will
help us'

(53) 6 qa-k'wuch-e7 t-etz qa-taaj
3sAbs-1pErg-PRAY-pot 3sErg-POSS 1pErg-FATHER
tan qa-ch'eey-e7n
Comp 1pErg-HELP-passive infinitive
'we will pray (it) to God so that he will
help us'
It can be seen in these examples that in purpose clauses, both A's and S's may be deleted by EQUI (as in 48, 49, 50, 53) if they are coreferential with a matrix clause S (as in 49 and 50), a matrix clause O (as in 48), or a matrix clause indirect object (as in 53). Examples (51) and (52) show that in at least some of these cases, EQUI is optional. If there is no coreferential S, O, or indirect object in the matrix clause (cf. 54), then EQUI cannot apply, even if there is a coreferential A (as in 54) [6].

Another thing to note about these examples is the form of the verbs in the purpose clauses. As noted above, purpose clauses have nominalized verb forms; however, there are three different nominalizations which appear in exx. (48-54). When the A of a transitive purpose clause is deleted by EQUI, as in (48, 50, and 53), the verb takes a form called the "passive infinitive". Passive infinitives are always derived with the nominalizing suffix -e7n and have a possessive prefix coreferential with the underlying O. Furthermore, those passive infinitives which are derived from transitive verb roots (as opposed to derived transitive verb stems such as the one in 53) show an overt passive suffix (e.g. -1 in 48 and -ch in 50). Thus, it appears that the transitive purpose clause has been passivized with the underlying O "advancing" to S and the underlying A being deleted. Since the derived passive intransitive verb form is nominalized, its S is crossreferenced by an ergative possessive prefix. When the A of a transitive purpose clause is not
deleted by EQUI, another nominalized verb form may be used. This is
called the "active infinitive", derived by a suffix -\textit{VV}, as seen in
(52) and (54). This is the same verb form that was seen in (10) and
(11). The ergative possessive prefix crosses the A, and the O
is crossreferenced by an absolutive morpheme following the verb. Actu-
ally, purpose clauses like those in (52) and (54) seem to be relatively
rare. More often one finds that the passive infinitive is used here
too; however, since the underlying A is not deleted by EQUI, it appears
as an oblique agent phrase. An example of this, corresponding to (54),
is (cf. line 80 of appended text):

(55) ṣqa-toy-e7 puuntiš tan
        ṣAbs-1pErg-FIND-pot WAY comp
       ṣ-xub's-eʔn q-aʔn
       ṣErg-TRICK-(pass)infin lErg-AGENT
'we will find a way to trick him'
(lit. 'we will find a way for his being
tricked by us.')

When the purpose clause is intransitive, as in (49) and (51), the verb
takes a form called the "intransitive infinitive". In most cases, the
intransitive infinitive is identical in form to the nominalized verb
forms seen in (8) and (9), which are called "intransitive verbal nouns".
However, there are a few intransitive verbs which have irregular infiniti-
tives but regular verbal nouns, e.g.

<table>
<thead>
<tr>
<th>verbal noun</th>
<th>infinitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>-wit 'sleep'</td>
<td>wit-eʔn waatl</td>
</tr>
<tr>
<td>-waan 'eat'</td>
<td>waan-eʔn waaʔn (Chalchitec)</td>
</tr>
<tr>
<td></td>
<td>waaʔaʔn (Aguacatec dialect)</td>
</tr>
</tbody>
</table>
If the S is not deleted by EQUI, it is crossreferenced on the intransitive infinitive by means of an ergative possessive prefix (as in 51). If the S is deleted by EQUI, there is no prefix on the infinitive (as in 49).

It can be seen, then, that since in purpose clauses A's and S's are both crossreferenced by ergative prefixes and O's are crossreferenced by absolutive morphemes, we have here another environment for split ergativity in Aguacatec. And it appears that this apparent nominative/accusative verb agreement in purpose clauses correlates with the fact that the syntactic rule of EQUI treats S's and A's alike as EQUI targets in purpose clauses. It is also interesting to note that, as was mentioned before, S's and O's (as well as indirect objects), but not A's, in the matrix clause are possible EQUI controllers. Thus, the rule of EQUI treats S's and O's alike in the matrix clause; and this seems to correlate with the ergative/absolutive verb agreement system in matrix clauses. There appears, therefore, to be a correlation between morphological ergativity or accusativity and the way in which the syntactic rule of EQUI treats A's, S's, and O's in these constructions. It is not the case, however, that morphological ergativity and accusativity always correlate with syntactic ergativity and accusativity in Aguacatec. For example, it would be possible for a purpose clause to be embedded in a matrix clause which was itself a time adverbial clause embedded within another clause. If the time adverbial clause were in the remote past tense, its verb, as we saw before in (8-11) would be nominalized and the verb agreement would be nominative/accusative. Nevertheless, the rule of EQUI would work in the same way as was seen in
(48-55). An example of this is

(56) ye t-opoon-e7n yaa\nTHE 3sErg-ARRIVE THERE-intrans.v.n MAN
\ntan wa-a7n comp EAT-intr.infin
'when the man arrived to eat,...'

Here it can be seen that in spite of the fact that the S of the matrix clause is crossreferenced by an ergative prefix (because of the nominative/accusative verb agreement in this environment) it still controls the EQUI deletion of the S of the purpose clause, just as it would if it showed the "normal" ergative/absolutive verb agreement (cf. 49) [7].

Another element found associated with verbs in Aguacatec is the directional. Directionals function much as do post verbal particles in English. There is a set of intransitive verbs of "motion" in Aguacatec, with each of which is associated a special directional particle. Some of these verbs and their associated particles are shown in Table 2. In the imperfective aspects (incompletive or potential) the directional element takes the form of one of the directional particles placed after the verb, as seen in (57) and (58).

(57) na \(\phi\)-s-tz'iib' ku7n
incompl 3sAbs-3sErg-WRITE DOWN
'he is writing it down'

(58) \(\phi\)-s-tz'iib'-e7 ku7n
3sAbs-3sErg-WRITE-pot DOWN
'he will write it down'

In the perfective aspect (that is, in the various past tenses), the
<table>
<thead>
<tr>
<th>intrasitive verb of motion</th>
<th>directional particle</th>
</tr>
</thead>
<tbody>
<tr>
<td>-b'een 'to go'</td>
<td>niin 'thither'</td>
</tr>
<tr>
<td>-saaj 'to come'</td>
<td>tzaaj 'hither'</td>
</tr>
<tr>
<td>-eel 'to leave'</td>
<td>leen 'leaving, &quot;out&quot;'</td>
</tr>
<tr>
<td>-ook 'to enter'</td>
<td>keen 'entering, &quot;in&quot;'</td>
</tr>
<tr>
<td>-kuu? 'to descend'</td>
<td>ku7n 'descending, &quot;down&quot;'</td>
</tr>
<tr>
<td>-kyaaq 'to stay'</td>
<td>kyeen 'staying'</td>
</tr>
<tr>
<td>-iky' (Chal.; Agc. dial.:-iik') 'to pass'</td>
<td>ky'een 'passing'</td>
</tr>
</tbody>
</table>

Table 2

directional element takes the form of one of the intrasitive verbs of motion showing absolutive agreement with the S, if the sentence is intrasitive, or with the O if the sentence is transitive. The main verb follows and, if transitive, takes the form of the active infinitive, which shows ergative agreement with the A:

(59) ja $^{*}$-kuu7 a-tz'ilb'-aal
prox.past 3sabs-DESCEND 2sErg-WRITE-act.inf
'you wrote it down'

(60) ja kxh-b'e'en w-uky'-aal
prox.past 2sAbs-GO 1sErg-CARRY-act.inf
'I carried you off'

(61) akxh-fee7 trimp-uj
2sAbs-RISE FALL-intr.suff.
'you fell (suddenly)'
(or maybe better: 'you up and fell')

140
The last example shows that directionals may have certain idiomatic uses. When the constructions shown in (59-61) occur in the environments in which nominalized verb forms are used, the directional element is nominalized and shows ergative agreement with the S or O:

(62) ye ə-kwe-e7n     a-tz'iib'-sal  
    THE 3sErg-DESCEND-intr.v.n 2sErg-WRITE-act.inf  
    'when you wrote it down...'

(63) ye a-b'een-e7n     w-uky'-aal  
    THE 2sErg-GO-intr.v.n 1sErg-CARRY-act.inf  
    'when I carried you off,...'

(64) ye a-j-e7n         trimp-uj  
    THE 2sErg-RISE-intr.v.n FALL-intr.suff  
    'when you fell...'

Note that in sentences like (62-64), A's, S's, and O's all show ergative agreement. With this in mind it would seem that Dixon's (1979:76-8, 97-8 n. 6) notion of an "extended ergative" system more accurately reflects what goes on in at least some of the types of subordinate clauses which are environments for split ergativity in Aguacatec than does the notion of "nominative/accusative" system.

Quotations are often followed by a special quotative verb chi: 'he says; it is said'. When the subject follows this verb, the j is usually dropped:

(65) Ø-chi        Lu7  
    3sAbs-SAY PEDRO  
    'said Pedro'

The subject shows absolutive agreement with this verb; however, if the absolutive prefix has a vowel, the stress uncharacteristically goes on
this vowel, and the i(ˌ) of the stem is dropped, as in

(66) che7-ch
    3pAbs-SAY
    'they said'

This verb is never marked for tense/aspect; however, if necessary, the
tense/aspect can be indicated by using the appropriate form of the verb

-ːb'an 'to do' [8]:

(67) che7-ch  ꭙ- ꭙ-b'an
    3pAbs-SAY 3sAbs-3sErg-DO
    'they said' (indefinite past)

(68) qa7-ch  ꭙ- ꭙ-b'an
    1pAbs-SAY future-3sAbs-3sErg-DO
    'we will say'

When chi(ˌ) is followed by the particle tzun 'so, then', they combine to
form stzun 'he said then'.

Scope of the Present Investigation

Having explained some of the basic terminology that I will be
using, I will now turn to the problem at hand. The present investiga-
tion was undertaken in order to see if certain syntactic and morphologi-
cal phenomena often found in Aguacatec texts, but not very often found
in sentence elicitation, could be explained from a functional point of
view. In particular, the following points were investigated:

1. The use of the particle -tz
2. The use of the genitive relational noun -eetz in contexts where it seems to have nothing to do with possession

3. The use of the "definite article" ye

4. The notions of "subjecthood" and "transitivity"

5. The use of certain subordinate clause verb forms in clauses that do not seem to be obviously subordinate.

The results of some of these investigations were more successful than others, as will be seen.

Most of these phenomena have already been discussed in the literature. For example, numbers 2 and 3 in the above list have been discussed by Harry S. McArthur at the third Mayan Workshop in Coban, Guatemala in July, 1978. McArthur, using data from AU, ATx, and JKgTx, describes the use of -eetz and ye roughly according to the chart shown in Table 3 [9]. Thus thematic participants (major participants which are "themes") are marked by -eetz. Themes are said to be either local (within a sentence or clause) or global (over a larger stretch of discourse). Minor participants, which are neither thematic nor active participants, are marked with ye. Major participants may also be marked with ye in occurrences where they have not yet come to dominate the action.

There seem to be a number of problems with this, however. For one thing, it is not clear exactly how the terms "theme" and "active participant" are being used. The term "theme", of course, has been used in a variety of ways by different authors. Halliday (1967), for example, uses the term to mean something like the point of departure of a sen-
<table>
<thead>
<tr>
<th>Thematic Participant</th>
<th>Active Participant</th>
<th>Non-Active Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erg-eetz + noun</td>
<td>Erg-eetz + noun</td>
<td></td>
</tr>
<tr>
<td>Subject (=agent S?)</td>
<td>Direct object</td>
<td>Direct object (= patient S or O)</td>
</tr>
<tr>
<td>of event clause</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-Thematic Participant</th>
<th>Noun</th>
<th>Ye + Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct object (= patient S or O), agent in passive construction</td>
<td>Direct object (= patient S or O), instrument, indirect object</td>
<td></td>
</tr>
</tbody>
</table>

Table 3

tence (which for Halliday conveniently turns out to always be the first element in a sentence, at least in English). McArthur does not seem to be using the term in this sense, however, since McArthur's themes do not always come at the beginning of a sentence nor does every sentence have one. Kuno (1975) defines theme as "what the sentence (or discourse) is about." One could imagine that many of McArthur's examples of themes fit this definition; however, in some cases, such as in example (89) below, it is not clear that the constituent marked with -eetz is either "the point of departure" or "what the clause is about". While Halliday seems to believe that every sentence has a theme, Kuno (1975:278) claims that there are, in fact, sentences without themes such as those which he calls "presentational sentences", e.g. 'Up jumped a rabbit.' Kuno
further claims (1975:326) that Halliday's notion of theme is completely independent of discourse functions and thus is a useless concept. Chafe (1976:28), on the other hand, believes that the term "theme", in any sense, is superfluous. Kuno's definition of theme seems to correspond to Chafe's definition of "subject". Also note that what Kuno calls "theme" is called by some writers "topic"; however, in this paper the term "topic" will be used to refer only to those phenomena which Chafe (1976:49-53) calls "topics, English style" (i.e., constituents which are fronted and contrastive), "topics, Chinese style" (i.e., the framework within which the main predication holds, a phenomenon found in many "topic-prominent" languages), and another kind of "topic" mentioned by Chafe, which seems to be something like a "premature subject", found in some other types of topic-prominent languages like Caddo.

As for whether or not the term "theme" is superfluous, it seems that there is a sense in which many kinds of discourses are "about" somebody or something. Of this there can be little doubt, and the person or thing that a discourse might be about might just as well be called "theme" as anything else. What really seems to be at issue, however, is whether or not such a notion of theme is ever overtly manifested in the grammar of any given language; and if it is overtly manifested, is it independent of all other grammatical categories found in the language. In many languages, such as English, this notion of theme seems to have been "grammaticalized" and is overtly manifested as "subject" as Chafe shows. In such languages, then, the term theme may indeed be superfluous since it seems to be pretty much subsumed under the notion of subject (though Kuno may disagree with this, since his
themes are not always subjects). I would like to keep an open mind about this, however, and not automatically assume that all languages have grammaticalized theme as subject. There may be languages which have not grammaticalized theme as subject; and among such languages there may be some in which the notion of theme is overtly manifested in some other way in the grammar [10], and there may be others in which the notion of theme is not overtly manifested in the grammar at all.

We might assume, then, that McArthur is using the term "theme" in some sense like "what the sentence (or discourse) is about" (since that definition seems to fit his use of the term a little better than does Halliday's) and that "active participant" means something akin to (potential?) agent (it is not clear how the "direct object", as seen in the lower left hand box of Table 3, fits in here). But even if this is so, there seem to be some even more serious problems with McArthur's interpretation. For example, in the appended text, the main character, Luw7 Tzu7, is introduced in line 4 and named in line 5. He seems to continue to be the (global?) theme in the assumed sense of the word through at least line 11. In these lines he appears as the grammatical subject, and also often as the semantic agent, of a number of transitive and intransitive clauses. Yet he is not once marked by -eetz until line 13. However, previously in line 13, the a'pyaa; have been marked by -eetz, thus apparently indicating that they are the theme. In fact, it does seem that the a'pyaa; which were first introduced in line 6, perhaps are the (local?) theme in line 12 and perhaps continue to be the theme through at least the first part of line 13. However, the a'pyaa; in line 12 are not marked by -eetz but are, in fact, marked by ye, which
according to McArthur indicates a non-active non-theme. And this in
spite of the fact that the a'pyaa' are the subject of the intransitive
time adverbial clause and of the transitive main clause, agent of those
same clauses, and furthermore "what those clauses are about". While it
is true that the a'pyaa', have not yet come to dominate the action in
line 12, thus perhaps making the use of ye appropriate, it nevertheless
does not seem to make much sense to say that the a'pyaa', marked by ye
in line 12, are not the theme of that line (but then what is?), and then
become the theme, marked by -eetz, in line 13 (suddenly dominating the
action?) only to be superseded by the reemergence of Lu7 Tzuu7 as the
theme, being marked by -eetz for the first time, near the end of that
same line. And in any case, to claim that this is in fact what is going
on here only because it would then agree with the scheme in Table 3
would be to argue in circles. What would be desirable would be an
independent and well motivated way of determining what things like
themes, active participants, etc. are and then see to what extent they
can be correlated with morphemes like -eetz and ye. I will not actually
attempt this at this point; however, impressionistically it seems that
while McArthur's treatment does perhaps account for at least some of the
data seen in the three texts that he referred to in his presentation, it
does not work in general.

Number 1 in the above list, the use of the particle tz, is dis-
cussed in L. McArthur (1979). According to this analysis (L. McArthur
1979:221),

Agua catec discourse in general organizes information according
to whether or not it is considered to be of primary interest. The
information of primary interest constitutes the MAINLINE of the
discourse, as opposed to information which is BACKGROUND, or supportive, in nature. In addition to the distinction between mainline and background information, there is a further category which consists of information which is HIGHLIGHTED by the addition of the -tz affix, and which intersects with the other types of information.

The disparity in frequency of occurrence of -tz in different texts can be attributed to the difference in orientation of varying discourses. A narrative which is event-oriented may have -tz marking almost every predicate, functioning as a type of progression marker highlighting succeeding events which form the BACKBONE of the narrative and which are to be built upon later. Other discourses, which are not primarily event-oriented, or are explanatory in nature, have relatively few occurrences of -tz.

Thus, in narrative texts, such as the one appended to this paper, the "events" of the narrative are said to constitute the "mainline" of the discourse. Mainline events are further divided into those which are of "primary interest" and those which are of "secondary interest". This difference between events of primary and secondary interest is generally signalled by the tense/aspect marking of the verb and is discussed in greater detail in H. McArthur (1979) [11]. Opposed to the mainline events is the "background" information. According to this analysis, tz may appear in a primary or secondary mainline event clause in order to "highlight" it, indicating that the event is one which "moves the story along" and which will be "built upon" later. These highlighted mainline events are said to form the "backbone" of the narrative. Tz also occasionally occurs in a clause giving background information in order to highlight it as "crucial background material". In other discourse genres, tz is not used as much as it is in narratives, and its function is somewhat different also; however, in all cases tz can be said to indicate a clause which is being "highlighted" for one reason or another.
While this analysis seems to work reasonably well for the data which McArthur presents, we might note a few problems with it. For example, consider lines 21-24 of the appended text. Pedro is trying to trick the merchants into believing that he has a magic pitcher that will boil water, roast meat, heat tortillas, etc. without having to build a fire. Unbeknownst to the merchants, however, Pedro has in fact built a fire under the ground on which he places his pitcher. In line 21 we are told that Pedro put his pitcher on the ground. This event seems to be one which advances the action (the stage is now set for Pedro to fool the merchants) and also seems to be one which is built upon later (this act allows the fire to heat the pitcher, thus enabling Pedro to perform the tricks described in lines 24ff). Thus, we should predict that this clause will contain the particle tz; however, inspection of line 21 shows that it does not. Line 22 seems to be background information; it does not advance the action; we already know that the fire is underneath, and this line is merely reminding us of that fact. Thus, we would predict, correctly in this case, that this clause would not contain the particle tz. The purpose of line 23 is debatable. It is not clear why Pedro piled a little sand around the pitcher; and thus, it is difficult to determine whether or not this clause advances the action or not. Certainly it is not something that is built upon later. It can be seen, however, that this clause does contain the particle tz. If this clause were intended to be background information, we would conclude, according to McArthur's theory, that tz is marking this clause as crucial background information. However, it seems clear that the information being given here is in no sense crucial. On the other hand, if it is the case that this clause is a mainline event, we would have to
conclude that tz is highlighting this clause as something that advances the action. Yet it is not clear why this clause would be advancing the action more than line 21, which we saw does not contain tz. Moving on to line 24, it seems that this line does advance the action (the fact that the pitcher appears to boil without being heated is part of Pedro's plot to trick the merchants), and yet it does not contain tz. Other problematical examples of this type can be found throughout the appended text. In the following section, an alternative analysis for the use of tz will be proposed, and then I will compare this alternative analysis with that of McArthur to see to what extent the two analyses can be reconciled.

It can be seen, then, that there are some problems with the analyses that have been proposed for some of these phenomena. We must therefore look into these and the other above mentioned matters somewhat more carefully in order to find a more adequate, and hopefully less circular, treatment. In what follows we shall consider each of the five phenomena mentioned above in turn.

The Particle Tz

In looking at texts in Aguacatec one often finds the enclitic particle tz appended to certain words, though this particle seldom, if ever, occurs in eliciting sentences from an informant. When examples of it are pointed out to them in texts, native speakers generally claim that the particle does not have to be there but otherwise seem to be unable to explain what difference it makes if it is there. If nothing else this seems to indicate at least that this particle is optional.
Though at first sight it seems to occur "scattered here and there" throughout a text, it is actually placed regularly according to a simple rule. Given that simple sentences consist of one or more of the following elements:

\[(69) (\text{PARTICLES})+\text{PREDICATE} (\text{+PARTICLES})+\{A\} S \\text{S} (\text{+PARTICLES}) (\text{+O}) (\text{+OTHER CONSTITUENTS})\]

the particle \(tz\), if it occurs at all, will be appended to the rightmost element of (69) excluding "OTHER CONSTITUENTS". The appearance of being scattered here and there is due to the fact that some of the elements of (69) are optional and some may not appear overtly even if present. For example, an \(A\), though obligatory in a transitive sentence, may be pronominalized, in which case it appears only as an agreement marker on the verb. All of this suggests that whatever the function of \(tz\) may be, it probably refers to the clause as a whole and not just to the constituent to which it is cliticized.

As a first approximation to the function of \(tz\), consider its use in the following excerpt from the appended text (lines 6-12) [12]:

(70a) b'een tilool Lu7 ye tetele7n tzaaj
HE-SAW-IT PEDRO THE ITS-LEAVING HITHER
chichoojo7n kob'ox ajpyaaj.
THEIR-PAY SOME MERCHANT
"Pedro saw some merchants(travelling salesmen) receiving their pay."

(70b) xe7te7n tzun-tz tan ixtumle7n
HIS-STARTING THEN-tz TO ITS-BEING-THOUGHT
juun tajtzaa7q1,
ONE HIS-IDEA
"So starting to have an idea,"
niin tzun b'e'en 117-tz tan k'otle7n
AND THEN HE-WENT HE-tz TO ITS-BEING-DUG
juun jul tzi b'ee7;
ONE HOLE AT-EDGE ROAD
'he went to dig a hole at the side of the road;'

niin kyaaq kyeen tq'ool q'aaq'-tz tk'u71 jul
AND HE-LEFT-IT FIRE-tz IN-IT HOLE
'and he left fire inside the hole;'

kyaaq kyeen tq'ool si17-tz;
HE-LEFT-IT FIREWOOD-tz
'he left firewood;'

niin tzun paqxi7-tz.
AND THEN HE-RETURNED-tz
'and then he returned.'

ej ma ye kyopoone7n ye
AND WHEN THE THEIR-ARRIVING-THERE THE
e7 ajpyaaq qa1e7 xmuqe7-t plural MERCHANT WHERE IT-WAS-BURIED-particle
ye q'aaq',
THE FIRE
'And when the merchants arrived where the
fire had been buried,'

niin tzun e7kuu7-tz
AND THEN THEY-DESCENDED-tz
'they sat down.'

Here it can be seen that the protagonist of the story, Lu7 Tzuu7 'Pedro
Tecomate', is the subject of each of the principal clauses in (70a-g)
(where "subject" is taken to mean 'either A or S'). Furthermore it can
be seen that tz appears in each of these clauses except the first one.
In (70g) the subject changes to 'the merchants' and tz does not appear.
In (70h), however, the merchants are still the subject and tz reappears.
Thus, it seems that tz functions as a "same subject" marker. This, in
fact, accounts for most of the occurrences of tz remembering, of course,
that native speakers judge its use to be optional and, thus, not every instance of same subject is marked.

There are, however, some occurrences of tz which cannot be characterized this way. For example, in the same text (line 43), after telling how the merchants went crazy over what they saw, we find [13]:

(71) tooke7n tzun chiyool ajpyaaj-tz
     ITS-ENTERING THEN THEIR-WORD MERCHANT-tz
te7:7
     ABOUT-IT
     'So the merchants, starting to discuss it,...'

Here, although we continue to talk about the merchants, tz cannot be interpreted as indicating same subject because ajpyaaj is not the subject but rather the possessor of the subject chiyool, which is here introduced for the first time. This suggests that perhaps tz does not mark same subject but rather same "theme", where theme is taken to be something like "the main participant that the discourse is about". As might be expected, the theme is usually the subject; however, there are some cases such as in (71) where this is not the case.

Another problematical example is the following (from N):

(72) ej niin tzun b'een naawloon tan
     AND THEN HE-WENT NAWLOON TO
     je7se7n tzaaj choklaat xe
     ITS-BEING-RAISED HITHER CHOCOLATE IN
     chikoo7k ye eegum kanteel
     THEIR-CRATES THE CARRIER CANDLE
     niin kuu7 q'aag'-tz ta7n
     AND IT-DESCENDED FIRE-tz BY-HIM
     'And so the "naawloon" went to take the chocolate from the crates of the candle carriers, and he made a fire.'
     (lit. '...fire descended by him.')

153
Here the subject of the last clause is q'aq' 'fire', and tz is appended to this. However, this is the first time that the fire has been mentioned; therefore, tz cannot be marking q'aq' as same subject. However, it seems clear that the subject of the first clause, naawloon, is the theme of that clause, and it seems likely that it continues to be the theme of the second clause in spite of the fact that it is not the subject but rather an oblique agent. Thus, once again tz can be taken to mean same theme. Another similar example is (73) from the beginning of AU.

(73a) Pwees i b'am chij, WELL Focus IT-DID-IT IT-IS-SAID at juun yaaj xa7q THERE-IS ONE MAN HE-WENT-OFF tan joyle7n kaab'ii1 wunaq txuk. Comp FIND-IT ITS-HONEY PERSON ANIMAL (=BEE) 'Well, it is said that what happened was that there was a man who went off to look for honey.'

(73b) Niin b'e'en tky'aal juun tzuu7 AND HE-TOOK-IT ONE TECOMATE t-etz ku71b'il ye kaab'. 3sErg-Genitive ITS-STORAGE THE HONEY 'And he took a tecomate gourd in which to keep the honey.'

(73c) Tpoone7n ii7-tz xo7l HIS-ARRIVING-THERE 3sPro-tz AMONG-IT xtzee7 WOODS 'Arriving in the woods'

(73d) nooj7n kyeen tzun juun ITS-FILLING STAYING THEN ONE wutzii7n kaab'-tz, BIG HONEY-tz '(and) then having found a large amount of honey,'
(73e) niin ook tan telse7n tzaaj,
AND HE-ENTERED Comp ITS-BEING-TAKEN-OUT HITHER
'he began to take it out,'

(73f) nooje7n t-etz tzuu7-tz tan kaab',
ITS-FILLING-UP 3sErg-Gen GOURD-tz BY HONEY
'the gourd filling up with honey,'

(73g) bween, paqxe7n tzun-tz xe ka7).
GOOD HIS-RETURNING THEN-tz UNDER HOUSE
'okay, and then going home.'

Here it seems clear that the man is the theme throughout this entire paragraph, and indeed every clause from (73c), which is the fourth clause in a row to have the man as subject, until the end of the paragraph contain tz except for (73e). What is interesting here is that in (73d) and (73f), even though these clauses contain tz, the man does not overtly appear at all although he must be understood to be the agent in both cases. Nevertheless, given that the man is the theme throughout the entire paragraph, it does not seem unreasonable to assume that tz in (73d) and (73f) indicates that the man continues to be the theme in these clauses in spite of the fact that he is not explicitly mentioned.

There is actually nothing strange about this considering that there are analogous situations in English. For instance, those clauses which have undergone "EQUI-NP Deletion" as well as some of those which have undergone "GAPPING" are "understood" as having the same subject as the previous clause in spite of the fact that no subject is explicitly mentioned.

We might consider another possible interpretation of these data, however. It might be that tz does indicate same subject but refers to deep subject rather than surface subject. This could perhaps account for all of the examples we have seen as long as we assume sufficiently
abstract deep structures and rules for deleting underlying subjects. Alternatively, we could take tz to mean same agent. However, example (74) should show that tz must be taken to mean same theme.

(74a) poro xhchiumutz niin ajpyaaaj ye
       BUT IN-FRONT-OF-THEM THITHER MERCHANT THE
       jee7n xtx'aanuu1 tiib' t-etz ch17b'aj
       ITS-ROASTING 3sErg-Gen MEAT
twutz ye xaruwu7;
IN-FRONT-OF-IT THE PITCHER
'But it was right in front of the merchants
that the meat roasted before the pitcher:'

(74b) xhchiumutz niin ye tooke7n xtxooli11
       HIS-BEGINNING-TO-LINE-IT-UP
       Lu7 ye 7waj te?; xaruwu7
       PEDRO THE TORTILLA ABOUT-IT PITCHER
       'it was right in front of them that Pedro
began to line up the tortillas around the
pitcher:'

(74c) xhchiumutz niin ajpyaaaj-tz ye jee7n
       MERCHANT-tz THE ITS-RISING
       sib'eel t-etz 7waj tan stz'e7e7n.
       ITS-SMOKE 3sErg-Gen TORTILLA BY-IT ITS-BURNING
       'it was right in front of the merchants that
smoke came up from the tortillas from their
burning.'

Here the tz in (74c) cannot be taken to mean same subject nor same agent since a'pyaa' is neither a subject nor an agent at any level but rather is the head of a locative noun phrase. Nevertheless, a'pyaa' does seem to be the theme throughout (74). Therefore, we can conclude that there is a notion of discourse theme which is relevant to Aguacatec syntax and which is distinct from, though closely related to, the notion of subject. Furthermore, we can conclude that tz is an optional marker of the fact that a clause has the same theme as the previous clause.
Even this characterization of tz is not completely correct. We should note four complications that must be dealt with in any complete treatment of the use of this particle:

(1) It is not always true that tz means 'same theme as the previous clause' since certain types of clauses are typically ignored in looking back at the "previous clause". Some instances of this can be found in the examples we have already looked at. One type of clause that is ignored is an object complement clause. Thus, when tz appears in (70b), it ignores the preceding object complement ye teele7n tzaa; chichoo7n kob'or a'pyaa in (70a) and instead refers back to the theme of the matrix clause b'een tilool Lu7. In the same way, purpose clauses such as tan k'otle7n 'wun 'ul tzi b'ee7 in (70c) are ignored. Examples of this can also be seen in (70b), (72), and (73). Though adverbial clauses are generally not ignored in this way, they are ignored if they are embedded within another subordinate clause. Thus, the adverbial clause (70g) contains another adverbial clause qale7 xmuqe7t ye q'asq embedded within it and is itself embedded within the matrix clause (70h). The tz in (70h) ignores the most deeply embedded clause in (70g) but not the less deeply embedded clause e; ma ye kyopoone7n ye e7 a'pyaa7. All of this suggests that tz may operate under some kind of "command" type constraint such as has been used in explaining pronominal anaphora in English; however, this needs further study.

(2) The second thing to note is that direct quotations are generally ignored in the same way as the clause types just discussed (see footnote [13]). Sometimes tz occurs in the quote formula formed with chi'. An example of this can be seen in line 14. Here the tz in the
quote formula *stzun *Lu7 niin-tz..., indicates that the theme of the
quote formula, *Lu7, is the same as the theme introduced at the end of
the line preceding the quoted material, that is, line 13. Note that the
theme of the quoted material in line 14 is not *Lu7. When tz appears in
the clause following a stretch of dialogue, as in line 108, it indicates
that the theme of that clause, in this case, the merchants, is the same
as the theme of the clause which preceded all of the quoted material
(including the quote formulas), in this case, line 104. Note that the
merchants as a group are not the theme of any of the intervening quoted
material nor in any of the quote formulas in lines 105–108. A similar
eexample has already been seen in lines 36–43.

(3) The third thing to note is that tz does not always refer back
to the theme of the previous non-ignored clause. In some cases a clause
with a particular theme may introduce a new theme near the end and then
the next clause may contain tz if its theme is the same as this newly
introduced theme. An example may be seen in:

(75a) niin kuu7 tq'ool-tz wi
AND HE-PUT-IT-DOWN-tz ON
txa7x ch'im tzwutz ye
GREEN GRASS IN-FRONT-OF-IT THE
xaaruu7 xchhiwutz e7
PITCHER IN-FRONT-OF-THEM plural
ajpyaaj.
MERCHANT
'And he put it on the grass before
the pitcher in front of the merchants.'

(75b) na chitze7een niin e7
THEY-WERE-LAUGHING THITHER plural
ajpyaaj-tz te7.j
MERCHANT-tz ABOUT-HIM
'The merchants were laughing at him.'
Here, tz in (75a) indicates that the theme of that clause, Lu7 Tzuu7, is the same as the theme of the previous clause. However, at the end of that clause, the a'pyaa: 'merchants' are reintroduced; and tz appears again in (75b) indicating that the a'pyaa: continue as the theme of that clause. Judy Aissen has suggested to me that it might make some sense to say that themes are not only something that clauses (or sentences?) may have but also something that adverbials may have. This would relate to the fact that (75b) contains the "same theme particle" tz even though the a'pyaa: were not the theme in (75a) but rather were reintroduced in the locative phrase xhchiwutz e7 a'pyaa:. According to this view, then, (75b) has the particle tz because it has the same theme as the preceding "theme containing constituent", namely, the locative adverbial phrase xhchiwutz e7 a'pyaa:. A similar situation was also seen in (74). This would then presumably also relate to the fact that adverbial clauses, as opposed to most other kinds of subordinate clauses, are not ignored in looking for the antecedent of tz. If there is some truth to all of this, it might not be just an idiosyncratic fact about Aguacatec. In Quechua, for example, there is a suffix -qa, which is usually called by quechuanists the "topic marker". Its function seems to be to mark contrastiveness on the NP which is the theme and also at times to indicate a shift to a new theme. However, one also sometimes finds adverbs which take the suffix -qa, thus suggesting that in Quechua too there is some kind of connection between themes and adverbials.

(4) The fourth thing to note is the above mentioned optionality of tz. It appears that tz tends to be used more often when the possibility of ambiguity may arise and less often otherwise. Thus, it tends to be
used less often in non-third person discourse than in third person discourse with two or more participants mentioned. Even in this latter situation, tz seems to be used less often when the third person participants are consistently mentioned by name than when they are consistently pronominalized.

A full treatment of these details is beyond the scope of this paper, however. For our purposes here, it is sufficient to observe that the use of the particle tz is based on some notion of theme. Then, looking at examples of the use of this particle we can note some interesting facts about the distribution of themes in Aguacatec. We have seen that in most cases themes are subjects, that is, A's and S's. We have also seen that there are some cases where the theme is not the subject but rather some other constituent such as possessor of the subject as in (71), oblique agent as in (72), and head of a locative phrase as in (74). We have also seen a case in (73) where the theme does not even overtly appear in the clause, though it is understood to be the agent. What has never been found, however, in any of the texts that I have looked at, is a case where an O was a theme. Therefore, given that A's and S's are included in the set of possible themes, and O's are excluded, it can be said that themes in Aguacatec are distributed on a nominative/accusative basis. This would seem to argue against a claim made by Plank (1979) that O's are grammaticalized topics (which is his term for theme) in ergative constructions. In fact in Aguacatec, O's are the one thing that can never be themes (or "topics" in Plank's terminology).
Having decided that the particle tz is an optional same theme marker, it would be well to examine certain potential problems which have been heretofor glossed over. For one thing, it may have been noticed that the tz in the phrase i tzun b'antz `so what happened was...' at the beginning of line 6 was conveniently ignored in the discussion of (70) above. Here the verb -b'en `to do' seems to be used in its impersonal sense (see footnote [8]), and thus we might imagine that the "same theme" referred to by tz is the story as a whole, which was talked about in lines 1-3. If this is true, however, we are stuck with explaining how this tz in line 6 can "jump over" lines 4 and 5 in order to refer to something last mentioned in line 3. Note that that lines 4 and 5 do not fit under any of our previously defined categories of clauses that are ignored in looking for the antecedent of tz. Another possibility is that -b'en is not impersonal here and that what i tzun b'antz actually means here is 'so what he (= Pedro) did was...'. If this is so, then the same theme referred to by tz is Lu7 Tzuu7, introduced in line 5. This would then correlate perfectly with what has been said above about the use of tz. Though this use of tz definitely needs further study, it seems likely that the theory advocated here, that tz is an optional same theme marker, can accomodate this case also.

Another use of tz is seen in the word i17tz. This word, which is probably derived from i17-tz (3sPro-tz) as seen in line 8, is used when the argument of an equational predicate undergoes a kind of "left dislocation". It is thus often translated as 'he/she/it is' although it is certainly not a copulative verb. An example from LTzp is seen in (76).
Here we might say that \( tz \) indicates that the immediately preceding NP is to be understood as the subject (= theme) of the sentence \( ii7 \) \( uuun \) \( paalee7 \) 'he is a priest' though this would appear to be an exception to the rule for the placement of \( tz \) as shown in (69). However, arguments of stative predicates ordinarily follow the predicate, and here we see that both \( ye \) \( yaa\) and \( ii7tz \) precede the predicate \( uuun \) \( paalee7 \). We might consider the possibility, then, that \( ii7tz \) in (76) is not only fronted but clefted. It was seen in examples (28-34) that clefted constituents are in a higher clause than the non-clefted part of the sentence. Thus, if \( ii7tz \) is in fact a clefted constituent, it would be the predicate of its clause; and thus, the \( tz \) is in fact placed according to (69) and indicates that the theme of the clause containing \( ii7tz \), namely \( ii7 \) (3sPro), is the same as the theme of the preceding clause, namely \( ye \) \( yaa\) [14]. Some evidence for this may be seen in line 66 of the text:

\[
(77) \text{ poro na } \ betael \text{ x-txuum } Lu7 \\
\text{ BUT incompl 3sAbs-LEAVE 3sErg-THOUGHT PEDRO} \\
t-eetz ye b'i7-tz na chi-txuum-uun \\
3sErg-Gen THE WHAT-tz incompl 3pAbs-THINK-antipass \\
e7 ajpyaaj t-e7; \\
\text{ plural MERCHANT 3sErg-SKIN} \\
\text{ 'but Pedro was imagining what the merchants were thinking of' (lit., 'but Pedro's thought was leaving of what the merchants were thinking about it.'
}
\]

Here \( tz \) appears on the question word \( b'i7 \) 'what' (note that, as shown in line 94, \( tz \) is not obligatory here). We saw in examples (38-40) that WH-question words are clefted; therefore, in the embedded question \( ye \)
b'i7tz na chitxuumun e7 a'pyss; te7; the question word b'i7-tz is a
clefted constituent and is the predicate of its clause. Thus, the
placement of tz conforms to (69) and indicates that the theme of that
clause, namely the questioned constituent, is the same as the theme
introduced at the end of the preceding clause, namely the antecedent of
b'i7tz, the possessor of the relational noun teetz.

In the previous section I briefly discussed L. McArthur's (1979)
theory of the use of tz as a marker of "highlighting" and showed that
there were some problems with it. In this section I have proposed an
alternative theory of the use of tz as a same theme marker. I also
claimed that tz was optional and that its use was to some extent deter-
mined by the potential ambiguity that there may be among the referents
of anaphoric devices. There is a strong possibility, however, that
these two theories are not incompatible. We might hypothesize that tz
is in fact a same theme marker as proposed here but that it is not in
fact optional but rather used only in "highlighted" clauses. That is,
tz marks the fact that a clause is both "highlighted" and continues the
same theme. Thus, clauses which seem to continue the same theme but
which do not have tz have the form they do because they are not
highlighted. Clauses which appear to "move the story along" and which
are "built upon later" but which do not contain tz have the form they do
because they do not continue the same theme. It must be remembered,
however, that in this paper we are only talking about a certain type of
narrative text. McArthur claims that tz works somewhat differently in
other genres. An examination of the examples that she presents from
these other genres seems to indicate that tz functions at least in part
as a same theme marker in these cases too. However, there are two cases where this does not seem to be true. One is an example from a "personal history narrative" (L. McArthur 1979:229). This text, however, appears to contain a number of non-third person participants, and it may be that the morphemes which crossreference these non-third person participants are enough to keep straight who the theme is. Thus, it may be that the function of tz as a same theme marker may be suspended in such cases so that it may be used strictly for highlighting. The other case is an example of "procedural discourse" (L. McArthur 1979:236). However, it might be expected that this type of discourse may not have a theme in sense in which I am using the term here, that is, a central participant which the sentences are "about". Thus, in this case too, it may be that the function of tz as a same theme marker has been suspended, and it is used strictly as a marker of highlighting [15]. All of this, of course, needs further study.

The Relational Noun -eetz

We will now turn to the examination of another morpheme, the relational noun -eetz. The basic meaning of this relational noun in Mayan languages generally is 'genitive'; that is, it indicates some kind of possessive relation between two NPs. In Aguacatec, however, this use seems to be limited to cases where the "possessive relationship" is of a highly "non-prototypical" nature, as in

\[
\text{\{78\} j\text{u}n \text{m}a\text{a}p \text{t-etz} \quad \text{tnum}
\]
\[
\text{ONE \ MAP \ 3sErg-Gen \ TOWN}
\]

\['a \ map \ of \ the \ town'\]
Here it can be seen that the "possessed" NP is followed by the relational noun, which always takes a possessive prefix coreferential with the "possessor" NP. If an overt possessor NP actually appears (as in 78), it immediately follows the relational noun, and the vowel of the relational noun is shortened. Usually the possessed NP does not take any kind of possessive prefix (as in 78); or if it does take such a prefix (as in 79), it does not refer to the "possessor" NP marked by -eetz. There is another use of this relational noun, however, which apparently is an innovation in the Hamean languages. Some examples of this use can be seen in

\[(80)\] xh-cheej (t-etz) yaaj 3sErg-HORSE 3sErg-Genitive MAN 'the man's horse'

\[(81)\] (q-etz) qa-cheej ~ qacheej (qetetz) 1pErg-Gen 1pErg-HORSE 'our horse'

Here it can be seen that the relational noun may sometimes appear before the possessed NP. Furthermore, it can be seen that the possessed NP does take a possessive prefix coreferential with the possessor NP. Also, as suggested by the use of parentheses in (80) and (81), -eetz is optional in these constructions. When -eetz is used it indicates contrastiveness in the same way as would adding heavy stress to the possessors in the English translations (see Chafe (1976) for a discussion of contrastiveness). Sometimes -eetz is used in Aguacatec, as well as in
other Mayan languages, without a head NP as a possessive pronoun as in:

(82) ye ka7l ya7stzun t-eetz
    THE HOUSE THAT-THEN 3sErg-Gen
    'the house is his'

(83) ej niin tzun jee7 tq'ool t-eetz wi txax7x ch'im
    AND THEN HE-SET-IT 3sErg-Gen ON GREEN GRASS
    'And he set his on the grass'

In (83), though not necessarily in (82), it seems that -eetz is also contrastive, indicating that of the possible things (in this case, pitchers) that we could be talking about, it was the one that belonged to him that he put on the grass. There are other uses of -eetz, however, that are not related to possession, as seen in

(84) ej jalchaan, niki' loo7 oor
    AND DAWN WHAT dubitative HOUR
    tiky'le7n t-eetz Lu7,
    HIS-HAVING-PASSED 3sErg-Gen PEDRO
    kyiky'e7n-t ky-etz
    THEIR-PASSING-particle 3pErg-Gen
    ajpyaaj
    MERCHANT
    'and at dawn (who knows at what
    hour Pedro moved on?) the merchants
    moved on'

Here tetz Lu7 cannot be understood as 'Pedro's' nor 'of Pedro'; similarly kyetz a'pyaa cannot mean 'the merchants' nor 'of the merchants'. It seems clear, however, that the use of -eetz in (84) has one thing in common with its use in (80), (81), and (83), namely, that it indicates contrastiveness. That is, what Pedro did is being contrasted with what the merchants did. Because of this, the English translation can be appropriately read with heavy stress on 'Pedro' and 'merchants' as indicated in (84). Actually, given that the verb forms tiky'le7n and
kyik'ey7n in (84) are nominalizations, one might argue that -eetz does have something to do with possession even here. However, the nominalized verb forms in (84) result from the clefting of adverbial phrases before verbs in the indefinite past tense. If the verbs had been in some other tense, they would not have been nominalized; however, -eetz could still have been used here. An example of this can be seen in line 72 of the text:

(85) sqeen chi-kuu? ky-etz ajpyaaj
    perfect 3pAbs-DESCEND 3pErg-Gen MERCHANT
    xe juun chin wi7 tzee7?
    UNDER ONE augmentative HEAD TREE
    'the merchants had sat down under a big tree'

Stress in English has other uses, of course, besides indicating contrastiveness. For example, in

(86) We got home late last night and found our front door open.

the heavy stress does not seem to indicate contrastiveness since the door is not being compared with anything else; but since we expect the front door to be closed, our surprise at discovering this unexpected situation is indicated by the heavy stress on door. In Aguacatec, -eetz is used in an analogous way, as seen in line 62 of the text:

(87) ma kyopoone7n, qopij taane7n
    BUT THEIR-ARRIVING-THERE OPENED ITS-FORM
    t-etz jul.
    3sErg-Gen HOLE
    'But when they arrived, the hole was open.'

In fact, it seems that -eetz can generally be translated as heavy stress in English. However, the converse is apparently not true. For example,
heavy stress can also indicate new information in English, as in

(88) Once upon a time there was a king. And
    he lived with his family in a large castle.
    (new) (given) (new)

However, -eetz is never used to introduce new information in Aguacatec unless it is specifically contrastive.

If we now examine the distribution of -eetz we find that it can be used with almost any NP. For example, we have seen -eetz used with noun possessors as in (80-81), and S's in (73f) and (84). It also occurs with oblique agents, as in

(89) ...juun sweerte7'c  ntx'amx
    ONE THIS-LUCKY-THING IT-WAS-CAUGHT
    w-eetz  wa7n
    1sErg-Gen BY-ME
    '(look at)...this lucky thing that I caught'

Examples of -eetz used with O's are extremely rare, but they do exist. One example was seen in (83), and two more can be seen in (90) and (91).

(90) niin b'aaq kyeen chitzooliiil ye ky-eetz
    AND THEY-RELATED-ALL-OF-IT THE 3pErg-Gen
    ky-ajb'il
    3pErg-DESIRE
    'and they related all of their wishes' 

(91) kun pilil ye w-eetz
    GO TRY THE 3pErg-Gen
    '(if you think your life is tough...)
    go and try mine!' (Shaw 1972:280)

Examples of -eetz used with A's are also fairly uncommon. One example is in the following, which is a common formula for beginning a story:
(92) at-$\theta$
  tuun yool
  EXIST-3sAbs ONE WORD $\theta$-w-ilt-naq w-eetz
  3sAbs-1sErg-HEAR-remote past 1sErg-Gen
  'there is a story that I have heard'

Here, however, it does not seem that -eetz is contrastive: it would not
be appropriate for me to read the English translation with heavy stress
on the pronoun 'I', singling myself out from the set of possible hearers
of stories. The weetz in (92) seems to be functioning as some kind of
non-contrastive pronoun in support of the ergative prefix on the verb.
It should be remembered that the ordinary independent pronouns in Agua-
catec are absolutes and are never crossreferenced by ergative pre-
fixes. Another example is seen in the following from AtTx:

(93a) niin tzun b'een b'eq'ooll,
   AND THEN HE-SWALLOWED-HIM
   'And then he (the animal) swallowed him
    (the patron),'

(93b) niin aaj junt tiirtz xe a7 chij.
   AND HE-WENT ANOTHER TIME-tz IN WATER quotative
   'and he went back into the water, they say.'

(93c) i  tzun b'eene7n tilool t-etz
   Focus THEN HIS-SEEING-IT 3sErg-Gen
   moos ye b'eene7n t-etz
   SERVANT THE HIS-GOING 3sErg-Gen
   patroon tan txuk,
   PATRON BY-HIM ANIMAL
   'when the servant saw that the patron
    had been taken off by the animal,'

(93d) niin tzun aajtz  lajqe7l...
   AND THEN HE-WENT-tz QUICKLY
   'he (the servant) went quickly...'

169
In (93c) *patroon* is the S of the intransitive verb *b'eene7n*, and *tetz* is apparently used with this NP because of the unexpectedness of his being carried off by the animal. This use of *-eetz*, then, is like that seen previously in (87). It can also be seen in (93c) that *tetz* is used with *moog*, which is the A of the transitive verb *b'eene7n tilool*. Once again, however, the use of *-eetz* with this A does not seem to be for indicating contrastiveness. The servant is not being singled out from the set of possible viewers of this event; in fact, it is clear that he is the only one who could have possibly seen it. What *-eetz* seems to be doing here is signalling a shift to a new theme. It is clear that the animal was the theme in the preceding lines, and it appears that the servant becomes the theme in (93c) as demonstrated by the fact that the same theme marker *tz* appears in (93d). Not all cases of theme shift are marked by *-eetz* in this way (cf. the theme shift in 75), but it does seem that the majority of instances of *-eetz* used with A's do indicate theme shift as in (93).

It seems unlikely, however, that examples like (92) can be explained in terms of theme shift. In a few "personal history" texts which I have briefly looked at, it appears that the first person singular genitive relational noun *weetx* is used very frequently in the same way as was seen in example (92); however, it seems that few of these instances of *weetx* can be construed as indicating either contrastiveness or theme shift. It is not clear why *-eetx* should be used so frequently in reference to the first person; however, such seemingly unwarranted emphasis given to first person participants is probably related to the notion of "empathy" (Kuno 1976, Kuno and Kaburaki 1977), which
apparently accounts for the kind of "egocentricity" in language use which seems to also be reflected in the notion of an "animacy hierarchy" with first person singular at the top. See DeLancey (1981:644-5) for further discussion of this point.

What seems clear, however, is that -eetz never appears to be used to indicate contrastiveness with A's. If an underlying A is to be made contrastive, it must be either clefted, as in (32), or made into an oblique agent of a passive and be preceded by -eetz as in (89). Alternatively, though, one might want to say that -eetz can indicate contrastiveness with A's but that it is a different kind of contrastiveness from that marked on S's and O's. With the O in (83), -eetz indicates that of the possible things (pitchers) that he could have set on the grass, it is the one that is his that is being talked about. With the S's in (84) -eetz indicates that of the possible people that could have passed on, it was Pedro (and afterwards the merchants) who did so. Thus, in both of these cases, -eetz is selecting out a member of the set of possible arguments of the verbs. In (93c), however, it could be said that -eetz is not selecting out a member of the set of possible "seekers" but rather is selecting out a new theme from the set of possible themes. From this point of view, then, it might make some sense to say that the theme shift marked by -eetz in (93c) is a "contrastive theme shift", whereas the theme shift in (75), which is not marked by -eetz, is a "non-contrastive theme shift".

It appears, then, that unlike thematization, which operates on a nominative/accusative basis, contrastiveness as marked by -eetz operates on an ergative/absolutive basis since S's and O's are included in the
set of things that can be made contrastive, while A's are excluded. It should be remembered, however, that even though O's may be contrastive, as seen in (90) and (91), there seems to be a great reluctance to do this in practice. Usually underlying O's which are to be contrastive are first put into derived S function as was seen in (93c) with tetz patroon.

We shall now take a very brief look at the introduction of new information. New information is often introduced by means of the "existential predicate" at 'there is/are; to be in a place' followed by a quantified noun phrase. The most common quantifier is the number ruum 'one', which also functions much like an indefinite article. An example of this construction was seen in (92). Arguments of at are always crossreferenced by means of absolutive suffixes. New information can be introduced in numerous other ways, however. For example, in (70a) kob'ox a'pyaa 'some merchants' is introduced as the possessor of an S. In (70c) ruum 'ul 'a hole' is introduced as the possessor/S of a "passive infinitive". In (70d) g'aaq 'fire' is introduced as an O. In (72) g'aaq 'fire' is introduced as an S. What is never found, however, is new information introduced as a surface A; ergative subjects of transitive verbs are always given NPs. Thus, since S's and O's may be new, but A's never may be, it can be seen that new information, like contrastiveness, but unlike thematization, is distributed on an ergative/absolutive basis.

It has been seen, then, that if we examine functional notions such as contrastiveness and new information, we find that ergativity is manifested in Aguacatec in other ways besides verb agreement and choice of
"pivot" (Dixon 1979) for certain syntactic rules. Furthermore, it appears that Plank's (1979) claim that ergativity can be accounted for in semantic and pragmatic terms is true, at least in Aguacatec. In this connection, it should be noted that one of the ergative syntactic rules we have seen is that which derives cleft sentences, and it is clear that clefting is another way of indicating contrastiveness. Furthermore, looking at the morphology of relative clause formation and WH-question formation, it appears that these two rules involve the same morphology as the clefting rule. In fact all three rules look like different versions of the same rule. This is not entirely unreasonable since cleft sentences, relative clauses, and WH-questions all deal with identifying a particular member from a set of possibilities; and this is essentially what contrastiveness is all about too. One could say, then, that relativization and WH-question formation bear a "family resemblance" (Rosch and Mervis 1975) to the kind of contrastiveness indicated by clefting; and, hence, the same syntactic rule (focus antipassive) is involved in these three constructions. Thus, it makes sense that these three rules should operate on an ergative/absolutive basis since, as we saw in the discussion of -etx, contrastiveness in general is restricted such that it can apply to S's and O's but never to A's.

It is also interesting to note that the reluctance to make O's contrastive with -etx seems to have spread to relative clause formation. We saw in (21) that O's may undergo relativization; and while (21) is a perfectly acceptable sentence, there seems to be a tendency not to use this construction but rather put the underlying O into "derived S function" before relativizing it, thus forming sentences like
(94) ja ʃ-w-il b'u7y (ye) n-ʃ-tx'-aj-x
                   prox.past-3sAbs-WASH-passive
t-an          xna7n
3sErg-Agent WOMAN
'I saw the rag that was washed by the woman'

As for the ergative verb agreement, we might wonder if this too
could have some kind of pragmatic basis. While it is not clear that
this is in fact the case, we might note the following facts. We have
seen that thematization in Aguacatec has a nominative/accusative basis,
but the verb agreement system does not seem to be based on this as it
apparently is in some other languages. The reason for this might be as
follows: Surface A's in Aguacatec seem to be subject to numerous res-
trictions, as we have seen. They may or may not be themes, and they may
be given, but they may never be new nor contrastive. O's may not be
themes, but they may be either given or new and either contrastive or
non-contrastive. Surface S's, unlike the others, seem to be under no
restrictions whatsoever. They may or may not be themes, may or may not
be contrastive, and may be either new or given. Thus, since O's have
fewer such restrictions on them than do A's, there may be a sense in
which they bear a greater "family resemblance", in the sense of Rosch
and Mervis (1975), to S's than do the more marked A's. And it may be at
least in part for this reason that S's and O's are categorized in the
same way in the verb agreement system while A's are treated differently,
thus motivating morphological ergativity [16]. It is also interesting
to note that in most treatments of ergativity it is said that ergative
NPs (i.e. A's) are "marked" whereas absolutes (S's and O's) are
unmarked (cf., e.g., Dixon 1979:71-9). However, it appears that in
Aguacatec, with respect to the functional parameters discussed here,
there is actually a "cline" of markedness with A's being highly marked (in the sense of being under a number of restrictions), O's being less marked, and S's being unmarked.

The Definite Article Ye

The particle ye in Aguacatec has a number of uses: it optionally introduces relative clauses, as in line 1 of the text (but cf. line 4); it often introduces complement clauses, as in line 6; it often introduces time adverbial clauses, as in line 13; and it is often used as the definite article. While all of these uses are undoubtedly related in some way (note that in the uses with subordinate clauses these clauses may be considered to be NPs), I will here consider only the use as definite article. Ye is said to be the definite article largely because native speakers translate it as such. However, an examination of the text will show that there are many NPs which seem to be definite (or at least are so translated) but which do not take ye (see Chafe 1976:38-43 for a discussion of definiteness). In fact, ye seems to be used relatively infrequently as a definite article. As noted above, H. McArthur considers ye to be the marker of a non-active non-thematic participant; however, we have seen that there are certain problems with this. We will examine here some of the pertinent facts about the use of this particle.

As might be expected, the use of the definite article ye has neither an ergative/absolutive basis like contrastiveness and new information nor a nominative/accusative basis like themes. It is found with A's, such as ye chitaac; 'their father' in line 56; it is found with S's,
such as ye e7 ḥpyaa: 'the merchants' and ye g'aaq 'the fire' in line 12; and it is found with 0's, such as ye ʔwa: 'the tortilla(s)' in line 30. It is also found with other types of constituents such as the object of a locative relational noun, ye xaaruuʔ 'the pitcher', in line 26. Nevertheless, there do seem to be a number of constraints on the use of ye. For one thing an NP does not seem to take ye if it is either an A or an S which is marked by -eetza. An 0, however, may sometimes take both ye and -eetza, as seen in (90) and (91). Another constraint on ye seems to be that it does not go on continued themes which could be marked by te (though it apparently can go on newly introduced themes; see line 12 of text). This seems to correlate, at least in part, with McArthur's claim that ye indicates a "non-thematic participant". Thus, if we consider ye to be the definite article, but subject to these two constraints, it can be seen why many apparently definite NPs do not take ye. However, there are still cases of other apparently definite NPs which do not take ye. For example, in line 29, xaaruuʔ 'pitcher' appears as the object of the locative relational noun tzwutz. This pitcher was first mentioned in line 17 and was also mentioned in lines 20-24, and 26. Thus, it would seem that in line 29 this pitcher should be "identifiable" (cf. Chafe 1976:39) and, therefore, should be definite. It can also be seen that this NP is not subject to any of the above mentioned constraints on the use of ye. Thus, this NP should take the definite article; and, in fact, it can be seen that xaaruuʔ is preceded by ye in line 29. However, in the following line, 30, the same NP, xaaruuʔ, which is the object of the relational noun teʔi, should take ye for all of the same reasons that it did in line 29; however, it can be seen that it does not, although the NP ʔwa: in that same line.
does take **ye** (presumably because "the tortillas" is a generic aggregate NP; see Chafe 1970:192). Whatever the reason for this may be, it is probably related to one further constraint on the use of the definite article: generally only one NP per clause may be marked with **ye**. There is an apparent counterexample in line 104; however, since **ye b’aa** "the first one" is in apposition to **ye k’ase7n** "the waking (one)", we may consider this to be like the repetition of the same NP and not a true counterexample. In fact, in all of the texts that I have looked at, there only seems to be one other counterexample, which is found in LTzP when the priest says:

(95) nik’na7 oor tz’-g’-uul ye yaaj
WHAT HOUR fut-3sAbs-ARRIVE HERE THE MAN
tu ye ko7k
WITH THE CAGE
‘when will the man come with the cage?’

This seems to indicate that the one-per-clause constraint is not entirely correct, although the fact that examples like (95) are so rare (one example in eight texts) surely must be significant. It is probably also significant, however, that in example (95) neither the man nor the cage seem to be the theme. If this sentence could be said to have a theme at all, it would probably have to be the ‘hour’. It does seem to be the case that when **ye** appears on an A, that NP is not the theme. Note the use of **ye** with the A **chitas:** ‘their (the merchants’) father’ in line 56 where the merchants, rather than the father, seem to be the theme. It would also seem to be the case that if **ye** appears on an S (as in 95), that NP is not the theme unless it is a newly introduced theme (as in line 12 of the text). But even if the constraint could be
revised along these lines, we would still be unable to predict which of several possible NPs would in fact receive the one occurrence of ye.

All of this seems to be vaguely reminiscent, however, of the situation in Tagalog as discussed by Schachter (1976, 1977). In Tagalog exactly one NP in a sentence may be selected as what Schachter calls the "topic" and marked by the particle ang. One characteristic of the "ang phrase" in Tagalog is that it is always definite, and the only way to indicate definiteness in Tagalog is with ang. While other NPs in a sentence might seem to be definite, only one may be selected to be marked with ang. It is possible that Aguacatec ye may be somewhat like Tagalog ang. They are quite similar in that there is only one per clause (although, as already noted, the one-per-clause constraint is not quite so rigid in Aguacatec as it is in Tagalog), and in that they both indicate definiteness. They are also similar in that neither of them indicates discourse theme. Schachter (1977:281-2) seems to indicate that in spite of the fact that the ang phrase in Tagalog is commonly called the topic, it does not indicate discourse theme in the sense that that term is being used here. Similarly, as was noted above, ye also does not ordinarily occur with themes in Aguacatec. One difference between Tagalog and Aguacatec, however, is that sentences in Aguacatec are apparently not required to have a "ye phrase" though most types of sentences in Tagalog do have to have an ang phrase.

Van Valin and Foley (1980:339), looking at Tagalog from the perspective of "role and reference grammar", describe ang as the marker of "pragmatic peak" where pragmatic peak is defined as the one nominal constituent of a clause that is "singled out for special morphosyntactic
treatment as the pragmatically most salient NP in the clause." (Van Valin and Foley 1980:338). They go on to say that

Pragmatic salience is established by two interacting factors, discourse prominence (i.e. definiteness, specificity and givenness), on the one hand, and...the speakers 'focus of interest', on the other, i.e. that participant which the speaker treats as the most salient in the situation under consideration. (Van Valin and Foley 1980:338)

A further characteristic of pragmatic peak in those languages which have them (referred to as "reference-dominated languages" by Van Valin and Foley) is that it is a "grammaticalized" constituent which may control rules of anaphora and deletion (Johanna Nichols, personal communication), such as the subject in English and the _ang_ phrase of Tagalog. Thus, in spite of the many similarities between Aguacatec _ye_ and Tagalog _ang_, _ye_ cannot be considered to be the marker of pragmatic peak in Aguacatec since the _ye_ phrase does not meet this latter criterion. According to the criteria mentioned here, it would appear that the absolutive NPs would have to be the pragmatic peak in Aguacatec (which appears to definitely be a "reference-dominated" rather than a "role dominated" language according to Van Valin an Foley's typology). Nevertheless, it does seem that _ye_ calls some kind of special attention to one of the definite and (usually) non-thematic NP's in the clause. Therefore, we might say that _ye_ marks a definite NP as being a "center of attention". That this is so is suggested, though not definitively proven, by fact that the examples of _ye_ seen in the text generally seem to appear in NPs which indicate one of the more important participants in a clause (along with the theme and any contrastive constituents). Thus, for example, in line 29 there are three participants mentioned: the merchants (theme),
the meat, and the pitcher. All three of these are definite, but the one non-thematic NP that the speaker chooses to single out as "center of attention" is the pitcher, ye xaaruul, which has played a fairly important role in this part of the story. In line 30, the speaker apparently chooses the tortillas, ye ʔwaʔ, as the "center of attention" rather than the pitcher or Pedro. The phrase teʔi xaaruul is apparently being treated here as an incidental locative expression. It seems to be the case that NPs of rather marginal importance, such as b'eeʔ 'the road' in the locative phrase of line 8, never get marked by ye even though they may be definite and even though there is no other NP in the clause that is so marked. This would presumably be because the speaker does not consider such NPs to be the center of attention. This idea, of course, needs further study and refinement. If there is any validity to it, however, we might then wonder if the use of ye with time adverbial clauses and some relative clauses might have to do with these clauses being "centers of attention" with respect to their matrix clauses.

Subjects and Transitivity

It has been seen that both the ergative/absolutive and the nominative/accusative distinctions are relevant in Aguacatec. The latter in particular is related to the notion of theme. One might wonder, however, just what is the significance of the notion "subject" in Aguacatec given that the notion does seem to be related to, though to some extent independent of, theme and the other notions we have looked at. I do not have a very satisfactory answer to this; however, it seems that whatever determines the selection of subjects in Aguacatec must be
somewhat different from what determines the selection of subjects in English. This can be seen from the fact that many clauses which are rendered best as transitives in English seem to be rendered in Aquacatec as intransitive clauses with patient subjects and optional oblique agent phrases. In general this phenomenon may result from the fact that absolutive NPs are less marked (i.e. under fewer restrictions) than are ergative NPs. However, this alone is insufficient to determine which clauses will be transitive and which intransitive. In any case it seems that whether an NP appears as a transitive subject or as an oblique agent of an intransitive verb is not determined by whether or not it is a theme. It should be remembered that, as has been seen, oblique agents as well as A's may be themes.

In some cases, determining whether a sentence should be transitive or intransitive may be based on lexical considerations. For example, in line 49 of the text one finds

(96) b'untz kyi1kt sii7
SO THAT THERE IS NO LONGER FIREWOOD
$-q$-eeq niin
3sAbs-1pErg-CARRY AND
kyi1kt qa-q'aag' $f$-kuu7
THERE IS NO LONGER 1pErg-FIRE 3sAbs-DESCEND
'...so that we will no longer carry firewood
and no longer make fire'

Here the first verb is transitive and the second verb is intransitive.
This may have to do with the fact that -eeq 'carry' is basically a transitive verb whereas -kuu7 'descend' is basically intransitive. And while there is a transitive causative stem -ku7saa; that can be derived from -kuu7, it happens that 'for fire to go down' is for some reason a common idiomatic way of saying 'to make a fire' (cf. ex. 72 above).
This will not, however, explain the use of the intransitive verb plus agent phrase in line 81:

(97) ḋ-kuu-q q-a7n t-Ḍ-k'u7l
     3sAbs-DESCEND-pot 1pErg-Agent prep-3sErg-BELLY
     ṣun saak
     ONE GUNNY SACK
     'we will stick him in a gunny sack' (lit. 'he will go down by us inside of a gunny sack')

But looking ahead to lines 91-93, it can be seen that the aʻpyes is not actually "stick" Pedro into a gunny sack but rather give him the sack and tell him to get in himself. Thus, the construction in (97) may be due to the fact that the aʻpyes are not, strictly speaking, agents here but rather something more like "indirect causes". Looking back to (96), we might wonder if the reason that 'to make fire' is expressed as 'for fire to descend (by someone)' might not be because humans are conceived of as the "indirect causers" of fires rather than their "makers".

This still will not explain, however, why in line 80 the merchants say that they will find a way to fool Pedro tan kyime7n 'in order to kill him' (lit., 'so that he will die'). In fact, it is fairly common in Aguacatec to say that 'someone dies by someone' rather than 'someone kills someone'. I suspect, though at present this is little more than a supposition, that the reason for this may be similar to the tendency seen in other languages to avoid admitting guilt. Thus, if a mother hears the sound of breaking glass coming from the kitchen and calls out, "What happened?", her young child is likely to respond with, "A glass broke" (not "I broke a glass"), or in Spanish, "Se me cayó un vaso" 'a glass fell from me' (not "Boté un vaso" 'I dropped a glass').

None of these explanations, though, will account for the relative clause in (89). Presumably the construction here is determined by the fact that the agent is contrastive (remember that As may not be contrastive). The problem with this, however, is that soon after this sentence in LTzP there is a very similar sentence where the oblique agent is not made contrastive with -eetz.

It should be noted in passing that DeLancey's (1981) attempt to account for voice changes (e.g. active vs. passive) in the world's languages in terms of "attention flow", "viewpoint", and "empathy" will not explain the Aguacatec constructions being discussed here. DeLancey claims that there is a "natural" order of "attention flow" from "source" to "goal" and that the order of NPs in a clause ordinarily reflects this natural order. The order of NPs in a clause is called "linguistic attention flow"; and, thus, linguistic attention flow is said to ordinarily follow natural attention flow. He also claims that events are most naturally viewed from the "viewpoint" of the participant with which the speaker has the most "empathy" (Kuno and Kaburaki 1977). Ordinarily the first NP in a clause is the source, starting point of natural attention flow, the starting point of linguistic attention flow, and the viewpoint NP. Sometimes, however, the viewpoint NP is not the starting point of natural attention flow. DeLancey claims that the function of passive in English and other languages is to change the order of the NPs such that the starting point of linguistic attention flow (i.e., the first NP) is not the starting point of natural attention flow but rather the viewpoint NP in the cases when the two do not coincide. It should be noted, however, that the order of NPs in a clause can often be
changed in Aguacatec without resorting to voice changes. Furthermore, when voice changes are resorted to in Aguacatec, it is not clear that the first NP is either the starting point of natural attention flow or the viewpoint NP. Thus, for example, we have seen that passive was used in (89) and the first NP to appear is "this lucky thing". Yet it would seem that the source of the action of "catching" would have to be the the first person singular agent. Furthermore, the speaker of this sentence is the first person singular agent himself; and one would expect him to have more empathy with himself than with the thing that he caught. Thus, it appears that the starting point of linguistic attention flow in this clause is neither the starting point of natural attention flow, nor the viewpoint NP. Although not all of the examples we have seen in this section involve passives, there is nevertheless a formal similarity between intransitive verbs with oblique agents and passives with oblique agents; and it appears that problems similar to those seen in example (89) can be found in each of the examples discussed in this section.

There may, however, be another explanation, which could perhaps account for all of the examples we have seen. Saksena (1980) argues that there is a semantic notion of "affectedness" which has consequences for the grammar of Hindi and other languages. This notion is also discussed by Ackerman (1981) and by Dahl and Karlsson (1976) (where it is referred to as "crucial change in the state of the object"). The general idea here is that in at least some languages the case marking of an argument sometimes depends on the extent to which the argument is "affected" by the action. Looking at the examples we have seen, it
would appear that one could make the claim for Aguacatec that patients which are very highly affected (i.e., those which undergo some major change of state) tend to appear as intransitive S's whereas patients which are less affected by the action tend to appear as transitive O's. Thus, for example, the fire in (96) (which undergoes creation), Pedro in (97) (who gets enclosed in a sack), Pedro in line 80 of the text (who gets killed) and the lucky thing in (89) (which gets captured) all undergo some kind of major change of state and all appear as S's of intransitive verbs. On the other hand, the firewood in (96) gets carried; but it does not really undergo any major change of state and, thus, appears as the O of a transitive verb. These examples, however, are merely suggestive. More evidence is needed to prove this claim. A possible problem with this idea is the fact that all of these intransitive verbs do have transitive forms. If the claim being made here is correct, why should such transitive forms even exist? One hypothesis that may explain this is that the transitive forms are used when one wants to draw attention away from the affectedness of the patient and concentrate more on the effect on the agent or experiencer. A suggestive piece of evidence for this hypothesis is found in the story JKyTx'. According to this story, at one time the coyote took care of the man's animals while the dogs lived in the woods. The dogs heard, however, that the coyote had eaten the man's animals. They decided to go report this to the king and suggest that they could do the job better than the coyote and, therefore, should be given a chance to prove themselves. The king then called the coyote and the man in order to check the dogs' story. Now it happens that the verb meaning "to eat something" in Aguacatec is an intransitive verb which may optionally take an oblique agent.
phrase. Thus, for example when the king asks the man if the coyote had in fact eaten his chickens, the man answers (among other things):

(98) ū-b’aaj w-aa7ktx t-aaq’un
prox.past 3sAbs-GET FINISHED OFF 1sErg-CHICKEN 3sErg-WORK
'my chickens were eaten by him'

Obviously food undergoes a fairly radical change of state on being eaten. Thus, it stands to reason that it would appear as the S of an intransitive verb rather than the O of a transitive. However, previously in this story, when the dogs are considering going to visit the king, they say to the man:

(99) ye e7 q-eetz ky1 1s
THE 1pPro 1pErg-Gen NEG incompl
ū-qa-baj-sa-aj ye e7 kne7r
3sAbs-1pErg-GET FINISHED OFF-caus-trans THE Pl COAT
tu e7 ky’1itx
WITH Pl CHICKEN
'As for us, we don't eat chickens and goats'

Here the verb appears as a transitive causative. However, it would appear that in this particular case, the dogs are less concerned with the effect of eating on the hypothetical chickens and goats and more concerned with the effect such an act would have on their reputations. Thus, this example suggests that the transitive form of such verbs is used to draw attention away from the patient and onto the agent as the affected participant.

Nominal Clauses

We have seen already that subordinate clauses in Aguacatec frequently have nominalized verb forms. There are clauses, however, which
use the same nominalized verb forms, but which are not obviously subordinate clauses. Thus, for example, in line 7 of the text the main predicate is the intransitive verbal noun ṣ-xeʔt-eʔn 'his starting' rather than the finite intransitive verb ṣ-xeʔt 'he started'. Yet it is not obvious that this sentence functions as a subordinate clause either in the sentence of line 6 or in the sentence of line 8. In most cases it seems possible to translate such "nominal clauses" (as I will call them) as gerundive clauses in English. Thus, line 7 has been translated as 'so starting to have an idea' rather than 'so he started to have an idea'. However, there are cases where one finds fairly long strings of these clauses (e.g. lines 83-86 and 96-100) where their treatment as gerundive clauses seems rather awkward, at least in translation. It should also be noted that these clauses are usually translated into Spanish by native speakers of Aguacatec as finite main clauses in the past tense rather than as gerundive clauses; though this could possibly be due to the fact that strings of gerundive clauses sound as awkward in Spanish as they do in English. Another thing to consider is that the nominalized verb forms used in these clauses are crossreferenced for subject by an ergative prefix, whereas the subject in English and Spanish gerundive clauses is deleted by EQUI [17]. For this reason native speakers may feel that nominal clauses in Aguacatec are more like finite clauses in Spanish than like gerundive clauses. This situation is not unique to Aguacatec, however, since there seem to be parallels in other languages. For example, in the North Caucasian languages Chechen and Batsbi, one finds strings of non-finite "conjoint clauses" together with finite "main clauses". Nichols (1981) relates these clause types to the notion of "grounding". Following this lead,
we might wonder if the use of nominal clauses in Aguacatec isn't related to grounding.

Grounding has to do with those parts of a discourse which are "foregrounded" as opposed to those which are "backgrounded". The foreground consists of the main points of a discourse and is apparently equivalent to what the McArthurs call "mainline". Thus, in a narrative text such as the one appended here, the foreground consists of the events of the narrative (see also Hopper and Thompson 1980:280-1). The background, on the other hand, is "that part of the discourse which does not immediately and crucially contribute to the speaker's goal, but which merely assists, amplifies, or comments on it" (Hopper and Thompson 1980:281). We might speculate that nominal clauses in Aguacatec represent either foregrounded or backgrounded material.

As a preliminary test of this hypothesis, I selected a number of lines from the text and tried to determine subjectively whether they represented foregrounding or backgrounding [18]. For a variety of reasons, this was not always an easy thing to do. For one thing, the distinction between foreground and background is not in all cases perfectly clear cut. For another thing, it was not clear what kinds of units I should be considering. I initially tried to do this by considering each individual clause as a separate unit and then determining whether the clause was foreground or background. This method turned out to be rather messy and did not give particularly revealing results. I next tried to consider whole sentences; however, in some cases this method seemed to group things together which intuitively seemed like they should be separate. It also begged the question of whether or not to
consider nominal clauses to be subordinate. As a compromise, the method I finally decided on was to consider nominal clauses and time adverbial clauses as separate units and otherwise to take whole sentences as units. I am not certain that I can adequately justify this method (I am not even entirely convinced that it is correct), but it did seem to be a fairly workable procedure. The results of this are shown in the first two columns on the left of Table 4 (In line numbers which contain a letter in Table 4, e.g. 12a, the number with "a" means the first clause of that line, the number with "b" means the second clause of that line, and so on; the line numbers of nominal clauses are circled). It can be seen in Table 4 that all of the nominal clauses are foreground, but there are also many foreground lines which are not nominal clauses.

In order to avoid one of the problems mentioned above, that of determining whether a particular line is foreground or background, it would be nice if there were some less subjective way of determining grounding. Hopper and Thompson (1980) argue that grounding is related to their notion of "Transitivity". Transitivity (with a capital "T", as opposed to the more traditional notion of transitivity with a lower case "t") has to do not only with the number of arguments that a verb may have but also with a number of other factors shown in Table 5. Thus, Transitivity is viewed as a continuum. A fairly "prototypical" transitive sentence, such as the farmer killed the duckling, should rate very high in Transitivity, that is, it should possess all of the properties shown in the "High Transitivity" column of Table 5. Any given sentence, however, may rate lower than this, anywhere down to the lowest Transitivity, such as in a sentence like would that sunsets not be so beauti-
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td></td>
<td></td>
<td>?</td>
<td>?</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>B</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>12a</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>12b</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>13a</td>
<td>B</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>13b</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>B</td>
<td>x</td>
<td>?</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>95a</td>
<td>B</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>95b</td>
<td>B</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>95c</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>?</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>?</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>?</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>100a</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>100b</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>F</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>B</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

Table 4

Hopper and Thompson (1980:283-4) then go on to claim that "the audience infers grounding not from a single morphosyntactic feature, but from a CLUSTER OF PROPERTIES, no single one of which is exclusively characteristic of foregrounding." They then go on to say that "this cluster of properties is precisely that which characterizes high Transitivity." Thus, assuming that this is correct, one could try to rate the sample lines from the text for Transitivity in order to see how well the results of this correlate with one's subjective judgements of foreground
and background. The results of this are shown in Table 4 where, for example, an "x" in column A means that this line was rated high with respect to property A in Table 5, and so on. Of course, some of the decisions to be made in rating these lines for Transitivity were also not always perfectly clear cut, and for this reason some of the squares in Table 4 contain a '?' rather than an "x" or a blank. The numbers in the column marked "T" are intended to be an overall Transitivity rating for each line, calculated by giving one point for each "x" and 0.5 point for each "?".

<table>
<thead>
<tr>
<th>A. Participants</th>
<th>High Transitivity</th>
<th>Low Transitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 or more</td>
<td>2 or more</td>
<td>one participant</td>
</tr>
<tr>
<td>participants,</td>
<td>participants,</td>
<td></td>
</tr>
<tr>
<td>A and O</td>
<td>A and O</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Kinesis</th>
<th>action</th>
<th>non-action</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>C. Aspect</th>
<th>telic (=action viewed from an endpoint)</th>
<th>atelic</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>D. Punctuality</th>
<th>punctual</th>
<th>non-punctual</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>E. Volitionality</th>
<th>action of agent volitional</th>
<th>action of agent non-volitional</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>F. Affirmation</th>
<th>affirmative</th>
<th>negative</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>G. Mode</th>
<th>realis</th>
<th>irrealis</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>H. Agency</th>
<th>agent high in potency</th>
<th>agent low in potency</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>I. Affectedness of O</th>
<th>0 totally affected</th>
<th>0 not affected</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>J. Individuation of O</th>
<th>0 highly individuated</th>
<th>0 non-individuated</th>
</tr>
</thead>
</table>

Table 5
for each "?". It can be seen that for the most part those lines which were subjectively judged to be background rated relatively low in Transitivity while those subjectively judged to be foreground rated relatively high. The only glaring discrepancies are in lines 3 and 102, which were judged to be background but received Transitivity scores of 6 and 8 respectively, and in line 7, which though judged to be foreground, only received a 4. Those lines which have an "x" in the column marked "Nom" are those clauses which are nominalized (including both nominal clauses as well as nominalized time adverbial clauses). The average Transitivity scores for the various types of lines are shown in Table 6.

Thus, on the average, there does seem to be a significant difference in Transitivity between foreground and background clauses, as seen in Table 6. It can also be seen in Table 6 that nominalized foreground clauses, though still relatively high in Transitivity, are on the average somewhat lower in Transitivity than are non-nominalized foreground

| Overall average | 6.05 |
| Average background | 3.45 |
| Average nominalized background | 4.00 |
| Average non-nominalized background | 3.31 |
| Average foreground | 7.23 |
| Average nominalized foreground | 7.06 |
| Average nominal foreground | 6.90 |
| Average non-nominalized foreground | 7.63 |

Table 6
clauses (note that the average Transitivity of nominal clauses is not much lower than that of nominalized clauses in general). This calls to mind H. McArthur's (1979) analysis of these clauses. According to McArthur, in "definite-event narrative", background is indicated by the use of any aspect other than the "§-aspect", which I have here called the "indefinite past tense". The mainline (=foreground) is divided into primary events, indicated by finite clauses in the indefinite past, and secondary events, indicated by nominalizations such as those found in what I have called "nominal clauses" and in nominalized time adverbial clauses [19]. Thus, it seems that McArthur's division of foreground into primary and secondary events, which was determined by morphological criteria, is reflected by the fact that, as shown in Table 6, the secondary events (i.e. the nominalized foreground) are slightly lower in Transitivity than the primary foreground. Of course, in order to demonstrate this conclusively, one would have to rate a larger sample of text material for Transitivity and then determine whether any differences found in the Transitivity of nominalized and non-nominalized foreground were statistically significant. Nevertheless, the preliminary figures shown in Table 6 suggest that there is something to McArthur's analysis [20].

It is also interesting to note that Nichols (1981) seems to have reached a similar conclusion in her analysis of Chechen. In analyzing texts in Chechen and Batsbi, Nichols found it necessary to distinguish two levels of grounding: "text-level" grounding and "episode-level" grounding. Text level background includes introduction, descriptive passages, scene setting, closing, and asides, while text level fore-
ground includes the episodes containing plot. On the episode level (= text level foreground), incidental information and restatements of known events are background while predications which advance the plot are foreground. It was found that in some Chechen texts, non-finite "con-junct clauses" were favored in text level backgrounding. However, in some other texts the conjunct clauses had relatively high Transitivity and were found in episode level background (which is part of the foreground relative to the text level) [21]. Thus, Aguacatec seems to be very much like Chechen and Batshi in that it distinguishes (at least) three types of grounding. And in particular, Aguacatec seems to be very much like those varieties of Chechen which use non-finite conjunct clauses for episode level background since Aguacatec uses non-finite (nominalized) clauses in foreground of lower than average Transitivity (or in McArthur's terms, mainline clauses of secondary interest).

One final thing to note about nominal clauses is that while they are fairly common in the appended text and also in the texts in L. McArthur (1973), they seem to be much less frequent in LTzP and N; and they do not occur at all in YQ'. While LTzP, N, and YQ' were transcribed from tape recordings, the texts in McArthur's book are said to have been written by native speakers. I am not certain of the origin of the appended text; however, I have been unable to find a tape recording of it, thus suggesting the possibility that it too may have been written down directly (so far, I have been unable to consult with Mendez on this). In studies which have compared spoken and written English (e.g. Kroll 1977, O'Donnell 1974) it has often been found that, among other things, written English (or more generally, "planned" English) often
contains more subordinate clauses than does spoken (or "unplanned") English. We might speculate, then, that the differences in frequency of nominal clauses in various Aguacatec texts is an example of the same phenomenon that has been reported for English.

Given that the use of nominal clauses in Aguacatec is related to grounding, we might further speculate as follows: In producing a "planned text", a speaker or writer has time to "stand back", as it were, and take a broad overview of the text as a whole (in writing, this could even be done after the first draft is produced), consider its structure, and make subtle decisions as to such things as grounding. Thus, in this type of text one would expect that all of the grammatical devices available for indicating grounding (such as nominal clauses in Aguacatec) could be exploited to their fullest. On the other hand, since spoken language is generally produced in sequences of short "spurts" (sometimes called "information units", "idea units", or "foci"; see Chafe 1979), the speaker does not have much of an opportunity to consider the overall structure of the text at the same time that it is being produced, and thus the speaker is less able to make global decisions as to grounding. This would be especially true in unplanned spoken language; and, therefore, one would expect to find that the grammatical devices for indicating grounding would not be fully utilized in this type of text.

Conclusion

We have looked at a number of morphological and syntactic phenomena in Aguacatec which, from the point of view of "sentence grammar", seem
to be somewhat mysterious. It has been seen, however, that by looking at these phenomena from a functional perspective and considering their use in the context of a discourse, some sense could be made out of these phenomena; and in many cases further interesting lines of research were opened up. It was also seen that ergativity has an important role to play in the organization of Aguacatec discourse. In conclusion, then, it could be said that the results reported in this paper represent a first step towards a functional grammar of Aguacatec. Tline7s.

Acknowledgements

I would like to thank Wallace Chafe and Johanna Nichols for taking the time to discuss with me many of the issues presented here. Without their help this paper would have fared much the worse. The usual disclaimers apply, of course. I would also like to thank Farrell Ackerman, Judith Aissen, Jack DuBois, and Catherine O' Connor for their helpful comments and suggestions.

Footnotes

[1] This paper represents the second substantial revision of a paper written in March, 1980, entitled "A preliminary look at Aguacatec narrative discourse". A portion of the first revision was presented on February 16, 1981, at the seventh annual meeting of the Berkeley Linguistics Society. That portion appears in BLS 7 under the title "Functional correlates of ergativity in Aguacatec" (Larsen 1981). Some of the ideas contained in this paper were also
presented at the West Coast Mayan Symposium at U. C. Santa Barbara in April, 1981.


[3] My use of these convenient labels, however, does not carry with it any commitment to Dixon's claim that A, S, and O are universal semantic-syntactic primitives.

[4] Roughly, an "ergative NP" is one that can trigger ergative agreement with a verb. Similarly, an "absolutive NP" is one that can trigger absolutive agreement with a verb. (But see footnote [7]).

[5] This sentence is not actually attested in my data, but has been included as a parallel to (55) below, which is attested. It has been modelled on actually attested sentences such as the following from Jky'Tx', p. 32:

(1) poro chin tajwe7n ku7n tz-$\phi$-in-chaq
    BUT VERY NECESSARY DOWN future-3sAbs-1sErg-ORDER
    ye yaaaj tan w-iit-zal...
    THE MAN Comp 1sErg-HEAR-act.infin
    'but it is neccessary that I send for the man so that I can investigate it'

See also (52).

[6] Note that it should be possible to say

(1) qajoye7 puuntii1 tan xub'se7n;

however, this would mean 'we will find a way for him to be tricked (by someone)'. It does not mean the same as (54): 'we will find a way (for us) to trick him' and, therefore, is not derived from (54) by EQUI.

197
[7] Thus, it is not clear whether the S of the verb -opoonen should be considered, following the rule in footnote [4], to be an ergative NP because of the verb agreement or an absolutive NP because of its syntactic behavior.

[8] -b'án is a transitive verb root; however, it is unique in that it has certain impersonal uses, such as the one shown here. See also its use in line 6 of the appended text.

[9] This chart has been adapted from that shown in McArthur’s handout in Spanish. I have added in parentheses some suggestions for changes in his terminology in order to avoid confusion with the terminology used in this paper (for example, McArthur seems to refer to patient S’s as "direct objects"). Hopefully, by making these suggestions I have not misrepresented his position.

[10] If the analysis to be presented in this paper is correct, it would appear that Aguacatec is one such language.

[11] I will not discuss H. McArthur (1979) in detail in this paper; however, I will on occasion refer to those parts of it which relate to points to be discussed here (in particular, number 5 in the list of topics). McArthur’s main point is that different genres of discourse in Aguacatec can be distinguished by the tense/aspect categories that characteristically occur in the mainline clauses, and that background clauses are distinguished from mainline clauses also by the use of other tense/aspect categories within each genre. His arguments seem to be basically sound although I find that I have certain disagreements with him as to just what the various tense/aspect categories are and how they are marked. Note that
what the McArthurs call "mainline" information corresponds more or less to what Hopper and Thompson (1980:280) call "foreground" (as opposed to "background"). The phenomena of foregrounding and backgrounding together are referred to by Hopper and Thompson as "grounding".

[12] The first part of line 6 of the text, *1 tzun b'antz*, has been ignored here for purposes of exposition. The significance of this phrase will be taken up later.

[13] Actually, it can be seen by looking at the text that there is a stretch of dialogue (lines 37-42) between line 36, where the merchants went crazy over it, and line 43, which is being discussed here. It will be seen presently, however, that such direct quotations must be ignored in looking back for the referent of *tz*. Thus, in this case, in order to find the referent of *tz* in line 43, we have to look back to line 36.

[14] Note that *17tz* can also refer to something said by another speaker, in which case it may be translated as 'yes, that is correct, it is him/her, he/she is the one, etc.' Some examples of this are seen in lines 77 and 79 of the text.

[15] The idea that a morpheme may "suspend" some of its functions in certain environments is not necessarily just an ad hoc explanation invoked to save this analysis of Aguacatec *tz*. Another example is the Lakhota conjunction *na* 'and'. This conjunction also indicates that the following clause has the same subject as the preceding clause if there is no full NP subject in the following clause and the verb agreement in the following clause is compatible with that
in the preceding clause. Otherwise, this conjunction is interpreted simply as 'and' with no implication of "same subject" (see Van Valin to appear for details). A further example can be seen in the Northern Pomo "different subject" marker, which when used on certain classes of verbs may still be used even if the following clause has the same subject provided that the subject of the second clause has a different semantic "role" from that of the first clause (see O'Connor 1981).

[16] Farrell Ackerman has pointed out to me that, since all Mayan languages have split-ergative verb agreement, one can not legitimately make the claim being made here unless it can be shown that other Mayan languages distribute themes, new information, and contrastiveness in the same way as Aguacatec such that these distributions can be reconstructed for Proto-Mayan. Otherwise, one would have to claim that each Mayan language developed ergative verb agreement independently for different reasons, which is not likely. My impressions of the few other Mayan languages that I know something about suggest that it is likely that these functional parameters are distributed as in Aguacatec. However, further investigation is obviously required to resolve this issue. (In this respect, now see DuBois 1981 and England 1981).

[17] Given that gerundive clauses in English and Spanish are "subjectless" and can therefore only be fully interpreted after finding the subject of the main clause, it would seem that long strings of such clauses would be more difficult to process than long strings of clauses which have subjects. This, then, may very well explain why
long strings of nominal clauses, which are marked for subject, can occur in Aguacatec while their translation as long strings of gerundive clauses in English or Spanish sounds awkward. In spite of the fact that the subjects of Aguacatec nominal clauses are not deleted by a rule of EQUI-NP DELETION, there often does seem to be an "equi-subject" requirement on strings of such clauses: note that the clauses in lines 83-86 and also in lines 96-100 all share the same subject. Thus, this equi-subject requirement on Aguacatec nominal clauses may be functionally equivalent to the use of EQUI-NP DELETION in gerundive clauses in English and Spanish.

[18] Lines 1-13 were chosen more or less arbitrarily although the decision to stop at 13 was not made arbitrarily; line 14 contains a direct quote, which I did not want to have to deal with. Lines 94-102 were chosen specifically because they contain a number of nominal clauses.

[19] Note that this applies only to the genre that McArthur calls "definite-event narratives".

[20] Note that Table 6 also shows a difference in Transitivity between nominalized and non-nominalized background clauses; however, since there were only two nominalized background clauses in the sample, it is not clear whether this is significant or not.

[21] Jones and Jones (1979) also argue that there are more than just two levels of grounding. They claim that all of the languages discussed in Jones (1979) exhibit a minimum of three levels of "information", which they call "background", "events", and "peak". In addition to these three basic levels, however, some languages
divide the background into "ordinary" and "significant" background
and/or they may divide events into "ordinary" and "significant"
events. Thus, some of the languages that they discuss may have up
to five levels of grounding according to their system.

References

ACKERMAN, FARRELL. 1981. Objects of possible interest. Special field
exam. Department of Linguistics. University of California, Berke-
ley.

ANDREWS, J. RICHARD. 1975. Introduction to classical Nahuatl. Austin:
University of Texas Press.

CHAFFE, WALLACE L. 1970. Meaning and the structure of language. Chicago:
The University of Chicago Press.

____.1976. Givenness, contrastiveness, definiteness, subjects, topics,
and point of view. Li 1976:25-55.

____. 1979. The flow of thought and the flow of language. Discourse
and syntax, ed. by Talmy Givon, 159-82. (Syntax and Semantics, vol.

DAHL, ÖSTEN and FRED KARLSSON. 1976. Verbal aspects and object marking:
a comparison between Finnish and Russian. International review of
slavic linguistics 1.1-29.

DELANCEY, SCOTT. 1981. An interpretation of split ergativity and related
patterns. Lg 57.626-657.


203


Text

The following text was written down and translated into Spanish by Gaspar Méndez López, an Aguacatec speaking student at the Proyecto Linguístico Francisco Marroquín. The English translation given here was made by me, comparing both Méndez's Spanish translation and the Aguacatec original. A certain amount of interpretation on my part has gone into this English translation (e.g. in the case of the "nominal
The Aguacatec version has not been tampered with other than to correct spelling errors. The punctuation in the Aguacatec version is Méndez's. The line divisions and numbering are mine.

**Juun Xtxoolb'il ye Xhe7n B'anaq Lu7 Tzuu7 Tyeemp Tzaaj**  
A Story about How Pedro Tecomate Did Long Ago

1. **At juun yool ye na chiyol wunaq jaalu7.**  
   There is a tale that the people tell now.

2. **mi b'intziij nqo kyi7 pe7,**  
   If it's true or not? (I don't know).

3. **e7 tu7 wunaq na yoolo teetz jaalu7.**  
   It's just that the people tell about it now.

4. **At juun yaaj na xoon wi muunt teen tzaaj**  
   There was a man who used to walk the earth long ago.

5. **Lu7 Tzuu7 b'ii7 le qayool,**  
   Pedro Tecomate was his name in our language.

6. **i tzun b'antz niin b'een tilool Lu7 ye teele7n tzaaj chichoofo7n kob'ox ajpyaaj,**  
   And it happened that Pedro saw some merchants (travelling salesmen) receiving their pay.

7. **xe7te7n tzuntz tan xtxumle7n juun tajtzaa7q1,**  
   So starting to have an idea,
8. niin tzun b'een ii7tz tan k'otle7 n juun jul tzi b'ee7
he went to dig a hole at the side of the road,

9. niin kyaa7 kyeen tq'ool q'aaq'tz tk'u7l jul
and he left a fire inside the hole,

10. kyaa7 kyeen tq'ool sii7 tz
left firewood,

11. niin tzun paqxijtz,
and then returned.

12. ej ma ye kyopoone7n ye e7 ajpyaa7j qale7 xmuqe7t ye q'aaq' niin tzun
e7kuu7 tz,
And when the merchants came to where the fire had been buried, they
sat down.

13. che7n ate7 kyetz ajpyaa7j tan wasa7n xe juun tzee7 ye tuule7n tzaaj
tetz Lu7,
The merchants were there eating under a tree when Pedro arrived.

14. cho7n pe7 xwit woq tzne7; stzun Lu7 niintz tzkyetz e7 ajpyaa7j,
"Are you going to sleep here?" said Pedro to the merchants.

15. cha7tz b'in che7ch ajpyaa7j.
"Here," said the merchants.

16. Aaj chinwitoq b'in nook tzixlaj,
"Ah, then I will sleep a little beside you."
17. ej niin b'eentz tan tiku'le7n tetz tk'aa7 tk'u7tj juun xaaruu7 tx'otx'
   And he went to get his water in a clay pitcher.
18. ma kyetz ajpyaaq na chi'iich' ye kyeetz,
   But the merchants, since theirs was of metal,
19. chim b'aalaj niin kyetz chixaaaruu7 ye e7 ajpyaaq
   the merchants' pitcher was very nice.
20. ma tetz Lu7 Tzuu7 na tx'otx' tu7,
   But Pedro Tecomate's was just of clay.
21. ej niin tzun jee7 tq'ool teetz wi tx'a7x ch'im,
   And he put his on the grass,
22. poro che7n at kyeen q'aaq' tzaq',
   but the fire was underneith.
23. ej niin jetzaaj juutz'iil mu7xh taal puqlaajtz
   And he piled up a little sand around it,
24. poro le raat niin pultz7n tetz xaaruu7
   but soon the pitcher boiled.
25. ej niin b'een tetz Lu7tz tan quxle7n tzaaj alaj chin xhchib'
   And Pedro went to cut off a piece of meat,
26. niin kuu7 tq'ooltz wi tx'a7x ch'im tzwutz ye xaaaruu7 xhchivutz e7 ajpyaaq

208
and he put it on the grass before the pitcher in front of the merchants.

27. na chitze7een niin e7 ajpyaaajtz te7:

The merchants were laughing at him.

28. mb'iit sb'ne7 nokx yaab',

"What's that crazy guy going to do?"

29. poro xhchiwutz niin ajpyaaj ye jee7n xtx'aanuul tliib' tetz chi7b'aj tzwutz ye xaaruu7

But it was right in front of the merchants that the meat roasted before the pitcher.

30. xhchiwutz niin ye tooke7n xtxoolii7 Lu7 ye 7wa7 te7: xaaruu7

It was right in front of them that Pedro began to line up his tortillas along the pitcher.

31. xhchiwutz niin ajpyaaajtz ye jee7n sib'eel tetz 7wa7 tan stz'e7e7n

It was right in front of the merchants that smoke came up from the tortillas from their burning.

32. poro kyi7k q'aaq' na kyil niin

But there was no fire that they could see.

33. poro cho7n at kyeen q'aaq' taq'un Lu7 'aq' tx'otx'

But Pedro had left the fire under the ground.

34. ej stz'e7e7n tetz 7wa7

And (with) the tortillas burning,
35. jukkaane7n tetz chi7b'aj tan tz'e7e7n atite7t
    and the meat shrivelling up from burning there,
36. qale7n kyeele7n yaab' e7 ajpyaaj te7j
    the merchants went crazy over it.
37. taa7 stzun juun b'an,
    "Papa," said one,
38. qaloq'e7 7uun qaxaaruu7atz.
    "Let's buy ourselves that pitcher.
39. kunu7 qa7j qale7n niin iil na ook q'as7q1 qeetz qasi17
    Just think how hard it is for our firewood to burn.
40. tz'uul ak'
    It'll come wet,
41. tz'uul tza7jo7t tan we7j
    we'll nearly get finished off from hunger,
42. go7q wi7 niin ab'aal
    or it'll be raining."
43. tooke7n tzun chiyool ajpyaajtz te7j
    So the merchants discussed it:
44. I lo7q sk'aay sqeetz
    "Would that he sell it too us."
45. poro ya7tz niin b'in tajb'il tetz Lu7 Tzuu7 tan k'asyyi1l ye xaaru7
    na pwoq na tsaaj Lu7 skye7j ajpyaaj
But that's what Pedro Tecomate wanted: to sell the pitcher, because
money was what Pedro wanted from the merchants.

46. ej ma chiwì7t waane7n niin tzun ook chitaaj e7 ajpyaaj tan jaqlë7n
xaaruu7.

And when they finished eating, the father of the merchants began to
ask for the pitcher.

47. k'aa7e7u7 ye xaaruu7u7 sqeetz

"Sell us your pitcher.

48. jatna7 miil na taaju7 sqaq'e7 te7j

However many thousand you want we will give for it."

49. aaj loq'e7 niinu7 taq7 b'untz kyì7kt si77 qeeq niin kyì7kt qaq'aaq'
kuu7

"Ah, buy it, Papa, so that we will no longer have to carry firewood
nor make a fire.

50. jalt kuntu7 lb'een quky'aj a7 niin1 qatuk'b'aa7 ku7n niin qatxool
keen qawaa7tz te7j xaaruu7 niin qaq' kyeen qachib' tzwutz b'untz
jaal ku7n lqiky'tz.

We'll just bring water and set it down and put our tortillas around
the pitcher and also put our meat in front of it so that we can
pass on quickly.

51. at pe7 sqeetz tuuch' ab'aal mpe naq tzaan ab'aal.

What does the rain matter to us is it's raining."
52. tooke7n tzun traat te7tj xaaruu7

So making a deal for the pitcher,

53. jat loo7 miil b'ân tetz xaaruu7

who knows how much the pitcher cost?

54. ej niin tzun kyaaj kyeen jatxij te juun k'oloj pwoqa7s tq'ab' Lu7 Tzuu7

So a pile of that money was left in the hand of Pedro Tecomate.

55. ej jalchaan, niky' loo7 oor tiki7le7n tetz Lu7, kyiky'e7nt kyetz ajpyaaj

and at dawn (who knows at what hour Pedro moved on?) the merchants moved on.

56. ma topoone7n oor tan tz'amle7n b'eene7n xhohaqool ye chitaaj juun tan tiki7le7n a7 tul xaaruu7

When lunch time came, their father sent one of them to bring water in the pitcher;

57. niin kuu7 kyiq'ooltz wi txas7x ch'im chi tuleej Lu7 teetz

and they put it down on the grass like Pedro had done with it.

58. na7 na pult ku7nt tooke7n chitxoolil1 chiwaa7tz juun tzi7i7 niin.

Around where it was going to boil they lined up their tortillas.

59. niiky'aane7n tetz xaaruu7 tan puqlaaj

(And with) the pitcher filling up with sand,

60. kyi7 niin na pult te7tz lo7q niin,
that thing didn't boil at all.

61. e\j chipaqxe7n tzun juun tiirtz tan tilwe7n ye xhe7n ku7n taa7n taane7n
    So they went back again to see how it had been.

62. ma kyopoone7n qopi7 taane7n tetz juj
    But when they got there, the hole was open.

63. che7n at kyeen q'aaq' tzaq' te7tz
    The fire was down inside that.

64. e\j i tzun kyaaltz aaj poro kyi niin tz'eel tzwutz
    And so what they said was, "Ah, but he won't get away from us.

65. alo niin inooj keent
    Some day we will find him."

66. poro na eel xtxuum Lu7 teetz ye b'i7tz na chitxuumuum e7 ajpyaa\j
    But Pedro imagined what the merchants were thinking of because he
    te7j na nachool b'in i17
    was intelligent.

67. na tb'iit niin Lu7
    Pedro heard everything.

68. ma teele7n juun k'oloj tyeemp nooje7nt keen Lu7 chiq'ab' ye e7
    After much time had passed, Pedro found himself again in the hands
    ajpyaa\ja7tz
    of the merchants.
69. kwee7n tzun junt tajtzaa7q1 Lu7 Tzuu7 tk'u71,
    So Pedro Tecomate, having another idea:
70. qale7 7nchaje7 wiib' xhchiwut7 stzun Lu7 b'antz
    "I will even show myself to them," said Pedro,
71. kwee7n tzun chikuuluul kyib'
    thus finding each other.
72. sqeen chikuul7 kyetz ajpyaa7 xe juun chin wi7 tzee7 stzl17 juun
    The merchants had already sat down under a big tree on the bank of
    chumaam a7,
    a large river.
73. i te7n niin yool Lu7a7's taaltz b'ajx tiir
    Pedro said the same words that he had said the first time:
74. che7n pe7 chiwitu7 stzun Lu7.
    "Are you going to sleep here?" said Pedro.
75. cha7tz b'in che7ch ajpyaa7
    "Here," said the merchants.
76. aaj taa7 ku7n ya7stzun juun taa7q1 ye k'aa7yiin ye xaaruu7 sqeetz.
    "Ah, Papa, that's the guy who sold us the pitcher.
77. kun i7tz, stzun juuna7tz
    Look! It's him," said that one.
78. nq'eera7tz stzun junt b'an tzaaj.

214
"No it isn't," said another.

79. kyiz 17tiz chi junt.

"Yes it is," said another.

80. aaj qa'e7 qa'oye7 puuntiil tan xub' se7n i17 qa'7n tan kyime7n

"Ah, we will look for a way to fool him in order to kill him.

81. kuuq qa'7n tk'u7 i jüen saak che7ch b'an

We will stick him in a gunny sack," they said.

82. ej ma yel wü niin tzun qab'okl niintz wü a7

"And when he sleeps, we'll throw him in the river."

83. kwëe7n tzun q'aq' tetz Lu7

So Pedro, making his fire.

84. tooke7n tetz Lu7 tan tx'ujte7n

starting to busy himself,

85. waane7n tetz Lu7tz,

(and) eating,

86. xa7q tzaaj Lu7tz skye7z

he (then) went to them.

87. jalchaan pe7 eeq chiky' kyeeru7 stzun b'an

"Are you going to move on early tomorrow?" said Pedro.

88. jalchaan che7ch ajpyya

"Early," said the merchants.
89. aaj b'a7n b'in stzun Lu7

"Ah, good," said Pedro.

90. qawitoq b'in na ja chinsaktij weetz stzun Lu7 b'antz

"Then let's sleep because I'm tired," said Pedro.

91. b'eene7n tzun kyiq'ool juun saaktz tetz Lu7

So giving a gunny sack to Pedro,

92. i tzun kyaal teetz nkween tul juun saake7; na wi7 niin che7w.

what they said to him was, "Get into this gunny sack because it's really cold."

93. kuuj chi Lu7

"Okay," said Pedro.

94. poro nsqeen eel xtxuum Lu7 teetz ye b'i117 kyajb'11 ye aqpyaaj te7;7,

But he had already imagined what the merchants were going to do to him.

95. ma ye tooke7n niin tetz aq'b'aal ye kyaaaje7n samp kyetz aqpyaaj tan waat1, niin tzun jee7 k'ootzoon tetz Lu7tz.

And when night came and the merchants had gotten numb from sleep, Pedro got up.

96. tooke7n tzuntz tan molche7n tzaaj kyeqiil e7chq b'e7ch kyeetz e7chq chiwaa7 e7chq chi7lo7x mb'i117 kunin ky'a7n kya7n,

And starting to gather all of their clothes, their tortillas, their napkins, everything they carried,
97. ryaatiil cheejtz kyeqiil ku7n b'eeene7n xajooltz tk'u7l te juun saaka7tz,
the horse's riata and everything he put in that gunny sack.

98. kwee7n tzun nuk'uuultz chi taane7n i17 teetz ye wite7n,
Then forming it like he had been while he slept,

99. tooke7n tzun tq'ool ye tw17 i17 xlaj
and then putting his hat beside it,

100. ma nuk'xe7n kyeen ta7n niin iky' tetz Lu7tz jala7 jye'en a7.
and when he had it arranged, he passed over to the other side of
the river.

101. ej jala17s tzun chinaachoone7n kyetz ajpyaaj
and then after that the merchants woke up.

102. poro sqeen iky'poon tetz Lu7 jala7 jye'en chumaam a7
But Pedro had passed over to the other side of the river.

103. ma chinaachoone7n ye e7 ajpyaaj
When the merchants woke up:

104. yaj stzun ye k'ase7n ye b'aax
"Hey you guys," said the first one awake,

105. oor tzijee7n oor tzijee7n
"Get up, get up.

106. tz'uul chaan sqiil.
It's already getting light.

107. kyil kxjiloon stzun juuna7tz b'an,

Don't talk," said that one.

108. jee7n tzun chipalool ye saaktz

So lifting up that gunny sack,

109. b'eene7n tzun chitrimpuultz wi a7

and throwing it into the river.

110. poro kyeqii7 chiwa7t kyeqii7 e7chq chitwl7 mb'ii7 kunii7n ky'a7n kya7n, ryaatiil cheej i kunii7n te7tz b'eene7n chitrimpuultz wi a7

But all of their tortillas, all of their hats, everything that they carried, the horse's riata and all of that they threw in the river.

111. i tzun b'eene7n kyiloottz tetz Lu7

So what they said to Pedro was:

112. cheeb' ku7n xhb'een Lu7 che7ch b'an niin

"Good bye, Pedro," they said.

113. kuuj ryaat niin tiir ku7n tamaal stzun Lu7 b'an tzaaj len jalaj ky'een a7

"OK, thanks, riatas and all of the tamales," said Pedro on the other side of the river.

114. poro nq'etzt Lu7 ye b'een chitrimpuul wi a7

But it wasn't Pedro that they threw in the river:
but rather it was their clothes, their tortillas, all of their things that Pedro put in that gunny sack; and they themselves threw it into the great river.

And in this way the merchants realized that it was not Pedro that they had screwed,

but they themselves who had gotten screwed by Pedro.

Thank you.
Maya Writing: Linguistic Evidence for Eastern Mayan Influence

Martha J. Macri

University of California, Berkeley

From approximately 300 A.D. to 900 A.D. the civilization known as the Classic Maya recorded in carvings and paintings their history, myths, and scientific knowledge. Writing done on perishable materials has not been preserved, and even many stone carvings have not withstood time. It is reasonable to assume that the hieroglyphic texts we do have from the Classic Period offer a small and probably unrepresentative sample of the entire corpus of Maya writing. In addition, three, perhaps four, bark-paper codices date from the Postclassic (900 A.D. - 1250 A.D.) and Proto-historic Periods (1250-1521 A.D.). Colonial sources report that many such books were destroyed by the Spanish.

Because of cultural similarities between the Classic Maya and Mayan speakers of the historic period, linguistic continuity has been assumed as well. As more is learned about the script, it becomes clear that the Classic texts do indeed record a language related to the thirty or so Mayan languages spoken in Mexico and Guatemala today. Several of these have been proposed as being more directly related to the language or languages of the Classic Maya than others. The Cholan and Yucatecan families have been the traditionally favored candidates (Thompson 1950: 160; 1977: 3), with more recent scholarship overwhelmingly preferring the Cholan (Kaufman 1976: 112, 117, Norman & Kaufman 220.
1979). In fact, the conference on phoneticism in Maya writing held at the State University of New York at Albany in 1979 seems to have been based on this premise.

Over the years, however, bits of information have been accumulating that hint at some kind of connection with the languages now spoken in the Guatemalan Highlands, particularly the Quichean family. Much of the information has been dismissed as flukes resulting from the imperfect knowledge we have of the genetic predecessors of the Yucatecan and Cholan families (Thompson 1950: 17; Justeson & Campbell 1979).

Recently a few scholars have begun to suspect more than a random relationship between the Classic Maya and the Eastern Mayan languages (Justeson 1978: 245-273 (rejected in Justeson & Campbell 1979); Dunning 1979: 183; James Fox, personal communication). In the following discussion I will demonstrate how internal evidence from the Classic Period texts suggests that certain signs were used by speakers of a language lexically and phonologically similar to Eastern Mayan, and that subsequently a more developed stage of the script was used to record a language or languages sharing phonological similarities with Yucatecan and Western Mayan.

Figure 1 shows the Mayan language family. Not all the relationships are shown in detail. Kaufman (1976: 107ff) gives the time depth for Greater Quichean (the Proto-Quichean of Campbell 1977) as 1400 B.C.-600 B.C. By 200 B.C. Greater Quichean split into Kekchi, Uspantec, Quichean Proper, and Pokom. At the time I propose for an early stage of Maya writing Quichean Proper would have been a single
Figure 1. Mayan Languages
language. Cholan begins to diversify around 600 A.D., and Yucatecan around 1000 A.D., well after hieroglyphic writing is established in the Lowlands.

A Mayan Sound Correspondence

Abraham Halpern in 1942 was the first to offer a historical reconstruction of Proto-Mayan. Since then Swadesh (1956), McQuown (1955, 1964), Olson (1964), Kaufman (1964, 1968, 1969), and Fox (1978) have proposed various sets of proto-sounds. With the exception of some of the earliest work, linguists have traditionally agreed that Proto-Mayan had three nasal phonemes: *m, *n, and *ŋ. Recently Fox and Justeson (1980: 209) have suggested *nw or *ŋʷ as more accurately describing the proto-sound represented by the so-called x/n correspondence. It is not, however, the proto-sound, but the reflexes of it as they existed at the time of the development of Maya writing which is important here. Therefore, while its actual features are the object of controversy, the symbol ŋ remains adequate for our purposes. Table 1 gives the reflexes in representative languages. *ŋ remains /ŋ/ in Greater Kanjobalan, becomes /n/ in Greater Tzeltalan and Yucatecan, and becomes /x/ in Proto-Quichean. In some Kanjobalan languages *ŋ has become /n/. The sound change *ŋ→x occurred before the diversification of Eastern Mayan around 1400 B.C. (Kaufman 1976: 106). The date for *ŋ→n in Yucatecan is unknown, but for Greater Tzeltalan it occurred after the split with Greater Kanjobalan around 1000 B.C. (Kaufman 1976: 107). In all probability it antedates the Proto-Classic Period (1 A.D.-300 A.D.).
<table>
<thead>
<tr>
<th>GLOSS</th>
<th>PROTO-MAYAN</th>
<th>JACALTEC</th>
<th>PROTO-QUICHEAN</th>
<th>YUCATEC</th>
<th>CHOL</th>
<th>PROTO-CHOLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>sky</td>
<td>*ka7n</td>
<td>sat kən</td>
<td>*kaaŋ</td>
<td>ká7n</td>
<td>čan</td>
<td></td>
</tr>
<tr>
<td>day,</td>
<td>*q'iin</td>
<td>*k'iinw</td>
<td>q'iŋ</td>
<td>*q'iix</td>
<td>k'ǐn</td>
<td>*k'in</td>
</tr>
<tr>
<td>sun</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>house</td>
<td>*ŋaah</td>
<td>ŋah</td>
<td>*ŋaah</td>
<td>nah</td>
<td>ŋa-3</td>
<td></td>
</tr>
<tr>
<td>four</td>
<td>*keen-</td>
<td>káŋeb'</td>
<td>*kəxib'</td>
<td>kan</td>
<td>čən</td>
<td>*čən</td>
</tr>
<tr>
<td>nine</td>
<td>*b'eleŋ-</td>
<td>b'áluŋeb'</td>
<td>*b'eeleŋeb'</td>
<td>b'olon</td>
<td>b'olon</td>
<td>*b'olon</td>
</tr>
<tr>
<td>ten</td>
<td>*lxuŋ-</td>
<td>láhuŋeb'</td>
<td>*laxuux</td>
<td>lahun</td>
<td>luhun</td>
<td>*ləhun</td>
</tr>
<tr>
<td>avocado</td>
<td>*ʔooŋ</td>
<td>*oonw</td>
<td>oŋ</td>
<td>*ooŋ</td>
<td>oon</td>
<td>*un</td>
</tr>
</tbody>
</table>

Table 1. Reflexes of *ŋ

Evidence from Maya Writing

There is evidence that this sound correspondence is reflected in the glyphs. Two combinations occur very frequently in Maya writing: T544:116 and T561:23 (Figure 2, 1-6). T544, k'jin in Yucatec and k'in in Chol, means 'sun, day.' Kelley credits Brasseur de Bourbourg with the identification (Kelley 1962: 6) which is now universally accepted. Fox and Justeson (1980: 212) assign the value ne or ni to T116. They quote Lounsberry as having independently assigned it the phonetic value -n.

The na reading for T23, originally proposed by Knorozov (1967: 79), now seems established beyond a doubt (for example, Lounsberry 1979). Kelley discusses T561 and T23 together:

As originally defined by Seler (1902-1923), the glyph was recognized as 'sky' in a context where it had T23 affixed. This very puzzling affix is sometimes present and sometimes absent with no obvious change in context. T561 had been read as caan 'sky' on the basis of its context and its association with the so-called planetary band (1976: 150).

Again he writes of T23:

It is one of the few affixes which are sometimes present and sometimes absent without obvious changes either in grammatical structure or meaning....If Knorozov is correct, T23 may be present on such glyphs as caan 'sky' as a phonetic determinant. However, it appears so rarely in such contexts that this explanation does not seem likely (1976: 185).

The suggestion that the affix is a phonetic determinant, a
1. T544 k’in
   -in
   'day, sun'

2. T116 ni-

3. T516 kaan 'sky'

4. T23 na-
   -an

5. T544:116

6. T561:23

7. Woman from Piedras Negras
   Stela 3

8. T281:23 k’an 'yellow,
   ripe' Yaxchilan, Lintel
   10, D2

Figure 2.

(Sources: 1-4, Thompson 1962; 5, Thompson 1950: Figure 26, 51 & 55; 6, Kelley 1976:148; Proskouriakoff 1961:17; 8, Graham & Von Euw 1977.)
suggestion which Kelley rejected, has also been proposed by Taak. In reference to the 'sky compound' as well as two other pairs of signs he writes, "Although it has not yet been determined why the above logographic signs require a VC determinative suffix, the most plausible explanation would be polyvalence of the CVC stem" (1977: 288). Justeson, commenting on the frequency of T544: 116 feels it, too, may simply be an extension of an otherwise functional practice of phonic complementation (1978: 273). This explanation accounts for neither the extensiveness nor the uniqueness of such a non-functional use.

On the earliest stelae T544 occurs without the -n suffix. Justeson makes an intriguing suggestion: "The absence of the T116 complement could indicate that the language being written on the Baktun 8 monuments was not Lowland" (1978: 273). That is, perhaps T544 was read as q'ix rather than k'in. This negative evidence from the earliest inscriptions is rather weak since examples are limited. However, examples from Classic texts of T544 and T561 being followed by an -n suffix are extensive. Following Justeson then, a very satisfactory explanation would be that prior to use by Yucatecans or Cholans, the signs meant q'iix 'day' and kaax 'sky' in Eastern Mayan. When the writing system was adopted by Lowland speakers, they often, but not always, indicated their pronunciation by the addition of a suffix containing n and the appropriate vowel.

The occurrence of an -n suffix also on signs for words which have reflexes of *n would prove that the suffix stands for /n/ and not simply for any reflex of *ŋ. For example, if T116 does stand for 'tail,' a Quichean speaker would read it xe, and a Yucatec speaker neh.
The T544.116 and T516.23 combinations might just as easily represent the words q'iiix and kaax. However, the T23 affix occurs with portrait glyphs of women to indicate na 'woman' (from *na) and with the glyph for 'yellow,' k'an (from *q'an) so its use is not limited to words containing reflexes of *ŋ (Figure 2, 7-8).

Conclusion

Therefore, the Classic Maya script cannot have recorded an Eastern Mayan language. The sign that represents /n/ would also represent /x/. In the language or languages recorded in the glyphs *ŋ had merged with *n. The only extant language families in which this has taken place are Cholan, Tzeltalan, Yucatecan, and in some Kanjobalan languages. It would seem that the language spoken by the Classic Maya comes from one of these groups.6

In addition, the fact that phonetic complements were at some point needed to indicate /n/ in logographs that would not have had /n/ in Eastern Mayan, supports the hypothesis that certain signs used by the Classic Maya were previously used by Eastern Mayan speakers. The first writers of the Classic script knew the pronunciations of these signs in the Eastern Mayan language, and modified the signs with affixes so they could accurately record the sounds of their own language.

The 'Lowland' language identification of the Classic writing system has, of course, been accepted as a working hypothesis by glyphicists for many years. The evidence presented here gives added confirmation to the hypothesis that the Classic inscriptions do not record an Eastern Mayan language, but one related to Western Mayan or
Yucatecan.

Kaufman (1976: 117) identifies the Classic Maya of the Peten as Cholan speakers, and those of Yucatan as Yucatecan speakers. There are, however, more than two groups whose cultures were distinct enough to suggest the possibility of language differences. Northern Yucatan, the Peten, Usumacinta River sites, Palenque, Quirigua and Copan, and a late foreign presence at Seibal probably represent more than just two linguistic groups. At this time the language of the Classic Maya remains an open question. Given the geographical and temporal span, it is likely that more than one is involved. It is also possible that the language or languages recorded in the inscriptions is not everywhere that of the common people. Furthermore, the language of the Peten Maya, while certainly related to Yucatecan or Greater Tzeltalan, may have left no direct genetic descendents. Given the thoroughness of the collapse in the tenth century, it could have included the extinction of a language.

The identification and distribution of Mayan languages at the time of the Classic Period is not an unsolvable problem. Careful analysis of archaeological, linguistic, and glyphic data can provide important information for our understanding of Mayan linguistic history, and of historical linguistics in general.

Footnotes

1 The term Classic Maya in this paper refers only to the civilization using Long Count dates and a style of hieroglyphic writing shared throughout the Peten and surrounding areas. It does not include many
sites in the Northern Lowlands.

The dates given here are from Kaufman's "Archaeological and linguistic correlations in Mayaland and associate areas of Meso-America" (1976). One of five criteria on which the dates are based is glotto-chronological calculations. Campbell (1977: 63-65) argues effectively against the use of this method. So while Kaufman's sketch of Mayan prehistory is the most detailed to date, and was done by a scholar whose knowledge of Mayan languages is unsurpassed, it remains theory rather than fact, and the actual dating of particular events is probably its weakest point. I give his dates as being reasonable estimates when not contradicted by other evidence.

\footnote{3} Occurs as a bound morpheme: 'base' c'Al'cib; 'base de una casa' nacilib.

\footnote{4} A 'T' followed by a number indicates a sign listed in Thompson's 1962 Catalog. A colon indicates that the following sign is beneath the first. A comma indicates the second sign is to the right. Most phonetic signs represent CV or VC syllables. Because of the CVC root structures in Mayan languages, the second consonant is often indicated by a phonetic syllable having the same vowel as the first sign. This is common practice among users of syllabic writing systems throughout the world.

\footnote{5} Baktun 8 covers approximately the first 400 years of this millennium, that is, the Proto- and Early Classic Periods.

\footnote{6} Y remains /ŋ/ in Jacaltec and Chuj. In other Kanjobalan languages such as Kanjobal and Tojolabal it becomes /n/. Any Kanjobalan language which had /ŋ/ during the Classic Period would group with
Eastern Mayan in having three distinct reflexes of Proto-Mayan nasals. Those not having /ŋ/ would group with Greater Tzeltalan and Yucatecan.

References

AULIE, H. WILBUR and EVELYN W. AULIE. 1978. Diccionario Ch'ol-Espanol, Espanol-Ch'ol. Instituto Linguistico de Verano, Mexico, D.F.


---. 1979. Workshop at the conference on phoneticism in Maya hieroglyphic writing. SUNYA April 20, Albany.


LOUNSBURY, FLOYD. 1979. Notes on some substitution patterns in Maya
hieroglyphic writing. Paper delivered at the conference on phoneticism in Maya hieroglyphic writing. SUNYA April 20-22, Albany.


---, and DONNAN CALL JEFFERS, eds. 1979. Tercera Mesa Redonda de Palenque Volume IV. Pre-Columbian Art Research Center, Palenque, Chiapas, Mexico.


---. 1977. A proposal for constituting a Maya subgroup, cultural and
linguistic, in the Peten and adjacent regions. In Jones 1977: 3-42.

The Use of Body-Part Terms as Locatives in Chalcatongo Mixtec

Claudia Brugman

University of California, Berkeley

The expression of locational notions is achieved in a variety of ways in the languages of the world. European languages, for the most part, express locations through a system of cases and/or prepositions (for instance, the box is in the room on the chair). Another widely-used mechanism is to use extended meanings of body-part terms in locating expressions.

English exhibits one kind of semantic extension in its use of body-part terms to refer to parts of noncorporeal objects, as in arm of a chair or pants leg. The English case, however, is somewhat uninteresting since the extended uses occur through transparent metaphor (based on perceived similarity and perceived close association of the objects, e.g. the leg on a body and a pants leg). Additionally, the extended English uses still name objects rather than locations. English examples like these occur sporadically rather than systematically.

A more interesting and regular system of correlations between body-part terms and locating expressions occurs in Chalcatongo Mixtec. In this language, body-part terms are used in at least four distinguishable ways: 1) to refer to parts of the body; 2) to refer to sub-parts of other objects, based on perceived similarities to the corresponding subparts of the body: these uses take the subpart of the object as an object; 3) again to refer to subparts of other objects, but taking the named subpart as a location rather than as an entity: these uses often express not only the shape of the location but also the particular type of locative relation (cf. in v. on); 4) to refer to areas outside the boundaries of the object, areas associated with the subpart named by the body-part term. These uses also often convey the type of locative relationship. As implied above, Mixtec body-part terms fill two sorts of roles in the language: the role of some "lexical" morphemes (cf. pants leg) and that of some morphemes which are arguably "grammatical" (cf. on the table). This fact suggests that the Mixtec
system of locative expression is quite different in kind from any European-type system of prepositions. Moreover, among the particular body-part terms themselves, some are more readily and regularly used in abstract, nonlexical locating expressions than are others. This broad range of variations in the level of abstraction achieved by each particular locative term can suggest which elements of meaning of a lexical morpheme are most amenable to extension or abstraction, that is, which semantic components are most likely to motivate the move to grammaticization of a morpheme.

Nominal Compounds in Mixtec

Body-part terms in their most literal uses exhibit all the semantic and syntactic behavior we would expect from true nouns. They name objects, can function as nuclear elements in a clause, can be modified, and so on. When used in either a literal, lexical manner or in their extended, grammatical meanings, they almost invariably occur as the first in a sequence of two nominals. It is important to establish the pervasiveness of the nominal compound construction in order to motivate some of the discussions to follow.

Macri (1981) has noted that in nominal compound constructions, "the second [noun] stands in a genitival relationship to the first." [p. 2] The concept "genitival", of course, needs to be explained, since the range of relationships that one could consider genitival have not all shown up in Chalcatongo Mixtec. I shall just define "genitival" here as being either a partitive or a possessive relationship, with the caveat that this definition will not prove completely adequate to characterize the data to follow. It is nevertheless an intuitive enough first approximation to allow us to proceed.

A definition of a "nominal" is also in order at this point, although, again, it is a definition that will prove troublesome with an exposition of the particular constructions under discussion here. I include as "nominals" here pronominal morphemes of both the full, unbound forms, and the greatly reduced, enclitic forms. The full forms occur syntactically just where
nouns can, i.e., as arguments of a predicate. The reduced, enclitic forms occur bound to verbs and predicate adjectives, indicating Subject; on prepositions which are not synchronically related to body-part terms, indicating the prepositional Object; and on nouns to indicate the possessor of the object named by the full nominal or the entire object whose part is named by the full nominal. Examples of each of these uses are given in (1)–(5).\textsuperscript{3}

(1) ndukoo-ri'
be+seated - 1sg.
'I'm sitting down.'

(2) lalif-6e
small-3sg.
'He is small.'

(3) fi?a  hii-rf
come+pot. with-1sg.
'Come with me.'

(4) s - ndoo  ßé?é-ro
cause-be left+pot. house-2sg.
'Leave your house!'

(5) ru?dì ni - nakaš-ri  nda?ù-ri
1sg. perfv.-wash - 1sg. hand-1sg.
'I washed my hands.'

The details of the distribution of the enclitic pronominal and that of the full-form pronominal are yet to be worked out; but what is at issue at this point is that the enclitics function in nearly identical syntagmatic positions as do nonpronominal nouns or full pronominals; thus to conclude that they are "nominals" for the purpose of the nominal compound construction is uncontroversial. Examples (5) and (6)–(9) demonstrate that including enclitics
as nominals in the context of nominal compounds allows a unified treatment of all constructions that express a genitive relationship between the two morphemes.

(6) soʔ-o-kif  ni - kaʔa tendaʔa
    ear -horse  perfv.-enter worm
    'Our horse has worms in its ear.'

(7) ndaʔ-a-ydn  kə kaʔnə
    hand -tree be big
    'The tree's branch is big.'

(8) šini-rə  ?də
    head-1sg. hurt
    'My head hurts.'

(9) nə - ndəʔ-e ini ydʔ-aʔnə
    perfv.-look in mouth-3sg.fm.
    'He looked inside her mouth.'

Body-Part Terms in Nominal Compounds

An understanding of the productive capabilities of nominal compounds is necessary in order to understand the significance of the use of body-part terms in such constructions. When used in locative constructions, the body-part term is the first nominal in the compound, a position which invites an interpretation consistent with those of the nominal compounds we have seen so far, i.e. where the first noun denotes a subpart of the entity denoted by the second noun. The specification of a subarea of an object readily suggests a use as a locational. It is also a position relative to the second noun which is syntactically parallel to the position held by prepositions that are not synchronically related to body-part terms.
Greenberg (1963) gives this syntagmatic parallelism as a universal: “In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes.” [p. 78] In the cases of a most literal partitive genitive relationship between the two nominals, the use of the body-part term itself is best motivated, since the extension of the body-part term is a metaphorical use applying a configuration of the human body to a similar configuration of another object. Example (7) above is a case in point. But this clear partitive case is not, strictly speaking, a location, at least in (7). It names what we can understand to be an entity rather than merely a place. A clearer locative noun phrase, and one again demonstrating the superimposition of a bodily configuration onto the configuration of another object, occurs in (10):

(10)  hiyā’ā  ḍe  šini - yuku  
be+located-3sg.m. head-hill

‘He is on top of the hill.’

My claim that šini-yuku names a location is supported by the presence of the verb ‘be located’.

Both (7) and (10) must be distinguished from (11), where the first nominal in the construction names not a subpart of the second, but an area in space associated with that subpart:

(11)  ni-ndeč’ ʔn sαd  šin-pyínu  
perlv.-fly one bird head-tree

‘The bird flew over the tree.’

This is the sort of case in which these nominal compounds begin to behave more like the preposition + nominal constructions that also occur in Mixtec. Here we have departed from the general scheme outlined above in two ways: first, obviously, we are no longer naming a subpart of an object when using the nominal compound (although we are naming an area in space which is easily associated with the object named by the construction). Secondly, šini-
yunu, even in its new "relational" meaning, names only a single point in the path which is expressed in the verb. We shall also see that other locationally-derived notions such as source, extent, and goal are expressed via these constructions.

Thus the range of extensibility of use of body-part terms to locationals depends upon many variables. We can use transparent metaphorical association to refer to a subpart of an entity by using a term for a perceptually similar subpart of the human body. Or we can step outside the confines of the strict genitival relationship and exploit the construction by using it to refer to an area in space rather than an area of an object.

The Role of Canonical Spatial Orientations

Each of the body-part terms under discussion has as one of its characteristics a location relative to the body as a whole in its canonical position. For instance, on the human body the head is the topmost tip. The back is vertically oriented and defines the dorsal surface of the body. The body and some other objects have inherent orientations: humans are canonically (though not unalterably) upright, with their faces and ventral surfaces determining their fronts. This is intuitively sensible, since our primary means of locomotion requires us to be upright and facing the direction of motion; our ventral side includes most of our perceptual and motor apparatus; and so on. Other objects have a front and back based on similar properties, such as direction of locomotion or perceived similarities to human fronts. Perceived similarities to human orientation also account for perceived uprightness of objects such as mountains and trees, which in Mixtec can be described as having "heads" (see Fillmore (1971) for more explication of these issues).

Some objects have no inherent orientations like those above: rocks and tables, for instance, have no inherent front or back. In Mixtec as in English, front/back orientation is assigned to such objects with the surface closest to the reference points being the front, and its further side being the back.
So both inherent and assigned spatial orientations are important for the use of corresponding body-part terms to refer to subparts of objects. In some cases, e.g. (11) above, the body-part term does not refer to a subpart of an object, but is used in other more abstract situations. Canonical spatial orientation provides some of the links in the motivation of the use of some of these body-part terms. (12) is an example.

(12) ni- haʔa - riʔ ?n kitî nū′i- seʔ-e-ro
perfv. pass-1sg. one horse face-son-2sg
'I gave a horse to your son.'

Other evidence in the language indicates that nū′i 'face' marks energy goals of all sorts, but this particular sentence is a well-motivated use of this convention. In the typical instance of giving exemplified in (12), the participants named as Subject and dative Object are facing each other so that the face is a salient subpart of each. In this case it is not merely the spatial orientation of the participants, which is inherent in the Subject and dative Object, which motivates the use of nū′i. What is being exploited is the usual configuration of the participants with respect to each other, i.e. that they are facing.

The Data

The body-part terms whose use I will be discussing here are the following:

šini 'head'
cii 'belly'
nū′i 'face'
yata 'back (human)'
ndeʔa 'hand'/'arm'
siki 'back (animal)'
haʔa 'foot'/'leg'
In addition, the element sêti 'road' plays an important part in the use of these terms and
behaves in roughly the same way as the body-part words, so it will also be discussed. De
Alvarado (1593), Pensinger (1974), and de los Reyes (1889) indicate that in other dialects of
Mixtec the word sêti, which the consultant, Mr. Cortés, glossed only as 'inside' or 'into', can be
translated variously as 'spleen', 'center of the emotions', 'stomach (the internal organ)', or
'heart', as well as with the prepositional use. This evidence suggests that sêti was extended
into these locational uses relatively early, and, at least in the idiolect of my consultant, has
lost its basic, literal meanings (he uses Spanish borrowings for 'heart' and 'stomach'). Simi-
larly, Hills and Merrifield (1974) claim that the word for 'with' in Ayutla Mixtec is nãñ, derived
from the word for 'side'. Again, my consultant does not recognize a relationship between his
word for 'with', Àñî (cognate with the above Ayutla form) as being related to any body-part
term. Note in example (18) below that he uses the word for 'hand' to express the area in
space that we gloss as 'side'.

Each body-part term has particular idiosyncratic limits on the extent of its usability in
nonliteral senses. In each case, important elements of the meanings of the body-part term
will be its overall shape and/or its location relative to the entire object named by the second
noun in the compound, which implies that the shape of the object will select the body-part
terms appropriate to it.

nda?a 'hand'/ 'arm'

The Mixtec word for 'hand'/ 'arm' is by far the most restricted term in its range of ex-
tended use. It can be used only to refer to subparts of an object where that subpart is a
limblike structure: thus it exhibits the most transparently metaphorical type of extension.
Some examples follow:
(13) nda?d-yuča
    arm - plant
    'twig of a plant'

(14) nda?d-yčku
    arm - river
    'tributary of a river'

(15) nda?d-yčnu
    arm - tree
    'branch of a tree'

As mentioned above, these subparts need not be construed as locations but may be considered to name objects in their usual relationships to a larger object. A case in point is in (16), where the nominal compound is in Subject position, acting as a patient rather than a location:

(16) nda?a-yuču tā?nuč
    arm - tree    split-off
    'The branches of the tree are splitting.'

But nda?a can express a subpart which marks a location:

(17) ndučšo-o-nda?a-yčnu
    sit - 1sg    arm - tree
    'I'm sitting on the branch of the tree.'

Again: in (17) nda?a refers to a subpart of the tree. I could not elicit any sentences where nda?a refers to an area in space except those, like (18), where nda?a is preceded by čči
‘road’, which itself usually suggests an interpretation of an extended location of motion rather than a stative location; and that interpretation requires a "relational" reading of nda?a:

(18) iʃi-nda?a βd-ro
    road-arm right-2sg.
    ‘to your right side’

In this case, as with most other relational uses of body-part locatives, the shape characteristics of the named body-part cannot be preserved, since what is being referred to is an area in space. But since that area in space is contiguous with the named body-part, something akin to the relative location component of the body-part term is preserved. I would guess that this preservation motivates what we would call nontransparent uses of the body-part terms as it does here, where I have translated nda?a freely as ‘side’ (in the sense of "direction" rather than "flank"). Arms are located laterally with respect to the body and are often used to indicate direction. The relative location, then, cooccurs with a common function of this location (with respect to the body and to front/back orientation—see again Fillmore (1971)) to motivate the use of ‘arm’ to indicate direction.

In any event the most easily and commonly elicited extended uses of nda?a were those exemplified by (13)–(15), which is the subarea type of extension: a linear subregion of the object which branches off from the main portion of that object can be referred to as nda?a.

ha?a ‘leg’/’foot’

The term for ‘leg’/’foot’ appears to be extended only in the ‘foot’ sense, as nda?a is extended only in the ‘arm’ sense. That is, both nda?a and ha?a can refer either to the limb or to the terminus of that limb, but, in its extended uses, nda?a is regularly used to refer to limblike parts and ha?a to the terminus. There are many possible reasons behind this, but at least it is a most efficient way of dividing the load of conventional extended senses. A consequence
of this division of labor is that the semantic component which serves as the basis for extension of each will be affected. As we can see in (13), the location of the limb relative to the whole object need not be analogous to the location of an arm relative to the body; that is, we cannot readily identify an upward extent of this river to which an arm would be appended. In the case of ndaʔa 'hand' the shape of the subpart is clearly more essential than the relative location.

haʔa 'foot', by contrast, does not take shape into account at all in its extended uses. Because haʔa encodes the relative location, any lowest extent of an object perceived as being upright can be expressed with haʔa, and the nominal compound with haʔa as its first nominal can be used to express areas in space as well for stative locations and extents or points in extents along which motion takes place:

(19) yuu wā hiyaʔa haʔa-mesa
    stone det. be=loc. foot-table.

   'The stone is at the table leg.'

(English: "foot of the table" is not a suitable gloss for the expression given: in Mixtec)

(20) ni-ndukoc-ri haʔa-yur. wāŋ
    perfv-sit - 2sg. foot-tree det.

   'I sat at the foot of the tree.'

(21) kiʔiʔ-ri iči-haʔa-yčkə
    go+pot.-1sg. road-foot-mountain

   'I am going around the base of the mountain.'
grammatical while (27) is not:

(26) səə wə̀ ni - ndeqi ću-yənu wə̀

bird det. perfv.-fly belly-tree det.

'The bird flew under the tree.'

(27) səə wə̀ ni - ndeqi haʔ-a-yaʔu wə̀

bird det. perfv.-fly foot-tree det.

'The bird flew under the tree.'

This fits intuition since, as noted before, the area in space associated with a foot is most probably the surface on which it is resting, but flying cannot take place on a solid surface of that sort.

However, the use of ću as 'under' in (26) is not simply a conventionalization of the use of ću as 'under' as exemplified in (23). That is, ću does not mean 'under' independent of the shape of the object: ćuɣənu in (26) really names a space associated with a "belly" subpart of the tree trunk:

(28) isu wə̀ saʔa ?n yə̀ ću yənu wə̀

rabbit det. make one hole belly-tree det.

'The rabbit is digging a hole in the tree trunk.'

This certainly refers to a space in the tree itself, and names a subpart different from that named in (29), in which haʔa names the subpart expressing the location:

(29) isu wə̀ saʔa ?n yə̀ haʔ-a-yənu wə̀

rabbit det. make one hole foot-tree det.

'The rabbit is digging a hole [at the base of/under] the tree.'

Here, haʔ-a-yaʔu refers to a location in the ground. In contrast to both of these situations,
(22) yīdō ke?nū te tu?nū ke?i ke?i ha haʔa iže-haʔa-ydku

there is one road large and split-off one road small compl. pass road-foot-mountain

'There's a big road and branching off from it
a little one that crosses the foot of the mountain.'

Unlike ndaʔa, which cannot by itself be used to indicate an area in space, (20) shows
that an area outside of the object itself can be referred to with haʔa. We must understand
the location referred to in that sentence to be a place on the ground rather than a subpart of
the tree (a predictable extension, since a lower extent of an object is usually contiguous with
a surface on which it rests). But to provide another parameter of extended use, a "locational
extent" reading requires iže preposed to haʔa, as evidenced by (21) and (22).

haʔa exhibits transparent analogy with the body insofar as it cannot be used for the
lowest extent of just any object. It must be an object that is relatively upright, or one that
has a substructure similar to a leg. The use of haʔa interacts with that of ẓii 'belly', some-
times overlapping in use with it, sometimes occurring in semantic contrast with it.

ẓii 'belly'

We have seen that haʔa regularly gets extended on basis of its location relative to the
entire object, with no concern given as to whether the location named by haʔa + noun resembles
a foot or a leg in its shape. ẓii 'belly' is extended rather as haʔa is in the respect that
the shape of the named area is again not so important as its relative location, but there are
two canonical body positions on basis of which the location of ẓii apparently is extended, that
of a person and that of a four-footed animal. Both allow a "relational (i.e. area-in-space)" use
of ẓii as well as a strict partitive reading. Most commonly the relational use can be glossed
in English as 'under'. In (23) the use of ẓii is apparently on analogy with a quadruped and the
location of its belly.
(23) yuwa wā hīyad ʔiʾ-месā
stone det. be+loc. belly-table
'The stone is under the table.'

This describes a situation in which the stone is lying on the same surface on which the table is resting (e.g., a floor), but is directly beneath the table top. By contrast, (19) describes a situation in which the stone is again resting on the same surface as the table, but is beside one of the legs. (19) is repeated here.

(19) yuwa wā hīyad haʔa-месā
stone det. be+loc. foot-table
'The stone is at the table leg.'

Here, because of the shape of a table and the transparent anology with quadrupeds, ʔiʾ and haʔa name quite different subparts and associated areas in space. (24) and (25) give a case where ʔiʾ and haʔa name roughly the same area:

(24) kwādā ʔiʾ-ʔynu wā
go+pot. belly-tree det.
'Go under the tree.'

(25) kwādā haʔa-ʔynu wā
go+pot. foot-tree det.
'Go to the foot of the tree.'

These are both grammatical sentences, and they denote about the same location, since the base of the tree is under its trunk and branches. However, even when referring to trees these two terms do not always name the same area, since there are various verbs that require or prefer the use of one or the other of the body-part terms. For instance, (26) is
(28) can be used to describe the situation in which the tree being referred to is fallen and the rabbit is digging in the ground under the fallen trunk. Here, as was the case in (23), the analogy of a quadrupedal animal is being taken as the basis of extension for \( \ddot{\text{c}i} \) and an associated area in space is being referred to. The ambiguity of (28) is therefore consistent with regular principles by which \( \ddot{\text{c}i} \) is extended.

\( \ddot{\text{c}i} \) provides very strong support for an analysis of these locative morphemes being based in body-parts. If we were to take the locative \( \ddot{\text{c}i} \) and its corresponding body-part term \( \ddot{\text{c}i} \) as (synchronously) separate, homonymous words, we could not provide a coherent account of uses of the locative \( \ddot{\text{c}i} \) which includes reference either to the middle of the front surface of a vertically-oriented object or to the underside of a horizontally-oriented object, and to areas in space associated with each kind of subpart. The locative \( \ddot{\text{c}i} \) would have to have two separate definitions, and each would have to have selectional restrictions stating the required shape of the object named by the noun following \( \ddot{\text{c}i} \). If one takes the body-part term as a basic sense, synchronically related to the locative uses, then we can apply the same general principles required by the rest of the system, and produce the two distinct locating areas for \( \ddot{\text{c}i} \) analogy with either a bipedal or a quadrupedal animal’s body, and allowing the usual place for extension to associated space (as in (23) and (26)). We are not required to postulate complex selectional restrictions for the second noun (not to mention the verbs and other elements in the sentence which would affect selection). Positing the analogy and recognizing the two types of body configuration which serve as bases for the analogy provides a more elegant and coherent means of capturing the phenomena.

There is independent support in the language for postulating the two types of body configuration as a basic parameter for extending body-part term usage. As we shall see below, the human-versus-animal distinction arises again in the lexical distinction made in \( \text{yata} \) and \( \text{siki} \); both translate as ‘back’ in English, but the first refers only to backs of humans and the second to backs of animals, and in their locative senses the analogous areas of inanimate objects are referred to by the two morphemes. While there is no corresponding lexical distinction made between the bellies of humans and those of animals, the single available word \( \ddot{\text{c}i} \) seems to perform in the way \( \text{siki} \) and \( \text{yata} \) together do for backs.
ču can be used as a locational extent or goal freely and without any further specification such as the use of ści 'road'. (30) and (31) give examples of this use of ču.

(30) čuku wá ni -ha?a či-yūnu wá
fly det. perfv-pass belly-tree det.
'The fly flew under the tree.'

(31) čuku wá ni ndeči či-yūnu wá
fly det. perfv-fly belly-tree det.
'The flew flew around under the tree.'

The idiomatic translation of (31) is intended to provide the contrast between the interpretations of the two sentences. In (30) ču-yūnu names a point in a linear path. In (31) the same nominal compound names an extended location over which the entire activity takes place.

Notice that no change in the nominal compound has occurred: the informational status of the locative expression is recovered from information given in the verb. Kuiper and Merrifield (1975) and Macaulay (1982) have studied verbs of motion such as the ha?a of (30) and have noted that relative location of three anchoring points is encoded into many of these verbs. ha?a requires that the nominal compound name a location which exists on or near the linear path which the verb itself gives. There is no similar restriction on the locative expression following ndeči, which encodes manner rather than direction of locomotion, and in this case information given by the subject is helpful too, since a fly is small enough that a relatively small area in space like that named by ču-yūnu could still be large enough to accommodate an entire act of flying (by contrast, Mr. Cortés told me that replacing ‘fly’ with ‘bird’ in (31), if grammatical at all, would force an interpretation of the locative expression given in (30) (i.e. "flew in a path under the tree").

There are some cases of ču which are glossed as ‘because’. Phonologically this morpheme is identical to the recognizably locative ču. Syntactically, it behaves entirely unlike the body-part locative elements under discussion here, since it can appear in a sentence before words of many lexical classes, e.g.
(32) .GetItem Kids-indat-u  go? \\
because pl-wait-3 2sg. \\
'Because they are waiting for you.'

(33) .GetItem Kids-nde?e-taata?u \\
because want-see -3 grandparent \\
'Because she wants to see her grandparents.'

(34) .GetItem Biha ki-ta \\
because now go-pot-1sg. \\
'Because I have to go now.'

These sentences show that the Kii of 'because' is not subject to the same morphosyntactic restriction as the Kii of 'under', etc., since none of the words following Kii in (32)–(34) are nouns; and even in the case that a noun does follow this sense of Kii, it cannot be interpreted as naming an object of which Kii names a subpart. It is included here because of the possibility that this word is historically related to the body-part and locative uses. One possibility is that via the use as "under", Kii came to signify supporting material, underwent a semantic shift (a shift in domain) and a reassignment of lexical category, with the corresponding syntactic behavior. Synchronously it seems clear that the word meaning 'because' is a homonym of the body-part locative term.

yata 'back (human)'

yata contrasts with one sense of Kii in that they both refer to a subpart of an object of the same general shape. If Kii refers to a subpart of an object that is closest to the speaker or other reference point (i.e. its "front"), yata is the corresponding area of that object that is furthest from the point of reference (i.e. its "back" surface). This again is a consequence

251
of the general principles by which front/back orientation is assigned to objects without inherent orientation.

The commonest occurrence of the locative *yata* is in the relational uses. Mr. Cortés was usually unwilling to assign a *yata* to an object without an inherent, upright back surface; but he freely produced many sentences in which *yata* is best glossed in English as 'behind'--i.e. cases where *yata* referred to an associated area rather than an object--but which, in a sense, presupposed that the corresponding subpart existed as such and thus could be used to refer to an area in space. This suggests that, while the appropriate use of *yata* does depend upon a recognizable analogy with the human body, the relational meaning has achieved a level of independence from the question of whether the named object has a *yata* or not.

In this sense of 'behind', *yata* is freely used as a stative location, locational extent, and goal:

(35) nde?e čađ ha hindi yatâ - hâsiʔi wô
    look-pot. man compl. stand back-woman det.
    'Look at the man who is standing behind the woman.'

(36) haʔa - rd yatâ - bêʔe uniʔ
    pass+pot.-1sg. back-house three
    'You go (to a place) behind the third house.'

(37) kwâʔa (iʔi-)yata-yânu
    go+pot. (road-)back-tree
    'Go behind the tree.'

(38) čâʔa kendâ yata - bêʔe
    go+hortatory exit+pot. back-house
    'Let's go [in back of/outside] the house.'

When using *yata* as a locational goal, it is preferable to have *vî* before *yata*, as in (37). *vî* is
also optional and occurs frequently with yata when expressing stative locations (this is a mystery which will be touched on again briefly in a future section).

(38) exhibits a further level of extension, from analogical spatial configurations to a more general location. The sentence is ambiguous over the two glosses provided. This is another case in which the deictic anchoring of the verb is illuminating, since in all attestations of yata where it was used for 'outside', the verb encoding the activity of which yata + noun is the goal indicates that the speaker is inside the named location at the time of the utterance. For instance in (38) kendo can be glossed as 'exit' or 'move out' (after Macaulay (1982)), which, along with the hortatory verb ěpoon, places the speaker and hearer inside the house. If in fact my attestations constitute an accurate sample, this deictic anchoring of the verbs motivates the use of yata as 'outside', since from the inside of, e.g., a house, the outside is on the other side of the wall from the speaker—a location which could easily be seen as an instantiation of 'behind'. This seems to be another case of a conventionalization of use, since it does not strictly fit the pattern of yata we have seen: if yata still meant just 'behind' here, we would expect 'wall' to be the second nominal in the compound instead of 'house', since houses have recognized fronts and backs not dependent upon the location of the speaker. It is a conventionalization which is nonetheless consistent with the processes of analogy and deixis otherwise exhibited.

siki 'back (animal)'

Siki corresponds to one sense of ěiti in just the way that yata corresponds to the other: it specifies a top horizontal surface as ěiti specifies the underside of a horizontal surface. siki overlaps in function with two other terms discussed below, ěiti 'head' and niti 'face'. All three of these terms can be used in the partitive sense in a way we would gloss as "on", and in those situations the shape of the locational object plays a large role in the choice of locative term. Other factors are conventions of the language and pragmatic considerations, both of which will be mentioned below. The body configuration that is relevant when using siki is more than just the presence of a horizontal back as of an animal. Either horizontal or
sloping surfaces can be considered as instances of sīkī, and the surface can be almost linear (as in (39) below) or can be planar. The object itself, however, must be three-dimensional—that is, sīkī cannot name part of a planar object (cf. (40)).

(39) ni-nddkod-ri'sīkī-hīkā-bē?ē
    perfv.-sit - 1sg. back - wall - house
    'I sat down on top of the wall of the house.'

(40) se?e-ri' hīdū
    son-1sg. lie
    [sīkī
        back
        ydu
    mat
        face
    ]
    'My son is lying on the mat.'

nūū 'face', discussed below, is regularly used for naming the top surface of a two-dimensional object. When the object is three-dimensional and planar, either sīkī or nūū is acceptable (although there appear to be pragmatic differences between the two forms):

(41) nukoo-ri' sīkī
    sit-1sg. back
    [sīkī
        back
        ydu
    rock
        face
    ]
    'I am sitting on a rock.'

(42) ni-kaa-ri'sīkī-bē?ē
    perfv.-be-1sg. back-house
    'I was on the roof of the house.'

(42) illustrates an important subsidiary point. It is clear from this sentence and others like it that the Mixtecs do not impose a single, unified model of body configuration on the object named in the locative expression. A house can be given relative subparts on analogy
with a human body model, resulting in yafa 'back' expressing "behind", and, by implication, a front/back assignment; or, as in (42), it can be assigned subparts on analogy with an animal body, in which case siki 'back' designates, roughly, the roof. So the associations made with actual bodies are far from fixed; subpart relations are exploited to their fullest by being applied selectively rather than, all at once, in a kind of superimposition of an entire body model.

Another interesting case of this principle also points up the abstraction under conventions of use that can be a consequence of selective analogy. Notice (43):

(43) haʔa siki-riʔ  
    pass-pot. back-1sg  
    'Climb over me.'

(43) was uttered within a scenario in which the speaker was lying down and the (imagined) addressee was lying next to him at the time of the utterance. The use of siki here is regular in many respects: it expresses a relational use of the body-part, referring to an area in space directly contiguous with the location named by the body-part term, and the spatial information encoded into the verb allows the interpretation of the locative expression as a path. What is interesting about (43) is that we are reinterpreting a human body, based on its configuration and position in the context of the utterance, relative to the canonical configuration of an animal body. If the language were merely using the appropriate body-part term to refer to an area in space closest to that named body-part, we would hardly expect siki to be grammatical here, since siki by definition cannot name part of a person. This indicates again that a level of grammaticization is entering into the use of siki, by which it is coming to have a meaning independent of the appropriateness of referring to an actual part of the object with siki. At the same time, it preserves attention to the particular shape taken by the object, since it is that shape, of which the animal is recognized to be an instance, which allows the use of the locative term.
Because šini names the topmost extent of an object perceived as upright, it will overlap in function with the other terms šīki and nīfū, both of which name top surfaces that are horizontal and planar. The usual subpart named as the šini of an object is something like the "tip" of that object, i.e. a part that is smaller along all dimensions than is the object as a whole (by contrast with šīki, which can be identical in two dimensions with the object as a whole). This general restriction is intuitively comfortable, based on the size and shape of a head relative to the whole body of which it is a part, and occurs in partitive constructions, notably (44) and (45), which go unanalyzed synchronically (and express not locations but objects) but which are nonetheless consistent with the other data:

(44) šini-ha?a
    head-foot
    'toe'

(45) šini-n-d-a?a
    head-hand
    'finger'

These uses of šini are clearly partitive, expressing objects rather than mere locations.

Another subpart use of šini occurs as a locative element in (46):

(46) hiyad-ē šini-pydku
    be+loc.-3sg. head-hill
    'He is on top of the hill.'

This contrasts with (39), where šīki is required, in conformance with the above restriction on the shape of the subpart:
a. (39) ni-nddkoo-ri'  
    *siki  
    back  
    *sini  
    *head  
    -hika-śde?e  
    -wall-house  

'I sat down on top of the wall of the house.'

When used in relational senses, *śini can be glossed as 'above' or 'over', and can occur whether the location is stative or not. In any case, the shape of the locational object decides the grammaticality of *śini as it is used like 'above': recall that *siki, too, is used like 'over' when the object is horizontally oriented.

(47) ndesa hiyad yōd *śini-yunnu wā
    how be-loc. moon head-tree det.

'How is it that the moon is over the tree?'

Both the partitive and the prepositional uses of *śini can be used to express locational path. Again, ḫi 'road' placed before *śini often specifies path or disambiguates between a path reading and a goal reading, as in (48): whereas (49) gives a case wherein *śini is clearly a goal:

(48) ḫi?i'- ri' ḫi-*śini-ydku
    go+pot.-1sg. road-head-mountain

'I'm going along the top of the mountain.'

(49) ni - kaa - ri' *śini-ydku iku
    perfv.-climb-1sg. head-tree yesterday

'Yesterday I climbed to the top of the tree.'

(50) is predictably ambiguous between the path and goal interpretations of the locative:
(50) ni-n덕을 ?nn.ṣa nga śini-ỵdu

perfv-fly one bird head-tree

a. 'A bird flew over the tree.'

b. 'A bird flew to the top of the tree.'

The lack of information about the direction of motion, as well as other characteristics of the semantics of the verb ni-n دق, allows for the ambiguity of (50). (49) and (50) demonstrate that ܝMari can encode the location of an activity verb without requiring ܕMari.

Mari 'face'

The word for 'face' in Mixtec is by far the most interesting of the body-part terms. It exhibits the largest number of clearly distinguishable uses, and in some of those uses, the greatest degree of abstraction from the basic meaning. These abstract uses can be classed as grammatical rather than lexical uses, and this grammaticization is a result of extensions from the basic meaning which are highly conventionalized.

Mari is the term whose extensions depend most heavily on an understanding of the application of canonical spatial orientations and their place in more general situations. For example, in (12), repeated here, we can understand why Mari marks the dative goal if we take into account the usual situation of interactions--that the participants are face-to-face--and further extrapolate as to why this situation is the usual one--that the face is the part of the body by which people are most easily identified; that it is the location of most of the perceptual organs by which we gather information; that it is oriented in the same direction in which we are most agile in movement and locomotion.
(12) ni - xàd -ri' ?n kiti nàvi-se?e-ro
perfv.-pass.-1sg one horse face - son - 2sg
'I gave a horse to your son.'

Most situations which involve a goal, either dative or locative, will require the encoding of some or all of these cooccurring elements of the interactional scene, and apparently the way this is done in Mixtec is, in a manner completely consistent with the general phenomenon of using body-part terms, by naming the body-part which has all these experientially based characteristics associated with it.

The shape and relative location of a face are, of course, still different bases for extensions to locating expressions. In the case of nàvi, unlike the other body-part terms, one or both of these components may be disregarded when the extension is based on the situational characteristics instead (recall that a similar situation obtains for ha?a and nda?a, except that in each, one component is consistently disregarded) So, for instance, while a location described with stiki - noun must maintain both the shape of the stiki (i.e. a horizontal or sloped two-dimensional surface) and its position relative to the object as a whole (i.e. the top-most surface of the object), the nàvi of (51) does not refer either to a front-top surface or to a roughly planar surface:

(51) ni - ha?a - rí nuč - βe?e
perfv.-pass.-1sg face-house
'I went to his house.'

Looking at nàvi a little more systematically now, we can find several cases of the partitive type in which the shape and relative location components are preserved, as in (52):
(52) ni-nduko-o nu-t -yuku
perfv.-sit - 3sg. face-hill

'He sat down on the hill.'

'On' here is far too general a gloss to suggest the real distribution of this use. nu-t here
designates a subsurface of the hill which is closest to the speaker (or other deictic
anchoring point) and is therefore the front of the hill by convention. However, (53)
and (54) maintain only the roughly planar shape of the face, but not its location or size
relative to the entire body:

(52) nukoo -ri be+seated-1sg. \[\begin{array}{l}
\text{nūt} \\
\text{face} \\
\text{sīk?} \\
\text{back}
\end{array}\] -yūl

'I am sitting on a rock.'

(54) ni -hika -ri iči - nūt - ṃdu-ʁd
perfv.-walk-1sg. read-face -town-2sg.

'I walked around in your town.'

(55) ni -nde?e -o nūt-8alse
perfv.-look - 3sg. face-purse

'He looked all over the surface of the purse.'

These sentences show that nūt can refer to a stative location or the extended location of an
activity, as long as that area is two-dimensional.

By contrast, nūt can also be used in a sense which preserves the relative location of a
face, in the front of a body, but with no regard for shape or relative size. In these uses it can
be glossed 'in front of':
(56) kwa kündi nūʔi-mesā
go+pot. stand face-table

'Go stand in front of the table.'

In this use, there seems to be no restriction whatsoever on the shape of the object given in the locating expression: it can be upright, as a person, completely flat, as a rug, or ambiguously interpretable, as a table. The last case often leads to an ambiguity in the sentence, since nūʔi can be interpreted as referring to the planar top surface or to the area in space in front:

(57) yūʔ wā hiyaʔi-ci-nūʔi-mesā
rock det. be=loc. road-face-table

a. 'The rock is in front of the table.'
b. 'The rock is on the table.'

(I do not understand why the (b) reading is possible when ʔi is present. I would expect that ʔi would be used to disambiguate between the two readings, being present to indicate an implicit directionality and being absent in the case of a non-directional reading. This proves not to be the case here, although it does work this way in other sentences.)

Because of the conventional assignment of front/back orientation on spatial deixis principles, nūʔi has acquired a pair of meanings that have become quite grammaticized, meanings that would seem to be almost entirely devoid of any active principle of analogy based on properties of the face. These two uses are in the marking of sources and goals. Examples of dative and locational goals are given in (12) and (41) above, but both of those examples involve nouns which name objects considered to have fixed fronts (in the case of a house, the front is the wall in which the door is cut). (58) and (59) provide cases of a front surface that is completely dependent upon the situational context, as well as demonstrating how the language linguistically distinguishes between sources and goals:
(58) sa’wa ni-ndeći nuu-ydnu
bird det. perfv.-fly face-tree
'The bird flew to the tree.'

(59) sa’wa ni- kenda nuu-ydnu
bird det. perfv.-exit face-tree
'The bird flew out of the tree.'

As noted earlier, deictic information about directionality encoded into the verb gives the path of motion in (59) and allows nuu-ydnu to be interpreted as source. As in (58), when there is no information to the contrary, nùnu - noun is interpreted as source when the ambiguity is between source and goal. The contrast between (58) and (59) demonstrates a series of compromises made in the language. Using nùnu to indicate both source and goal is consonant both with the basis for extending body-part terms and with the assignment of front/back orientation based on the situation. But this double duty of nùnu leads to the possibility of serious ambiguities in particular instances. Having a lot of spatial information encoded into the verb, as it is in kenda, haʔa and other verbs of motion serves to disambiguate the meaning of nùnu, but at the expense of the expression of manner of motion, which is indicated by ndeći in (58) but is lost in favor of directional information given by kenda in (59).12

A similar set of compromises is demonstrated below. English distinguishes source from goal by means of two different prepositions, so understanding the semantic role of the locative element is independent of the deictic anchoring of the verb, so long as the information is consistent. For this reason, in English one can use both bring and take with either source or goal phrases, since the verbs differ only in deictic anchoring.

(60) I brought the bread from my uncle.
(81) I [brought \_ took ] the bread to my uncle.

The variability in deictic anchoring shown above for English is impossible in Mixtec, since it is deictic information in the verb that identifies the locating expression as either source or goal—the locating expressions themselves are marked the same in either case. Only the following versions of the possibilities given in (60) and (61) are possible in Mixtec:

(62) ni - hi - nda?d - ri'\_tastila nuu-stdo-ri'
    perfv.-with-hand-1sg. bread face-uncle-1sg.
    'I took the bread to my uncle.'

(63) ni - ki?\_ - ri tastila nuu - stdo - ri'
    perfv.-move-from+PLA-1sg bread face-uncle-1sg.
    'I took the bread from my uncle.'

Again, as we saw with nde\_ki and kenda, zinda?\_a 'with the hand' / 'carry' and ki?\_ move away from Place of Locutionary Act' (after Kuiper and Merrifield and Macaulay) encode different kinds of information: one gives manner, while other gives directionality, and the directionality given by ki?\_ requires the source to be specified, thereby allowing the locating expression to be interpreted as source.

While the face-to-face interaction scene exemplified in (12) is evidently the basis on which nu\_nu is used as a goal marker, synchronically nu\_nu marks abstract goals as well as goals of location or transfer:

(64) ni - s - na?a - ri' nuu - ad?d - ri' ba ad\_nu
    perfv.-cause- know-1sg. face-son-1sg. compl. work
    'I taught my son to work.'

The use of s\_ki in (43) indicates that these body-part terms were not being used purely
metaphorically and associatively every time, because the (human) object in (43) does not have a body-part named sti. Here is another graphic counterexample to the claim that body-part terms are pure nouns—purely lexical morphemes—which are used associatively. (65) is ambiguous over the two readings given, though (a) is the preferred reading:

(65) nuñ-haśiʔi ha pedr̥ ni - haʔa kuñ̥u ka kāʔnū

face-woman compl. Pedro perfv.-pass meat be fat

a. ‘The woman that Pedro gave the meat to is fat.’

b. ‘The face of the woman that Pedro gave the meat to is fat.’ 15

It seems to me that if the goal reading in simple clauses were merely a consequence of an extension from the face onto an associated area, placing nuni- has in subject position would result in a preferential reading of (b), if not a downright exclusion of (a). That is, the straight partitive reading should be preferred when the nominal compound occurs in subject position, if no conventionalization of the function of nuni is involved. The fact that the (a) reading is preferred suggests to me that, on the contrary, nuni + noun has achieved an easily recognized status as a goal phrase, and that in the sentence context, that is the most easily recoverable interpretation of the construction. On the other hand, the fact that the nuni gets pushed up front from its place in the lower clause suggests that Mixtec speakers would do something as strange as putting an apparent goal in subject position rather than break up what is obviously a very tightly-knit construction—as tightly-knit a construction as a clear case of a partitive or possessive nominal compound would be (note that a version of (65) in which the nuni were “stranded” before the copula of the higher clause would be ungrammatical—nuni and hašiʔi have to stay together).

To return to the more general issue of the distribution of nuni: it is used as a goal in such a variety of situations that it seems to have established itself as a marker of goals of all sorts. This wide application of a body-part term to such an abstract semantic role exhibits a level of grammaticization that has so far not arisen in these data. The body-part term nuni has lost virtually all of its function: as an indicator of a subarea of an object, since in the sentence above nuni marks the semantic role of the named object as a whole.
I have mentioned that when нун is used as ‘on’ it often refers to a two-dimensional space that can be either vertically or horizontally oriented. This seems to be the most neutral locative function in the Mixtec system, since it makes no restriction on orientation, shape of the object as a whole or of the subpart being referred to, or relative size. It is not the least marked means of expressing a locative, since one can place some location names directly after the verb without using a body-part term at all,14 e.g.:

(66) кити сквела
   go-pot.-3sg school
   ‘He will go to school.’

That нун is the most neutral locative expression with respect to the parameters discussed here is consistent with its use as source and goal, whereby it refers to a general place. Sources and goals usually have a salient surface—either horizontal, as in a locative goal like a city, or the front, vertical surface, as in a dative goal like a person, or both, like a location at a table. Thus it is easy to see how нун could become abstracted—less specified in terms of its semantic components—to designate simply the relevant location in an activity or state. In (67), this general meaning of ‘place’ or ‘location’ gets applied metaphorically:

(67) ни - сатн - ри - нук - моника
   perfv-work -1sg. face-Monica
   ‘I worked instead of Monica.’

Mr. Cortés informed us that this use of нун is highly restricted, occurring only with the word for ‘work’ and with semantically related words. This is understandable, since the space that нун ostensibly refers to is not even a physical space associated with Monica or some part of Monica, but is a position of activity and responsibility associated with Monica. Here нун literally refers not to a space at all, but rather to a position within a non-spatial domain—a further abstraction away from those occurring within the spatial domain.

265
'iči 'road'  

While not a body-part term, 'iči behaves enough like the body-part locatives and is frequent enough as an indicator of location to warrant some attention here. 'iči, when used in locating expressions, always occurs before a body-part term—never simply with a locational object. Using this fact and applying to 'iči the general principles of abstracting locational elements from the semantic components of body-parts, one would expect that the locative use of 'iči would exploit the implicit motion and directionality characteristic of roads, and that the use of 'iči would indicate a locational path or implied directionality to locational expressions. In the contrast between (68) and (69) 'iči behaves in this nice way:

(68) hindii-ri nuu-brid
stand-1sg. face-Maria

'I am standing in front of (facing) Maria.'

(69) hindii-ri iči-ncu-brid
stand-1sg. road-face-Maria

'I am standing ahead of Maria.'

The addition of 'iči in (69) gives a direction or implied path relating the speaker to Maria. The directionality of 'road' is, reasonably enough, the road ahead of Maria, and the speaker standing in that path is facing the direction of the path as well. (68) describes a simpler situation which instantiates the usual configuration of two people interacting (and resembles in that way the sentences with nuu marking goal). One can give several cases in which 'iči serves this predictable purpose of naming a path or linear extent associated with the point named by the body-part locative - Object. (55) is such a case: where the sentence without 'iči would probably be interpreted as including a goal phrase, inclusion of 'iči allows the locating expression to be interpreted as an extended area over which the activity of motion takes place.
However, $\tilde{\tilde{c}}i$ also occurs where one would not expect it to under this analysis. For instance in (57), the (a) reading indicates a nonpartitive space associated with the núñ (front surface) of the table. 'In front of' is a natural consequence. But the (b) reading of (57) is just as acceptable to the consultant, and the location there is purely stative and partitive with no implication of either a front-directed surface or of a nonpartitive space. The fact that the two readings are possible for this single sentence indicates that it is not simply that I have missed some feature of shape or position that requires or prevents the occurrence of $\tilde{\tilde{c}}i$. Mr. Cortés told me that he uses $\tilde{\tilde{c}}i$ often just to emphasize, further specify or make more precise the location being referred to in the body-part + noun compound, irrespective of whether the locative is stative or active or extended or otherwise oriented. This explanation works for some of the cases, but even if one could be satisfied with such a disjunctive definition of $\tilde{\tilde{c}}i$, it would not easily account for all the cases in which $\tilde{\tilde{c}}i$ changes the meaning of the locative and those in which it does not.

Some Issues Suggested by the Data

Up to this point I have been assuming that the locating expressions discussed here are derived from the nominal compounds used so productively in Mixtec to signify genitive relations. While I believe the data to have demonstrated this position, it may yet be subject to dispute. There are also other implications and difficulties of this position that require exposition.
Summary of the Principles of Extension from Body-Part to Locative

It might be possible to give an accurate representation of relational expressions without having to postulate a synchronic or diachronic relationship between the obviously nounlike body-part terms and the clearly relational locative expressions. It would require something like strict subcategorization and selectional restrictions, in order, for instance, to prevent (39b) from being predicted as grammatical, which it is not:

\[
\begin{align*}
\text{a. } & \text{ni-ndukco-rf} \\
& \text{perfv.-sit-1sg} \\
\text{b. } & \text{si}k\text{f} \\
& \text{back} \\
& \text{-hika-βe?e} \\
& \text{*sini} \\
& \text{-wall-house} \\
& \text{head}
\end{align*}
\]

'I sat down on top of the wall of the house.'

Furthermore, building into the lexicon the constraints for particular senses of particular locational objects would be extremely complicated. For instance, \(\text{nūnū} \) in the sense of 'in front of' can take an object of any shape, but in the sense of 'all over' (cf. (55)) requires a horizontal, planar object, and in the sense of 'from' also requires verbs of exiting.

This is information which can be coded more easily and in an intuitively more pleasing way by positing the body-part as the basic sense, explicating principles of canonical orientation and deixis that are independently required in the language, and determining which semantic components of the basic body-part terms serve as bases of extension to more abstract uses.

Another advantage to positing body-part meanings as the bases for the relational uses is that the principle of extension from the body-part meanings accounts for both partitive and relational uses of the terms, which cannot be given a consistent or unified treatment if there is no motivation for a synchronic relationship among the senses. If the body-part terms were not taken to be the basic senses from which are derived both partitive and relational locatives there would be no justification for calling the relational elements "nominal" to any
extent and we would lose the availability of the already-existing nominal compound to characterize the relational uses of the construction.

There are several other advantages to viewing the body-part terms as semantically basic and the locational senses as historically derived and synchronically related. The most general advantage is that of economy: by this analysis one can substantially decrease the number of rules and constraints on rules in the grammar required for descriptive adequacy. Rather than having just lexically-specified constraints as suggested above, we can give general principles in which all the body-part terms behave in a similar fashion, when extended and abstracted, after which we can specify the boundaries of extension peculiar to each term. We can thus provide a unified account of the literal uses of the body-part term, the partitive uses by which the subareas of inanimate objects are referred to, and the relational uses by which areas in space are referred to, thereby capturing an important generalization about the systems in the language.

Body-part terms as they are used literally provide most of the information that allows for the abstract uses of the terms. Since canonical body structure and orientation are such important notions in the extended uses of these terms, it is necessary to recognize that body-parts themselves are relational in an important way. For example, head cannot be fully understood without making reference to an entire body of which it is an integral part. Similarly, part of what we know about a foot—aside from its characteristic shape—is where it belongs on the body. Shape of the object as a whole, shape of the subparts referred to, and location of the subpart all serve as semantic components by which objects are perceived as resembling the body-part whose name is used in the reference.

Nevertheless, each body-part term exploits its possible range of extensions to a different degree. haʔa ‘foot’, for example, extends both to partitive and to relational uses almost exclusively on basis of the relative location aspect of its literal meaning. By contrast, the correct use of šini for partitive and relational uses relies heavily on the shape and size of the subpart as well as its relative location.

Sometimes the extension on basis of one semantic component requires that another semantic component of the body-part term is not preserved. For instance, naʔi ‘face’ is
regularly used to refer to various locations with respect to a planar surface ('on', or 'all over', etc. Cf. (55)–(57)). This usually requires that the surface referred to with ܢܘܐ is not toward the top of the front surface of an upright object, as a face is with respect to a body (cf. (58)). When ܢܘܐ is used in the sense of 'in front of', canonical body position, i.e. uprightness, and part of the relative position, i.e. frontness, are preserved, but relative size is not.

A different parameter of extension along which these body-part terms exhibit individual differences is the kinds of locations they can be used to express. All of them exhibit some uses as partitive locatives, e.g. (17), (28), (42), (48), (65b). In these cases the nominal compound functions as usual by referring to a part of an object and then to the object as a whole. I have also given some examples where the nominal compound does not refer to the expected partitive relation—what I have referred to as the "relational" uses. They are semantically related to the partitive uses, but some ambiguities, such as that below, show that they must be distinguished.

(50) ܢܝ-ܢܘܐ?ܐܢ ܐܕܕ ܢܐܝ-ܝܢ\\nperfv.-fly one bird head-tree
a. 'The bird flew over the tree.'
b. 'The bird flew to the top of the tree.'

(50) gives an example in which the relational meaning is easily recovered from a partitive meaning. Given the compound 'head-tree' and the partitive meaning 'top of the tree', we can easily make the extension from the top part of the tree, the partitive meaning, to the relational meaning—'above the tree'. Below I will summarize some of the cases where one cannot so easily recover a relational reading of the locative from just observing the partitive use and its associated space.

But (50) exhibits another point important for describing the range of distinguishable meanings of the nominal compounds used as locatives. Notice that the ambiguity of (50) between a partitive and a relational reading of the locative is contingent upon the fact that either activity goal or locational path can occur directly after the verb. Of course, this depends greatly also on contextual information, since many verbs require either source or
goal to be specified, in which case there is no possible ambiguity. Here, the semantics of \textit{fly}
allows either a goal or a path interpretation for the locating expression, and that semantic
role will determine whether it is taken as a partitive or a relational locative.\textsuperscript{15}

\textbf{Shift in Status: The Move from Lexical to Grammatical}

Most people would agree that the process of extending a body-part term to apply to a
corresponding subpart of an inanimate object involves a productive process of recognizing
and exploiting perceived similarities between the body and the object named. Most would
also agree that this is a creative process rather than a simple one of correct semantic com-
bination, and that even in the event that the extended use of the term becomes fixed, as it
does in English with the \textit{leg} of \textit{pants} \textit{leg}, that word remains a lexical morpheme, though prob-
ably a synchronically polysemous one.

A grammatical morpheme, on the other hand, is defined as being a member of an essen-
tially closed class whose function is not to provide semantic information but rather to give
information about how the lexical morphemes in a sentence are to be understood as relating
to one another. Under this definition, one could consider English prepositions as being in the
class of grammatical morphemes. Prepositions serve to mark semantic or case relations like
instrument, source, goal, location, accompaniment, and so on. Some further specify these
general relationships, including some lexical information of their own (compare English \textit{on}
and \textit{in}); but at least a subset of them can be considered straightforward grammatical mor-
phemes. The same could be said for some uses of Mixtec body-part terms. Given general
meanings of certain of them, as, for instance, marking location or goal, one could consider
them to be "function words"—grammatical morphemes. Their basic relational meanings
paired with the fact that the expressions are not referential in the way nouns usually are,
leads to a conclusion that they are not in fact totally nominal elements and therefore abso-
lutely lexical morphemes.

An example of this is in sentence (43), in which, as noted before, \textit{sita} cannot possibly be
being used referentially.

(43) ha?a sîkê-ri
    pass+pot. back-1sg.
    'Climb over me.'

(43) demonstrates that sîkê is first extended in the expected way from animal backs to roughly horizontal, planar surfaces of objects, and then gets extended further to the space associated with the planar surface, what we would translate as ‘over’. This relational use relies upon the already abstract partitive use, rather than being a projection directly from the animal back to the associated space.

A clearer case of an obviously grammatical morpheme deriving from an obviously lexical morpheme is when nûû marks a goal. This is least controversial as a marker rather than a partitive locational, since it marks energy goals that are not always goals of direction of motion, as in (70):

(70) ina wû nde?e nû -sê?e wû
    dog det. look face-house det.
    'The dog is looking at the house.'

(64) ni- s - na?â- ri' nûû - âdê - ri' ha sattû
    perfv.-cause-know-1sg. face-son-1sg. compl. work
    'I taught my son to work.'

These are cases of marking a semantic role rather than anything faintly resembling a location, so it is a departure from those uses more readily called lexical. Nevertheless, given assignment of front/back orientation done regularly in the language, and the normal configurations of participants in these activities, we can easily understand what motivates the use here of the word for ‘face’. Energy or activity is directed toward the object; the goal Object either has an inherent front, canonically closest to the source, or is assigned as front
that surface which is closest to the source, so that the surface of a goal object nearest to the source of the energy is, one way or another, its front. Faces are a highly salient part of the front surface of people, besides being important for other reasons. Therefore it is sensible to mark a goal whose front surface marks the endpoint of the activity or energy with the body-part term (namely, 'face') that invokes both the appropriate relative location in front/back orientation and other important information about the interactional setting as a whole.

The point of this discussion is not to place particular uses of body-part terms on one side or the other of the lexical/grammatical split, since it is clear that such a dichotomy would run counter to the very processes by which Mixtec speakers encode spatial (as well as more abstract) relationships. The fact that convention plays a big role in determining exactly how body-part terms get abstracted to locatives implies that there were semantic components in the first place to be conventionalized, which itself makes a lexical/grammatical dichotomy arbitrary. It is just as important to recognize that conventionalizations of use are there in the abstract sense, which amounts to saying that there are uses of the terms on the pole that we would call "grammatical". The system shows that there is an entire continuum existing synchronically in the language and that on the level of the behavior of individual terms there are varying degrees of movement toward the grammatical end of the continuum.

The fact that the consultant does not recognize a body-part reference of ūnū 'in', while as late as 1974 Pensinger lists both body-part and relational uses, indicates that for youngest speakers of the language, the movement is fairly far advanced for that lexical item. It is unlikely that all of the terms will lose their literal meanings, but ūnū represents the furthest stage of the same process of grammaticization that seems to be taking place in all the terms, at least in the idiolect of the consultant.
Problems with Lexical Categorization

On the lexicosyntactic level the same point about category shift needs to be made. As far as I have been able to determine, previous references to this phenomenon in the Mixtec literature (e.g. Dyk and Stoudt (1965), Pensinger (1974), Bradley (1970), Reyes (1889), Alvarado (1963)) have categorized these terms simply as nouns. In a representative quotation, Dyk and Stoudt state:

Los conceptos de tiempo, posición o localidad se expresan comúnmente en Español por frases preposicionales. En Mixteco estos conceptos se expresan por un tipo de frase en que unos nombres, principalmente de las partes de cuerpo humano, se untan con otros nombres para midificarles para expresar estos conceptos.

None of the descriptions I have found have made reference to the range of extensibility shown here. The examples contained therein make it impossible for the reader to realize that this is a problematical characterization, though throughout, the functions of the body-part "nouns" are compared with English or Spanish prepositions.

Daly (1973) recognized the distinction between what I have called "partitive" and "relational" uses, but in his generative grammar he distinguishes them only formally, simply calling the first element in what I have called the "partitive" type of construction the "head," so that the meaning of the first noun is referential when that noun is the head, and abstract when it is not. But since he gives no independent criteria for defining "head," this is a circular definition. He also confuses the issue by classifying formally the body-part terms with the feature \[ +loc \] (\{ +loc \} in more usual notation). This forces him into one of two implausible positions: on one hand he might say that the body-part terms used as locatives are not in the same lexical class as those not used as locatives (while I am arguing that such a distinction is artificial, despite there being clear cases of each). On the other, he might say that a body-part term was the same word whether or not it is used in a locative phrase, in which case he would also be saying that on both readings of (65), \[ nuw-hast' \] is a locative, even though its...
(b) reading is purely literal and is grammatical Subject. And like the others, Daly seems to consider it an unproblematic task to classify these elements as nouns, despite the opposite being indicated by the (a) reading of (65). In short, Daly's formal classification does nothing to suggest all the interesting issues involved in the system.

Throughout the discussions cited here, descriptions and examples have been such as to obscure the complexities and variations in usage. They have especially omitted any reference to the most grammaticized or abstracted uses of these terms such as the use of yata for 'outside'. All examples are given with glosses exemplifying only the partitive senses or relational senses which are the clearest cases of naming associated space from actual partitive use—the most recoverable of the range of uses. In the earliest of these studies, such shortcomings are probably attributable to an understandable lack of linguistic expertise, as well as to an actual difference in the data: it being nearly 400 years ago, it is likely that the most abstract of the uses found in contemporary speech were not in use then. However, four of the grammars and dictionaries I looked at were written within the last twenty years, which is recent enough that at least some of the more abstract uses must have been appearing. In fact, Persinger, in her dictionary, gives prepositions as glosses for the relational uses, but in her grammatical sketch gives mostly examples of partitive uses, and does not relate her two descriptions in any way, even as much as to mention whether she considers any pair of forms a case of homonymy or polysemy. Needless to say, none of these studies mentions the fact that there is variation across the set of terms in exactly how abstract the meaning can get.

The omission of these details is excusable, given the fact that in every case the description of locative expressions was embedded in an all-encompassing and very general description of the language as a whole. But the theoretical conclusions made in each of these cases seem drawn on the very limited range of data mentioned, rather than on the enormously varying range that these linguists must have encountered. I think it is misleading simply to label body-part terms as nouns and mention, as in the quotation above, that Mixtecs use nouns where Spanish- or English- speakers would use prepositions.

There are many justifications for such conclusions. As mentioned, these locative expressions function syntactically as nouns in both their internal structure and their
syntagmatic relationships within a sentence. Syntactically, the body-part terms, which in many cases are clearly recognizable as nouns (being obviously referential), always occur in the nominal compound. This construction is characterized semantically as involving a genitive relationship between the two nouns, and while it is clear that the "genitive" characterization does not quite work for relational meanings of the construction, it is similar enough to a true genitive relationship to warrant our taking the semantic and syntactic resemblance to clear nominal compounds as significant.

What then are arguments against unequivocally calling body-part locatives full-fledged nouns? My major argument against it has been the semantic one. There are very clear cases in which the relationship of the relational usage to the partitive one is perfectly transparent: the relational meaning is taken just from interpreting the reference of the body-part term as an area that preserves at least some of the locational components that the body-part referent has relative to the object.

Less clearly associational readings, however, are much less nounlike in their semantics, since they do not refer to a subpart of an object. Taking again the example of nuku designating source or goal in the environment of the appropriate verb, it is not merely that the conventional front surface of the goal or source is named, but that that naming process is exploited in order to specify the semantic role being played by the object. The process of abstraction occurring in this phenomenon as a whole goes in several different directions. In the case of nuku, where what is marked is not necessarily the goal of a physical path (cf. (63)), the abstraction seems to have extended beyond the spatial domain. The least transparently metaphorical uses of yata and stiki (those glossed as 'outside' and 'over', respectively) are still within the spatial domain despite the fact that both have gone fairly far afield of the most obvious partitive meaning. Both have achieved an abstraction from the body model to a more general physical model which then provides the basis for further extension.

Yet I think previous investigators of this phenomenon in Mixtec erred on the other side of the noun/preposition classification dilemma by simply equating these "nouns" with Spanish or English prepositions. Characterizing the semantics of these terms as prepositional while classifying them lexically as nouns hides the fact that these terms express the
relations between objects and between activities and objects in a way that is fundamentally different, in many cases, from the way it is done with prepositions. While in English, for instance, the grammaticality of a preposition may depend upon the shape of the prepositional Object (cf. *'Put the dishes in the table'), as it does in Mixtec for body-part locatives, English prepositions give details of the nature of the relationships between objects in a way that Mixtec locatives do not. Both systems can describe both "partitive" and "relational" kinds of locations in a broad sense of those terms; but as I have argued that in Mixtec the partitive senses are basic and the relational ones are derived,\textsuperscript{16} so in English the basic uses of prepositional phrases are relational though one can get a subpart reading from them.

It is oversimplifying the issue somewhat to say that English prepositions contain relational information while Mixtec body-part terms do not. But when given in isolation, and English prepositional phrase is pretty much unambiguous over a partitive or a relational interpretation, while a Mixtec locative is ambiguous:\textsuperscript{17}

(71) sikt\textasciitilde mes\texttilde

back - table

(72) a. on the table
    b. over the table

(71) gives an isolated nominal compound potentially interpretable as a locative element, two English glosses of which are given in (72).\textsuperscript{18} Notice that (71), when given without the usual information occurring in a full sentence to disambiguate it, could be given either interpretation in (72) and at least one other distinct interpretation ("top of the table"). The fact that in previous descriptions (72a) was the only gloss given does not mean necessarily that that is the most basic interpretation of (71), especially given the level of generality with which this phenomenon has been described.

By contrast, the English prepositional phrases are unambiguous in their reference. (72a) unambiguously corresponds to a Mixtec partitive reading (since it designates a location which is properly a part of the named object, the table). (72b) most readily corresponds to the Mixtec relational interpretation since it refers to an area in space not part of the named
object.

If this demonstration does point to a substantial difference in the way Mixtec and English speakers express such locational concepts, it is still the case that both languages can do both kinds. This paper has been devoted to showing that Mixtec can do both, and one or two examples should be sufficient to suggest the same conclusion about English to a native speaker:

(73) We flew over the hill.
(74) He lives over the hill.

Without attending here to details like pathhood versus stative location, we can see that (74) refers to a location by means of the relation called "over"--exploiting relations to name a location (see Brugman (1981)).

This is not to say that Mixtec and English work in exactly the same way, however, merely with the directionality reversed, for I think that in detail the two kinds of processes are rather dissimilar. The point here is just that while Mixtec and English can both describe a large range of locational configurations, the English system works basically by coding spatial relationships in prepositions. Mixtec, however, makes the spatial relationships inferable from the semantics of the verb, while the locative expression names only the relevant area in or around the locational object. So Mixtec locatives are not merely prepositions.

As mentioned above (fn. 6), it may be that there was no category of preposition in earlier stages of the language, since all morphemes serving preposition-like functions seem to be historically derived from other words (though not all of them body part words--cf. fn. 6). Possibly under the influence of European languages, and for whatever reason motivates Greenberg's universal, the process of grammaticization of these morphemes is such that once obviously full nouns in a particular construction became increasingly abstracted in meaning as their ranges of grammaticality broadened correspondingly. These considerations taken together make tenuous the position taken by those who have described these Mixtec data previously: the status of body-part locatives either as nouns--lexically--or as
prepositions—functionally—is subject to question:

**Ross’ Proposal: Nominal Decay**

In various publications (e.g., Ross 1973), Ross has proposed that lexical categories such as "noun" and "auxiliary verb" are not discrete categories, but are fuzzy-bordered; in his terms, "squishy". In these papers he has given substantial syntactic evidence from English for his claim that nominals can be "noumy" just to the extent that they allow certain syntactic rules, characteristic of nouns, to apply. In "Nouniness" (1973a), he gives a continuum of nounhood for English from *that*-clauses to monolexemic nominals, the former (e.g. *that Mary gave the letter to Frieda*) being the least noumy, and the latter (e.g. *spatula*) being most noumy. While the particulars are not relevant for present purposes, the general point of these papers was to demonstrate degrees of nounhood, an issue which it seems important to consider here.

A later paper of Ross' (1981) bears some interesting similarities to these; these Mixtec data. It is most generally concerned with the syntactic properties of "fake NP's" (after Ross (1973b)--) the toe of "stub one's toe," for example. But he remarks in passing that he considers the fake-NP facts to be a special case of a more general phenomenon of "ego-loss" among nouns that appear often in conventionalized phrases. The following examples from that paper demonstrate this point, and resemble the Mixtec facts insofar as they too suggest a diachronic shift from nounhood to prepositionhood:

\[
(75) \quad \begin{cases} 
\text{on the top of} \\
\text{or top of} \\
\text{on top} \\
\text{on top of} \\
\text{atop} \\
\text{the box}
\end{cases}
\]
(76) It is
\[
\begin{align*}
&\text{in the front of} \\
&\text{in front of} \\
&\text{*in front} \\
&\text{*afront}
\end{align*}
\]
the box.

(76) It is grammatical with the meaning of \textit{before the box}

(77) It is
\[
\begin{align*}
&\text{in the side of} \\
&\text{*in side of} \\
&\text{inside of} \\
&\text{inside}
\end{align*}
\]
the box.

(77) It is grammatical with the meaning of \textit{within the box}

(78) It is
\[
\begin{align*}
&\text{by the side of} \\
&\text{*by side of} \\
&\text{*beside of} \\
&\text{beside}
\end{align*}
\]
the box.

(78) It is grammatical with the desired meaning of \textit{next to the box}

What Ross refers to as "ego-loss" is the loss of the usual characteristics of nouns: unbound-edness of the morpheme, presence of an article, and presence of the following of (which, it should be noted, marks genitive relations in English). In some cases the nominals are almost fully nounlike (as in \textit{on top of}); in others fully prepositionlike (as in \textit{beside}); and in other cases there is variability in how nounlike or prepositionlike the locative expression can be. These seem to be relics, frozen in various stages of the process, of some historical shift from noun to preposition.

The syntactic cues for loss of nounhood in English have no correspondents in Mixtec. As we have seen, body-part locatives behave syntactically like full nouns. In fact the change itself seems to depend upon their exploiting the syntactic environments of nouns, since without the nominal compound construction it could not take place at all.

My arguments for the reduced nounhood of the body-part terms are three: 1) the wide variation among individual terms as to the range of syntactic and semantic roles in which they can appear unaided (i.e. without \textit{to} to mark, for instance, locational paths); this fact should also answer any claims that there may be no real conceptual or semantic distinction.
between partitive and relational uses of the terms in Mixtec, since if that were the case one should be able to find broad relational uses for all the terms, which I cannot; 2) that the terms do not always refer to the object they name, as when \u00e1nti 'head' is used to mean "space above"; 3) the two clear cases from Mr. Cortés where he has only "prepositional" or "relational" uses for terms which other studies indicate are derived historically from body-part terms.

Although Ross does not mention this fact outright, it is clear that only in the least decayed of the sets of locating expressions given in (75)–(78) above is the full prepositional phrase (truth-conditionally) synonymous with the decayed version, the one without the article (the contrast in (75)–and there, there seems to be a slight semantic difference). In all other cases the full prepositional phrase designates an area different from the one designated by the phrase which has lost its article (e.g. in the front of the box vs. in front of the box). English uses the reduced forms of locative prepositional phrases to give "relational" locative meanings, as opposed to the partitive meanings given in the full prepositional phrases (with the exception of the first version of (78), where the semantics of by prevents a partitive reading). Ross distinguishes the degree of ego-loss achieved by the examples in (75)–(78) from those in which the once-referential and free morphemes appear synchronically only as bound forms, e.g. beneath and underneath, behind, between. Even the form in his examples which has undergone the most decay, beside, still has within it a morpheme which is easily recognized as being related, if not identical, to the free morpheme side. And in the best case, on the top of, we can recognize top as referring to an actual subpart even in the most decayed version, atop.

Ross' short-term aim in discussing these phenomena was to correlate loss of pluralizability in fake NP's with the loss of the definite articles in these locating expressions and to demonstrate how variable the degree of loss of nouniness is at a particular stage in the history of a language. His fortuitous use of quasi-nominal locating expressions in English suggests that loss of nounhood among names of physical objects—in terms of their syntactic signals—correlates with loss of referentiality to the designated object, which in turn correlates with increased levels of abstraction and increase in grammaticization.

281
The totality of the Mixtec data coupled with Ross' observations about English and Greenberg's stated universal suggest to me that these Mixtec body-part terms in their partitive uses are completely nounlike, while in their relational uses they are in various stages of "decay". For Mr. Cortés' idiolect, in the cases in which words have completely lost their ability to refer to objects, one could argue that synchronically these terms count as prepositions, while still recognizing the degree of nounhood that the other words exhibit.

I feel comfortable calling "nouns" the body-part terms that I have discussed in detail here, doing so with the understanding that they are not fully nounlike and that they are undergoing a process of abstraction and grammaticization. Even in a synchronic analysis this historical process must be recognized, since simply to lump these words unproblematically into the category "noun" without attention to the historical process suggested here would be to ignore the intricacies of the system and would give no insight about the probable future direction of semantic shift.

**Motivation and Prediction**

Linguistic theories have traditionally relied on the assumption that there is a dichotomy between complete arbitrariness and complete predictability. For instance, the semantic association between the phonological string [kʰaet] and the concept CAT is a completely arbitrary one. Nothing within the concept itself predicts what the phonological shape of the word representing that concept should be. The converse is also true. This is an example of complete arbitrariness. On the other hand, the meaning of the expression "The cat is on the mat" is completely predictable from the meanings of cat, mat, the, is, and on. This kind of association between semantic and phonological chunks is completely predictable. It is from the presence of cases of these two types that it has been common to bifurcate linguistics. However, among others, Fillmore (1982), Lakoff (1980), Langacker (1982), and Haiman (1980) have argued on various grounds and for quite different realms of language that the apparent dichotomy represents merely the two endpoints of a continuous scale. Intervening cases on that scale include linguistic structures whose meanings are fully or partially predictable.
from the meanings of the component parts, but for which there is still some linguistic convention involved in the structure.

The Mixtec locative phenomena represent another area on this scale. These constructions are highly motivated from extralinguistic convention about orientation and directionality and by principles of extension of sense already existing in the language. Semantic components of the basic, literal meanings of the terms always serve as the sources for extension of the usage of the term. In that respect, the abstract locatives are motivated by the literal meanings of the terms. But we have also seen that the extension on basis of one semantic element or another is a result of some conventionalization.

As the semantic components are more abstract, they require greater degrees of conventionalization in order to be understood in a nonliteral use. The case of nūy in particular has both the greatest number of distinct senses and the highest degree of abstractness in several of its distinct uses, since it is extended on basis of several (mutually conflicting) semantic components of the face. In these cases the meaning has become highly conventionalized (as when nūy marks sources of highly abstract as well as highly physical sorts) and in places highly context-dependent (as when nūy can mean “instead of”).

The degree of motivatedness and conventionalization involved in an individual locative use of a body-part term varies highly across the individual cases, as we have seen. Only a theoretical stance which allows a continuous scale from full arbitrariness to complete predictability can even address the question of where each use of a term occurs on that scale. Thus the idea of motivated convention makes possible the synchronic and diachronic claim of relatedness between literal and extended uses of body-part terms. It also justifies the claim that some of the extended uses are more easily recognized as “metaphorical” extensions of the basic senses that do others.
Acknowledgements

Many thanks go to Learne Hinton, George Lakoff and Wallace Chafe for their comments on earlier versions of this paper, and to Mariscela Amador, Nick Faracas, Charles Fillmore, Susan Lindner, Martha Macri, John R. Ross, and especially to Monica Macaulay for their help at various stages of the development of this work. None of them is responsible for the opinions advanced here. Of course none of the work could have been done without Mr. Nicolas Cortés, who served as the group’s linguistic consultant.

Footnotes

1For detailed expositions of similar phenomena in Zapotec and Tarascan, see MacLaury (1976) and Friedrich (1969).

2Mixtec is an Otomanguean language comprised of over thirty mutually-indistinguishable dialects. It is spoken largely in the state of Oaxaca. Chalcatongo Mixtec, the language on which the present study is based, is spoken in the village of Chalcatongo in the Oaxaca highlands.

This study is based on elicitations from a single, young, and strongly bilingual (in Spanish) consultant, and is subject to the restrictions that naturally result from such a limited corpus; so while my description of this phenomenon should characterize the present state of a process which developed over time, this description is in its details basically synchronic and descriptive of an idiolect more than of the language or dialect as a whole. There can be no discussion of individual speakers’ variations that would suggest diachronic processes. Likewise the comparison of Chalcatongo Mixtec with other dialects of the language will be minimal, since the literature on this topic is so limited.
There are interesting cases where both nominals in the construction are body-part terms, as in examples (44) and (45). In those cases the second body-part is the object of which the first term can be understood as naming a subpart.

Macrau also mentions (1981 and this volume) the possibility that an earlier stage of Mixteco included proclitic, nominal classifier-like elements which were derived from full nouns. While her data do allow that interpretation, thereby providing a set of nominal compounds that do not fit my definition of "genitive", I am deliberately excluding them from consideration for now. There is some justification for this, since synchronically they are far less analyzable as compounds than the ones I am studying, and are synchronically quite idiomatic in their meaning (see Daly (1973) on this point, and Macrau's criticism). Nevertheless one would still have to construct diachronically-explanatory definitions of "genitive" and of "nominal", since the "classifier" proclitics have fewer of the characteristics of nouns than do the already borderline ones I discuss.

5 All transcription of Mixteco phrases are written phonemically in APA, except for the tones, which are recorded here as originally transcribed (there are complex morphophonemic processes in Mixteco which are beyond my understanding, and which should account for apparent inconsistencies in tone marking across several occurrences of the same word. See Faro, this volume). The three postulated tones of this dialect are transcribed as follows: high '!', low '!', with mid tone left unmarked. Chalcatongo Mixteco distinguishes two and sometimes three stem forms of verbs. All of the three arise here, but one of them, the "hortatory" mood stem, occurs only once, and is so marked in the gloss. The other two occur much more frequently. Of these I have marked "pot." the "potential" form, the stem used for future tense and imperative and subjunctive moods. The "realized" stem form, used in past and present tense meanings, is unmarked in my glosses.

Sometimes Mixteco monolexemes have had to be glossed with more than one word. An example is an exposition of stem form as described above, e.g. 'go-pot'. In such cases the plus sign indicates that both words of the gloss refer to the single Mixteco word. Hyphens are used both in the Mixteco and in the English gloss to mark the bound morphemes which form a single word, e.g. ni-ha?e 'perfv.-give'. Spaces indicate word boundaries.
It is possible that in Mixtec there are actually no prepositions that are not derived historically from lexical morphemes. Locative elements which are not recognized by my consultant as body-part terms, i.e. ini ‘in’ and hii ‘with’, seem to be derived from ‘heart’ /‘center of the emotions’ and ‘side’, respectively, according to Pensinger (1973) and Hills and Merrifield (1974). The other preposition-like element I have discovered, and the only one seemingly not derived from body-part terms, is haku ‘for (benefactive)’. But it does not have the canonical form of monolexemic words in the language, since almost invariably the two vowels in a monolexemic word are identical. My suspicion is that haku derives historically from ha + ku, complementizer + copula, to mean something like "such that it is", and that there are no primary prepositions in Mixtec.

MacLaury (1976) claims that the Zapotec conceive of their body parts as locations rather than as entities, so that no transition is involved between naming body parts and naming locations outside the body. I have not determined whether the same conceptualizations hold for the Mixtecs.

The use of this term reflects less linguistic chauvinism than terminological desperation. MacLaury (1976) uses "direct" and "positional" to make roughly the same distinction, but I have not found these to be mnemonic terms. I beg the reader’s indulgence in this matter and ask him to translate “relational” as “with an ‘area-in-space’ interpretation”, until I can find a better set of terms.

Although I strive for my analysis to represent conceptual realities, I do not have enough support for a claim that Mixtec speakers conceptually make this distinction. There are cases later where the high level of grammaticization—conventionalization of use—would suggest the necessity for such a distinction in their minds; but there may be subtle constraints on the use of these conventionalized meanings of the terms which elude me.

Not all languages perform the assignment in this way: see e.g. Hill (1975) for a discussion of other language-specific conventions of orientation and some variations among individual users.

It should be noticed here that there are three nominals in sequence in this sentence. This is an instance of the nominal compound described before, except that the construction
is recursive—theoretically, one can embed an indefinite number of possessives in such a compound. Here the sequence is loosely translated as “road of the face of the mountain”. The construction is not confined to body-part terms or to locating expressions as it is in this example: note also kii-teki-wiri ‘horse-brother-lsg’, “my brother’s horse”. Any of the details noted for two-noun compounds including locative elements apply to more deeply embedded compounds as well.

12Hills and Merrifield (1975) present a slightly different analysis of these verb + noun - noun sequences for Ayutla Mixtec which may be applicable to the Chalcotongo dialect as well. They claim that deictic information doesn’t completely disambiguate the interpretation of noun in such sentences as (59); rather, the verb forces a preference for one semantic role (in their examples, goal), but that in the appropriate context the locative expression could be interpreted the other way.

13George Lakoff has pointed out to me that this sentence is interesting for two reasons not within the scope of this paper: in the (a) reading, the Subject of the higher clause is marked for the semantic role it fills in the clause it was complementized out of. That is, the woman who is Subject of be fat is the dative goal of Pedro gave the meat... But the noun in Subject position is marked as goal of the lower clause. This poses problems for a claim that the higher Subject exists in a deep structure and the goal phrase in the lower clause is deleted under identity with it. The nominal compound in Subject position cannot have been in a deep structure, since goals cannot be Subjects. The (b) reading of this sentence is interesting because the Subject of the higher clause is face, but the noun deleted under identity in the lower clause is identical not with face but with the modifier of face—that is, woman, in apparent violation of the spirit of the Left Branch Constraint (Ross 1967).

14Kuiper and Merrifield (1974) claim for Duixi Mixtec that each verb chooses whether to mark its goal with noun or to leave it morphologically unmarked. This is not the case in Chalcotongo Mixtec, where there are attested cases of a single verb taking an unmarked goal in one case and a goal marked with noun in another, with no semantic difference in the two discernible to me. I have not discovered any principle according to which one could determine whether or not to use noun in a particular goal expression.
This seems to me to be just the same phenomenon as discussed by Langacker (1982) as "active zones". In his terms, the verb requires an active zone of the area specified by the nominal compound.

I use the term "derived" metaphorically: I do not intend to suggest that there is a rule to which one form serves as input and the other output.

Since I do not have the intuitions of a native speaker for the Mixtec data, I am extrapolating from the judgments of my consultant.

"Over" here is in some respects an unfortunate choice for my purposes, since it has senses, related to but distinguishable from this one, in which it could be argued that it is basically a "partitive" sort of locational (see (74)).

References


288


MACRI, MARTHA. 1981. Implications of nominal constructions for the definition of "Word" in Mixtec. ms.


Two Noun Class Systems in Mixtec

Martha J. Macri

University of California, Berkeley

Words in Mixtec languages are typically disyllabic. A number of nouns having general meaning can substitute for other nouns and so function as pronouns. When they occur as a subject following a verb, or as a possessor following a noun, and are not further modified themselves, they are shortened to a single syllable (Pike 1944:131). These pronominal onclitics effectively divide Mixtec nouns into genders which minimally include feminine, masculine, animal, and usually sacred.

Another feature of Mixtec is the use of superordinate terms in compounds with nouns (in a few cases with verbs or adjectives) to indicate certain lexical classes. Both full and cliticised nouns occur in the function of classifiers. A further shortening of certain classifier plus noun compounds has resulted in their being reduced to a single disyllabic word. This process of morphological layering in which affixes tend to coalesce with the root has been described by Gibson and Bartholomew (1979:309 n2) as characteristic of the Otomanguean family in general.

There are, then, two independent noun class systems: one, a gender system, the other a partial system of lexical classification. As we examine the two we will see that even though the classifying terms of both systems are, in a few cases, phonologically identical and semantically related, they have very different syntactic functions. And while the pronominal onclitics are not closely enough attached to the words
they follow to be termed affixes, the monosyllabic morphemes of the lexical class system are intimately connected to the word they precede and are properly called prefixes.

Mixtec is, in fact, not a single language, but a complex of perhaps as many as thirty mutually unintelligible dialects. Data for this study are from Atatlajuca (Alexander 1980), San Miguel (Pike 1944 and Dyk and Stoudt 1965), and Chalcatongo (author's field notes), three neighboring towns in the former district of Tlaxiaco; Peñoles (Daly 1973) spoken in the district of Etlá; and two southern dialects, Jicaltepec (Bradley 1970) and Chayucu (Pensinger 1974). Figure 1 gives their approximate geographical locations, and identifies their subgroups as given by Bradley (1970:1).  

The Gender System

Greenberg (1978:50) defines a noun gender system as:

a system in which the noun stems of a language are divided into a set of genders, the distinction being based on the fact that the choice of a noun belonging to a particular gender determines the choice among a set of alternative "agreeing" forms in one or more other classes of morphemes or words, e.g., articles, demonstratives, adjectives, unbound anaphoric pronouns, pronouns incorporated in a verb complex, etc.

In the case of Mixtec, the agreeing forms are the pronominal enclitics which mark possession on nouns and subjects on verbs. In most cases the derivation of the enclitics from a full noun form is obvious. Table 1 shows the third person pronominal enclitics of six Mixtec dialects and their full noun forms. Some of the blanks in the chart would probably be filled in if more complete data were available. All six dialects have a feminine, masculine, and animal gender. All but Chayucu have a
<table>
<thead>
<tr>
<th>DIALECT</th>
<th>SUBGROUP</th>
<th>DISTRICT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peñoles</td>
<td>eastern</td>
<td>Etlá</td>
</tr>
<tr>
<td>San Miguel</td>
<td>central</td>
<td>Tlaxiaco</td>
</tr>
<tr>
<td>Chalcatongo</td>
<td>central</td>
<td>Tlaxiaco</td>
</tr>
<tr>
<td>Atatlahuca</td>
<td>central</td>
<td>Tlaxiaco</td>
</tr>
<tr>
<td>Jicaltepec</td>
<td>southern</td>
<td>Jamiltepec</td>
</tr>
<tr>
<td>Chayucu</td>
<td>southern</td>
<td>Jamiltepec</td>
</tr>
</tbody>
</table>

Figure 2. Mixtec dialects cited
<table>
<thead>
<tr>
<th>Gender</th>
<th>Atlatlahuca</th>
<th>San Miguel Chalcatongo</th>
<th>Peneles</th>
<th>Jicaltepec</th>
<th>Chayuco</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Feminine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron</td>
<td>-ña</td>
<td>-ña</td>
<td>-ña</td>
<td>-ña</td>
<td>-ña</td>
</tr>
<tr>
<td>Noun</td>
<td>Ṇã7q</td>
<td>Ṇã7q</td>
<td>Ṇã7a</td>
<td>Ṇã-d7q</td>
<td>Ṇã7q</td>
</tr>
<tr>
<td>Gloss</td>
<td>woman</td>
<td>woman</td>
<td>woman</td>
<td>woman</td>
<td>woman</td>
</tr>
<tr>
<td><strong>Masculine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron</td>
<td>-de</td>
<td>-de</td>
<td>-dé</td>
<td>-rá</td>
<td>-ra</td>
</tr>
<tr>
<td>Noun</td>
<td>tēc</td>
<td>tēc</td>
<td>rāi</td>
<td>rayēt</td>
<td>man</td>
</tr>
<tr>
<td>Gloss</td>
<td>man</td>
<td>man</td>
<td>man</td>
<td>man</td>
<td></td>
</tr>
<tr>
<td><strong>Animal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron</td>
<td>-tī</td>
<td>-tī</td>
<td>-tī</td>
<td>-ri</td>
<td>-tī</td>
</tr>
<tr>
<td>Noun</td>
<td>kitī</td>
<td>kitī</td>
<td>kitī</td>
<td>kitī</td>
<td>kitī</td>
</tr>
<tr>
<td>Gloss</td>
<td>animal</td>
<td>animal</td>
<td>animal</td>
<td>animal</td>
<td>animal</td>
</tr>
<tr>
<td><strong>Sacred</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron</td>
<td>-ya</td>
<td>-ya</td>
<td>-ya</td>
<td>-yā</td>
<td></td>
</tr>
<tr>
<td>Noun</td>
<td>yāw</td>
<td>i7yā</td>
<td>i7a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gloss</td>
<td>god</td>
<td>god</td>
<td>god</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Inanimate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron</td>
<td>-ši</td>
<td>-či</td>
<td>-ši</td>
<td>-si</td>
<td></td>
</tr>
<tr>
<td><strong>Human, collective</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron</td>
<td></td>
<td>-yo</td>
<td>-yu</td>
<td>-nī</td>
<td>-ŋu</td>
</tr>
<tr>
<td>Noun</td>
<td>Ṇayśu</td>
<td>Ṇayśu</td>
<td>Ṇiwí</td>
<td>Ṇiyivi</td>
<td></td>
</tr>
<tr>
<td>Gloss</td>
<td>people</td>
<td>people</td>
<td>people</td>
<td>people</td>
<td>people</td>
</tr>
<tr>
<td><strong>Child/person</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron</td>
<td>-i</td>
<td>-i</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noun</td>
<td>sēčí</td>
<td>sēčí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gloss</td>
<td>child</td>
<td>child</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tree/vehicle</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-tu</td>
</tr>
<tr>
<td>Noun</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gloss</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yutu</td>
</tr>
<tr>
<td><strong>Water/rain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron</td>
<td>-te</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noun</td>
<td>ndute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gloss</td>
<td>water</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* -či is realized as /či/, /če/, or /ča/ in various environments.

Table 1. Third person pronouns in Mixtec
sacred gender. Atatlahuca and San Miguel have an -ı enclitic derived from sùčí 'child.' In Atatlahuca it also has the more general meaning of 'person.' Peñoles, Jacaltepec, and Chayuco have an inanimate gender, not shared by the central dialects. Several groups also have a human collective gender 'people.' Chayuco has a separate gender for trees and vehicles. This unlikely pairing may be explained by the fact that trees and wooden objects are often classed together, and that the first wheeled vehicles seen by the Mixtec were undoubtedly made of wood. Atatlahuca is unique in having a separate gender for water and rain.

Peñoles has two encliticized forms for the feminine. -a may be derived from 'woman' ŋa-diă as is the case with the other dialects, and is used for women of high respect. The other form -di is the less formal pronoun, and may be related to šita 'grandmother.'

The following examples show the enclitics used with nouns to indicate the gender of the possessor:

(1) sò7o - i
    ear    3S Child
    'his/her (child's) ear'

(2) k'ayú - de
    horse  3S Masculine
    'his horse'

(3) nu7u - de
    tooth  3S Masculine
    'his tooth'
(4) sáta - di (Péones)
    back 3S Animal
    'its (animal's) back'

(5) sa7ma - rá (Jicaltepec)
    clothes 3S Masculine
    'his clothes'

(6) za7ma - ŋa (Chayuco)
    clothes 3S Feminine
    'her clothes'

Pike (1944:132) illustrates how in San Miguel full noun forms are
used whenever there is further modification by another noun or adjective:

(7) yá7a kúu ve7e - i
    this is house 3S Child
    'This is the child's house.'

(8) yá7a kúu ve7e súcí lúlí
    this is house child little
    'This is the little child's house.'

These examples show how the onclitics function as person markers
with verbs to indicate the gender of the subject:

(9) ñu7q ñ viko - ŋa stañ nuñ maestro (Atatlahuca)
    woman the sells 3S Fem. tortillas to teacher
    'The woman sells tortillas to the teacher.'
(10) Hu'oves ŋako - yo ḋayáu ṭuu (Chalcatongo)
Thursday come from 3S Coll.people town(s)
'On Thursday people come from many towns.'

(11) nì kišì - yà (San Miguel)
Past sleep 3S Sacred
'She/he (the god or saint) slept.'

(12) nì' - ŋešì - di ndà-yọ7o (Peñoles)
Past eat 3S Animal dried cornstalks
'It (the animal) ate dried cornstalks.'

(13) kišì - ŋa (Jicaltepec)
came 3S Feminino
'She came.'

(14) ndito - tu (Chayuco)
live 3S Tree
'It (the tree) is living.'

In many, perhaps all, Mixtec dialects, the use of the enclitic following the verb is optional if the subject is given in its full form as in (9) and (10) above.

The Lexical Class System

For the purposes of this paper we will define a lexical class system as a noun class system in which the nouns of a language are divided into classes by the use of superordinate terms, which are combined with the noun itself, but do not involve agreement with any other class of
morphemes. It also differs from a gender system in that it does not include every noun in the language, unless we posit a large and heterogeneous class with a zero marker. Greenberg (1978:51) notes that such a class system has no syntactic function, and therefore is not grammatically relevant. In Mixtec the superordinate term can be either a full noun or an affix which precedes the noun in question.

Examples of noun-noun compounds, in which the second stands in a genitival relationship to the first are frequent in Mixtec. The following are from Chalcatongo:

\[
\begin{align*}
\text{ñuù bikò} & \quad \text{town + cloud} & \quad \text{'town of clouds'} \\
\text{čii kêtê} & \quad \text{belly + animal} & \quad \text{'belly of the animal'} \\
\text{nda7a yuča} & \quad \text{hand + river} & \quad \text{'tributary of the river'} \\
\text{nduči í7a} & \quad \text{eye + god, holy one} & \quad \text{'the eye of the Virgin'}
\end{align*}
\]

This can extend to even longer sequences such as:

\[
\text{rayu í7a ndikandi} \quad \text{ray + god + sun} \quad \text{'rays of the sun god'}
\]

To the degree to which certain body-parts used as locative prepositions can be called noun-noun constructions, they do not violate this pattern:

\[
\begin{align*}
\text{šínlì yuna} & \quad \text{head + tree} & \quad \text{'above the tree'} \\
\text{nuù yuku} & \quad \text{face + mountain} & \quad \text{'on the mountainside'} \\
\text{čii mèsa} & \quad \text{belly + table} & \quad \text{'under the table'}
\end{align*}
\]

The same construction occurs with the lexical class system in which nouns of general meaning are modified by other nouns to create what Greenberg calls lexicalized phrases (1978:51). In Chalcatongo, for example, we find the following town names:
nuyu kava  
town + boulder  
'Boulder Town'

nuyu viko  
town + cloud  
'Cloud Town'

nuyu nde7eyu  
town + mud  
'Mud Town'

and in San Miguel:

nuyu yoskuía  
town + ?  
'Juxtlahuaca'

nuyu ŋinj nuyu  
town + head + town  
'Chalcatongo, Town above the Towns'

nuyu ŋinj viko  
town + head + cloud  
'San Juan Mixtepec, Town above the Clouds'

In the Chayuco dictionary nine different kinds of buildings are listed, all beginning with ve7e 'house.' For example:

ve7e kaa  
house + metal  
'jail'

ve7e ndiši  
house + whiskey  
'bar'

ve7e tyiño  
house + authority  
'town hall'

Other classes marked in Mixtec by compounds of two nouns include 'beans,' 'snakes,' 'words,' 'flowers,' 'birds,' 'sicknesses,' and 'men.'

A second type of compound consists of a classificatory prefix followed by a noun. In Atlatlahuca the word for 'tree, wood' is yuñu. It is shortened to nu and used as a prefix on names of trees and on objects made of wood:

nunde7e'  
'fruit tree'

nuuyhe'  
'ocote grove'

nu7isá  
'loom'

nuhií'  
'rifle'

In a few cases classifying prefixes are phonologically and semantically related to encliticized pronouns.
The following examples from Chayuco begin with the shortened form of ṇa7a 'woman':

- ṇa tə7a yo woman + relative 'female companion; cousin of a man'
- ṇa tyivaa woman + noisy one 'girl, young woman'
- ṇa yoko woman + marriageable 'senorita, virgin'
- ṇa zi7a woman + mother 'woman'

A large majority of animal names in all Mixtec dialects begin with ťi- or ti-. Words for objects such as 'blanket' and 'onion' are also included in this group. According to Longacre (1957:148), the prefix ťi-/ti- represents the merger of two historically different morphemes. One is the Proto-Mixtecan classificatory noun *tu 'egg, fruit, round object' (Rensch 1976:213) which has extended its range of meaning to include 'thing.' The other is *kitu 'animal' which has become kıť. So while the animal gender suffix -ťi is a shortened form of the word 'animal,' at least some of the classificatory prefixes which are identical to it come from an entirely different source. In fact, in the examples below, etymologies of words beginning with ťi- are offered using only the more general meaning of 'thing.' These words are from San Miguel, but nearly all of them are also used by Chalcatongo speakers (often with a difference in tone).³

- ťa kə 'thing' + kaa 'go up' 'grasshopper'
- ťa kəť 'thing' + kači 'cotton' 'blanket, poncho'
- ťa kəkə 'thing' + kókó 'swallow' 'maggot, beetle'
- ťa kəhə 'thing' + kəkə 'snarl' 'disturbance'
- ťa kəsə 'thing' + kəsə 'toast' 'corn bread'
- ťa ňu 'thing' + ňu 'face' 'owl'
Longacre (1957:66) notes that tvy- becomes $c$- in the central dialects, and ty- in Jicaltepec (Chayuco). This amounts to a fusion of the tâ- prefix with words which begin with y. San Miguel has both fused and unfused forms.

- tîyókó, čókó (San Miguel)
- tyoko (Chayuco)
- 'ant''
- tîyô7ô, čô7ô (San Miguel)
- tyo7ô (Chayuco)
- 'flea''

For some words in San Miguel only the fused form is used:

čuku rî $t$-yuku rî $t$â + yuku 'louse' + rî $t$â 'lamb' 'tick'

A number of animal names and objects also begin with ndâ-/ndî- and lâ-/li. Many are analyzable in the same way as words beginning with ti-, although the meanings of the prefixes are not clear. The cognate sets in Longacre (1957) show that ndâ- and ti- prefixes were separate at the level of Proto-Mixtecan. Some examples from San Miguel with recognizable second elements are:

- ndikô' ndi- + koô 'snake' 'lizard'
- ndîyunu ndi- + yunu 'tree, wood' 'an insect'

In one example the prefix is added to a Spanish loan:

- ndîši7û ndi- + Spanish chivo 'goat' 'goat'

This is evidence that the process of prefixing has been operative at some period since Spanish contact.

In the six Mixtec dialects examined in this paper, there are two words for 'tree': yunu in the central dialects, and yutu in Peñoles and
Jamiltepec. What is interesting is that \textit{yu-}, \textit{-nu}, and \textit{-tu} have all been used as classifiers for 'tree.' In San Miguel we find the following:

\begin{itemize}
\item \textit{yuwa} \ 'herb'
\item \textit{yułu} \ 'shrub, herb'
\item \textit{yuha} \ 'foliage'
\item \textit{yușa} \ 'foliage'
\item \textit{yuși} \ 'hay'
\end{itemize}

Most of these words are also found in Chayuco and other dialects. However, in languages that have \textit{yunu} for 'tree,' \textit{-nu} is also used as a classifier as in the examples from Atatlahuca given above, and in San Miguel:

\begin{itemize}
\item \textit{nundóko} \ 'zapote grove' \textit{nu-} + \textit{ndoko} \ 'zapote'
\item \textit{numa} \ 'tree leaves' \textit{nu-} + \textit{ma} \ ?
\item \textit{nupélé} \ 'pear tree' \textit{nu-} + \textit{pélé} \ 'pear'
\end{itemize}

The last example is another case of a classifier added to a Spanish loan (\textit{pélé} from \textit{pera}).

In Chayuco and Peñoles where 'tree' is \textit{yutu}, \textit{-tu} is used as a prefix:

\begin{itemize}
\item \textit{tu yavi šuu ve7e} \ 'cross beam' \textit{} \textit{(Chayuco)}
\item \textit{tu yoo} \ 'reed grass' \textit{} \textit{(Chayuco)}
\item \textit{tu yusa} \ 'ocote stick' \textit{} \textit{(Chayuco)}
\item \textit{ty-tlēči} \ 'avocado tree' \textit{} \textit{(Peñoles)}
\item \textit{ty-yoo} \ 'cane' \textit{} \textit{(Peñoles)}
\end{itemize}

The use of the first syllable of 'tree' as a prefix is probably older since the \textit{yu-} compounds are now disyllabic, while compounds with \textit{nu-} and \textit{tu-} are usually trisyllabic, often analyzable. The incorporation
of a Spanish loan confirms the fact that this process has been productive since the Spanish conquest.

In Chayuco, the one dialect we have looked at that has a 'tree' gender, it is -tu, the second syllable of yutu 'tree' which is the encliticized pronoun. While the lexical class system seems to have used different prefixes at different points in history, we do not have evidence of a similar sequence in the gender system.

Discussion

A certain amount of caution must be exercised in trying to establish the relative ages of the two noun class systems. For example, since nearly every gender is so clearly identified with a full noun form, the gender system would appear to be a very recent development. And since the prefixes used in the lexical class system have, in some instances, fused with the noun itself, it would seem to be older. However, both gender systems and the tendency to use lexical classifiers are found throughout the Otomanguean family. The external realization of the gender system in Mixtec may not be ancient; the idea of the system is. Within the lexical class system there exist noun-noun compounding which is productive at the present time, prefix-noun compounding which has been productive at some time since the Conquest as attested by the incorporation of Spanish loans, and much older disyllabic nouns in which the prefix has completely fused with the stem.

The fact that the pronominal enclitics have not undergone fusion with the words they follow, can perhaps be explained by their syntactic relationship with them. When they serve as subject markers following
verbs, they are part of the subject noun phrase, separated from the verb phrase at the highest node. For example:

```
S
   /\   \\
  VP  NP
   \ /   \\
   kusu  te
```

will sleep it (the animal)

'It (the animal) will sleep.'

The superordinate terms of the lexical class, however, are always part of the same noun phrase as the word they classify, and are thus more intimately bound to it, and more easily affected by it on a phonological level. Therefore, the difference in the degree of fusion of the two noun class systems, which might appear to indicate that the lexical class system is older, is actually explained by syntactic differences. 4

In summary, Mixtec appears to have two independent noun class systems which, although they sometimes define similar semantic classes, function in very different ways. The gender system is a covert system in which the marker does not appear on the noun itself, but is an enclitic which marks subjects of verbs and possessors of nouns. The lexical class system, on the other hand, is an overt system in which the classifying terms are marked directly on the noun itself. This prefixing may take the form of a lexicalized phrase, in which the classifier word occurs before a full noun, a trisyllabic word which includes a classifying prefix, or a disyllabic word which consists of a prefix and a cliticized root.
What are needed now are more detailed studies of the noun class systems in Mixtec itself, as well as descriptions of the systems in neighboring, related languages. The rich variety of such systems in the Otomanguean family may provide insights into the more general problem of their genesis and development.

Acknowledgements

I would like to express my appreciation to Leanne Hinton for her encouragement and comments during the time I was writing this paper. I would also like to thank Nicolas Cortez from Chalcatongo, Oaxaca, who shared with our class his language and his friendship.

Footnotes

1The orthographies of the sources have been kept, with the following exceptions: the voiceless velar stop is represented by k rather than c and qu, the glottal stop by ʔ rather than h, the glottal fricative by h rather than j, the voiceless palato-alveolar fricative by ʁ rather than x, the voiceless palato-alveolar affricate by Ç rather than ch, and the voiced palatal fricative/approximant by y rather than ü. Vowel nasalization is indicated by ворот. High tones are marked by ˥, low tones by ˥˥. Mid tones are unmarked.

2In this paper we are considering only the third person singular.

3Identifications of the second elements in the following compounds should be considered tentative.

4I am grateful to Leanne Hinton for this important observation.
References

ALEXANDER, RUTH MARIA. 1980. Gramatica Mixteca de Atatlahuca. Instituto Linguistico de Verano, Mexico, D.F.


DYK, ANNE and BETTY STOUDT. 1973. Vocabulario Mixteco de San Miguel el Grande. Instituto Linguistico de Verano, Mexico, D.F.


LONGACRE, ROBERT E. 1957. Proto-Mixtecan. Indiana University, Bloomington.


306
Preliminaries to Tonemic and Tonomechanical Analysis

for the Chalcatongo Dialect of Mixtec

Nicholas Farclas

University of California, Berkeley

This study is a phonetic description of the phenomenon of tone in the Chalcatongo dialect of Mixteco as well as a preliminary analysis of tone sandhi phenomena.

The transcription system used in this study is that of the international phonetic association, as revised in 1979. Chao's system of tone representation is used throughout, only three levels of tone ($\acute{1}$ hi, $\acute{1}$ mid, $\acute{1}$ lo) being necessary in most cases. However, when data from the San Esteban dialect is cited, $\acute{1}$ is used to represent level 3 (between mid and lo). Acoustic data is always given in Hertz and always refers to the level of the fundamental frequency of the audio signal.

The data upon which this work is based was collected during the course of some sixty sessions with language consultant Nicolás Cortes, without whose most patient and gracious cooperation, such a project would be an impossibility. Any inaccuracies or oversights herein are due solely to the writer's lack of care or experience in collecting or analyzing the data.

Preliminaries to Tone Analysis

In the dialect of Chalcatongo, as in other dialects of Mixtec, the basic morphemic canon is CV.CV or CVV. CVV morphemes are disyllabic,
even if both vowels are of the same quality. Each vowel is associated with a tone. Because all non-cliticized morphemes in the Chalcatongo dialect are disyllabic, two tones are associated with each of them. It is convenient, then, to consider as a single unit each of these pairs of tones commonly referred to as tone couplets.

In his classic study of tone, Kenneth Pike (1948) makes considerable reference to the variability of tone at the phonetic level (Pike, 1948). The realization of tonemes, the tonal equivalent of phonemes, depends on two major types of factors. On one hand, tone may be affected by the pragmatic or segmental context in which it is found. In some languages, particular phonetic environments, such as following a voiced stop or utterance-final position, cause change in tone. In all languages such factors as the physical and emotional state of the speaker are known to influence tone (Pike, 1948:27). On the other hand, tones may be altered by the position they occupy in the general tonal pattern of the utterance or syntactic unit of which they are a part. In this type of perturbation of tone, called sandhi, the tone of certain syllables is determined by the tone of adjacent or nearby syllables. The capacity of a given toneme (or, in our case, of a given tone couplet) to cause a perturbation or to be perturbed by another toneme determines the tonomechanical relationship of that toneme to the other tonemes of the language in question.

In order to determine the nature of tonemes and their tonomechanical interaction, it is necessary to know precisely in which environments and to what degree the physical and phonetic contexts of the utterance are influencing the realization of tone.
In order to do just this in our study of the dialect of chalcatongo, several elicitations were made under controlled conditions at the phonology laboratory of the University of California at Berkeley. Data thus obtained was then analyzed for fundamental frequency level (the primary acoustic cue for tone, abbreviated $F_o$) using a pitch extractor and an oscilloscope printer.

(Tables are at end of paper)

Table 31 indicates that tone level, as predicted by prior analysis by means of minimal pairs and test by ear, clusters around three frequency levels. The samples in Table 31 were recorded on two separate occasions. The fact that tone range varies from session to session is graphically illustrated here in the dramatic difference in overall range of tone between the samples. Nevertheless, the validity of the analysis of the tonemic system of the chalcatongo dialect as consisting of three tones; one high, one low and one at roughly mid-level, is confirmed by the almost identically proportional relationship of the level of one tone with respect to another between the samples. The level of tone so determined for each elicited morpheme was checked against the tone level determined by ear and cross-checked against the tone level recorded for the neighboring dialect of San Miguel El Grande by Anne Dyk and Betty Stoudt (1965). Results were remarkably consistent, except for morphemes classed in our sample as having lo-lo tone (a category which does not exist in San Miguel) which in almost every instance had been reported as lo-mid by Dyk and Stoudt.

Since the position of the syllable in which a given toneme finds itself in the utterance may also affect its realization, an analysis of
the data was made in terms of which syllable of the couplet a given toneme at a given level occurred. The results are shown in Table 92.

Mid tone and especially low tone are significantly lower in the second syllable of a morpheme than in the first, but it has proved unnecessary to consider tonemic neutralization or to posit new tonemes to account for this phenomenon.

Next, an analysis was made of each tonal level in terms of couplet structure and of vocalic and consonantal composition. The results are represented in Tables 93, 94, and 95. Tonemes seem to be influenced slightly, by the level of the other toneme of the couplet in which they are found. In fact, the overall pitch patterns for many morphemes with tonemes at different levels looked more like gliding contours from one level to the other than like couplets of two level tones. Of course, such glide formation is predictable and in no way bars us from analyzing the tonemes of the Chalcatongo dialect as being level.

Tone measurements were made in conformity with the guidelines set forth by Pike (1948:6) at the endpoints of glides and at points where the direction of $F_0$ variation changes. Partly because of the fact that measurements were taken at these points rather than elsewhere, the pitch level for low tone is dramatically different for the low-low couplet in relation to low tones that represent the beginning of a rising contour, or especially the end of a falling contour. This difference, however, is not significant, since the lowest mid tones in the same environment are not only a full twenty Hertz higher than the highest low tones, but are phonetically if not tonemically endpoints also of downward glides.

Tone level does not appear to be significantly affected by the canonical structure (CVCV or CVV) of the morpheme concerned. Variation
in terms of vocalic and consonantal makeup of the syllable associated with a given toneme, while notable in some cases, is never significant.

In order to determine the pitch differences inherent in Chalcatongo vowels an experiment was devised where in the language consultant read sixty random sentences, including morphemes identical in structure and tone, varying only in the quality of their vowels (twenty contained [ə], twenty contained [u], and twenty contained [i], which were embedded in identical frames. The results were analyzed much like the other data thus far presented and are shown in Table 86. It appears that variation due to vowel quality, while comparable to that found in other languages, is not great enough to cause confusion in toneme perception.

Mariscela Amador (personal communication) reports that, in the chalcatongo dialect, syllable duration, closely related to syllable structure, is not correlated to tone or amplitude. While the majority of morphemes in our sample had higher amplitude on the first syllable, first syllables were almost invariably shorter than second syllables. Stressed second syllables were not notably longer than unstressed second syllables in relation to first syllables and unstressed initial syllables bore roughly the same relation to final syllables in duration as did stressed initial syllables.

Amplitude seems to be correlated, though not perfectly, with the beginning of a falling contour or with the end of a rising contour, that is, with the highest tone in the morpheme. Morphemes with two tonemes at the same level are very commonly realized with near equal amplitude on both syllables, especially when the structure is CVCV.

This positive correlation between amplitude and toneme level renders toneme identification relatively easy in most cases, but in a
significant minority of cases, where the correlation does not hold, a great deal of concentration was required to determine the couplet type of the morpheme concerned. Since such 'aberrant' morphemes have similar \( F_0 \) levels as others of their class and, as we shall see shortly, behave as others of the same couplet type, we may conclude that the primary cue for toneme level is \( F_0 \) level, but that amplitude may be a secondary cue of some importance (at least to the ear of an English-speaking linguist trying to identify tonemes!)

Amplitudes were extracted from the sample with the same machinery used for fundamental frequency extraction. The importance of experimental verification of audial impressions cannot be stressed enough here. Native speakers of English rely on change in pitch primarily, and on duration, absolute pitch, and amplitude secondarily to determine stress patterns. In the Chalcatongo dialect a similar correlation exists between absolute pitch and amplitude, and, to a limited extent, pitch change. However, duration of syllables is completely disassociated from these other variables, and furthermore, it is not at all clear that the phenomenon of stress has anywhere near the importance in the Mixtec of Chalcatongo as it does in English, vowel quality being much more stable in the former than in the latter.* In any case, further investigation in the areas of amplitude and stress is sorely needed.

Mixtec couplets may at one time have been two or more separate morphemes. A possible trace of this type of development is the

* In many cases, pitch and amplitude are not correlated, and these are precisely the cases where speakers of English will have most difficulty in distinguishing tonemes in the chalcatongo dialect. (in the silacayopayan dialect (North and Shields, 1977) stress falls regularly on the first syllable of all couplets).
existence of two vowels of different quality in the same morpheme, especially the CVV type. If the percentage of morphemes containing vowels not differing in timbre is checked against couplet types, some types emerge as more likely than others to have arisen from this type of process than other couplet types. Table 87 shows that couplets with rising contours are not only rarer than other types of couplets, but are also more likely to contain vowels that are not identical in quality.

In this connection, we may also consider the percentage of each couplet type in the data sample used in this study. A total of 177 couplets were used. Each type was represented in the following proportion:

<table>
<thead>
<tr>
<th>Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{\textdagger})</td>
<td>18%</td>
</tr>
<tr>
<td>(\text{\textdagger}\text{\textvirgolette})</td>
<td>5%</td>
</tr>
<tr>
<td>(\text{\textdagger}\text{\textdagger})</td>
<td>9%</td>
</tr>
<tr>
<td>(\text{\textdagger}\text{\textdagger}\text{\textdagger})</td>
<td>20%</td>
</tr>
<tr>
<td>(\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger})</td>
<td>28%</td>
</tr>
<tr>
<td>(\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger})</td>
<td>5%</td>
</tr>
<tr>
<td>(\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger})</td>
<td>9%</td>
</tr>
<tr>
<td>(\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger}\text{\textdagger})</td>
<td>6%</td>
</tr>
</tbody>
</table>

The data sample, however, reflects only indirectly the actual proportion of a given couplet type in the dialect, since rare couplet types were more actively elicited than more commonly occurring couplet types in order to ensure an adequate data base for each type. Therefore,
couplet types rare in the sample are most probably even rarer in proportion to the entire Chalcatongo lexicon.

Preliminary Analysis of Some Sandhi Phenomena in Certain Chalcatongo Constructions

Because of time limitations, a complete study of tonemic perturbation patterns in the Chalcatongo dialect has thus far not been possible. Analysis of sandhi phenomena is extremely time-consuming, since data must be checked and rechecked to control for variation due to style, utterance rhythm and speed, special behavior of particular morphemes or classes of morphemes, and variations in degree or type of sandhi due to syntactic context.

In our sample, all possible couplet types were found, another indication that our analysis of the tonemic system as consisting of three level tones is correct. In the San Miguel dialect, however, the sequence lo-lo is not found (Pike, 1948:79) and in the dialect of San Esteban Atatláhuca, as described by Mak (1953:86) which contains four level tones, lo-lo sequences are also not found. Whether the sequences lo-lo and lo-mid may be collapsible in the Chalcatongo dialect is an as yet unsolved problem (see below).

In the Chalcatongo dialect, as in other Mixtecan dialects, a genitival relationship may exist between two nouns in sequence. For example:

```
\[\text{[sn?n tfaa]} \quad \text{[hi to tfaa]}
\]
\text{leg man} \quad \text{bed man}

'the man's leg' \quad 'the man's bed'
```

314
the tonemes associated with the morpheme [tʃaː] in isolation are J.J.
We may conclude, then, that in this type of noun-noun construction, the
couplet J.J is perturbed to .tagName after .tagName and to J.J after J.J, at
least in the morpheme [tʃaː].

In Table 88 the general patterns of tone perturbation in noun-noun
constructions are schematically represented. Across the top row of the
table are found the perturbing couplets (i.e. the couplet types of the
first noun) and the horizontal rows represent the perturbed couplets
(i.e. the second term of the noun-noun construction.)*

On the basis of the data in Table 88, a preliminary classification
of couplet types into tonomechanical groups can be made.

Class I - .tagName, -tagName, J.J, and -tagName.

a) .tagName and -tagName never perturb any couplet type, except -tagName which
goes to -tagName. We shall call this type of perturbation, where two suc-
cessive tones are each raised by one tonemic level double one tone
raising.

b) J.J - never perturbs any couplet types except another J.J which
is perturbed to -tagName. In this way, a sequence of four tones at the same
level is avoided J.J + J.J + J.J + J.J. This seems to be a common process
in the Mixtec of Chalcatongo and is everywhere accomplished by raising

* Abbreviations used in Table 8: N.C. - no change, Var. - variation,
N+N - noun-noun construction, B/P - body part locative construction,
ADJ - noun-adjective construction, CVV, CVCV, etc. - canonical
structure.
the first tone of the perturbed couplet by one tonemic level, a process which we will call one tone raising.*

c) -J - never causes perturbation, except if the following couplet ends in ไกล, in which case one tone raising applies to the first toneme while the final ไกล goes to -.

Class I, in general, is a non-perturbing class. In the limited number of instances where class I couplet types do cause perturbations, these always involve one tone raising (simple or double).

Class II - \text{-1-1, 1-1} \\

a) When the sequences 1+1+1 and 1+1+1 occur underlyingly the second couplet is perturbed by means of one tone raising. Notice that once again one tone raising is used to avoid a sequence of four tonemes at the same level in the case of 1+1+1 + 1+1+1.

b) When class II couplets precede \text{-1} or \text{-1}, all low tones of the perturbed couplet are changed to high tones, as if a mirror image on an axis perpendicular to the frequency axis were being made of the contour. We shall call this phenomenon two step raising. An example of lo-hi raising is:

\[
\begin{array}{c}
\text{ไกล} \\
\text{[ndu tja]} \\
\text{water} \\
\end{array} + 
\begin{array}{c}
\text{านี} \\
\text{[na ni]} \\
\text{brother} \\
\end{array} \rightarrow \begin{array}{c}
\text{านี อิ} \\
\text{[ndu tja na ni]} \\
\text{'the brother's water'} \\
\end{array}
\]

* In the San Miguel dialect, what we call "one tone raising" above is the most common, indeed nearly the only type of morphotonemic perturbation reported for noun-noun constructions.

** Similar lowering phenomena are attested to in other dialects of Mixtec, as well as in other Otomanguean and non-Otomanguean languages (Daly, 1977:19) where they are allophonic in nature. We judge these to be allophonic here also. Further study is necessary.
c) \( \mathbb{J} \mathbb{J} \) and \( \mathbb{J} \mathbb{J} \) regularly perturb \( \mathbb{J} \mathbb{J} \) couplets to \( \mathbb{I} \mathbb{I} \). In fact, except for the above-cited perturbation of \( \mathbb{J} \mathbb{J} \) by \( \mathbb{J} \mathbb{J} \), a perturbed \( \mathbb{J} \mathbb{J} \) couplet always goes to \( \mathbb{I} \mathbb{I} \), in noun-noun constructions.

Class II couplets always perturb tone level upward, by the processes of one tone raising, two step raising, and special lo-lo raising \( (\mathbb{J} \mathbb{J} \rightarrow \mathbb{I} \mathbb{I}) \).

**Class III - \( \mathbb{I} \mathbb{I}, \mathbb{J} \mathbb{J}, \mathbb{J} \mathbb{I} \)**

a) Just as class II couplets, class III couplets perturb \( \mathbb{I} \mathbb{I} \rightarrow \mathbb{I} \mathbb{I} \) and \( \mathbb{J} \mathbb{I} \rightarrow \mathbb{I} \mathbb{I} \) by one tone raising.

b) As after class II couplets, \( \mathbb{J} \mathbb{J} \rightarrow \mathbb{I} \mathbb{I} \) after class III couplets.

c) The tone patterns for \( \mathbb{J} \mathbb{I} \) and \( \mathbb{J} \mathbb{I} \) following class III couplets are reversed. Example:

\[
\begin{array}{cccc}
\mathbb{J} & \mathbb{I} & \mathbb{J} & \mathbb{I} \\
[t\text{\&a} \text{ka}] & [n\text{a} \text{ni}] & [t\text{\&a} \text{ka} \text{na} \text{ni}] \\
\end{array}
\]

We shall call this type of perturbation **pattern reversal**.

d) \( \mathbb{J} \mathbb{I} \) - differs in its perturbing effect, with a particular four-tone series occurring in almost all couplet pairs with \( \mathbb{J} \mathbb{I} \) as the first component, yielding the tonal pattern \( \mathbb{J} \mathbb{I} \mathbb{I} \mathbb{J} \) (that is, \( \mathbb{J} \mathbb{I} \mathbb{I} \mathbb{J} \) plus its mirror image on the axis parallel to the frequency axis.) Therefore \( \mathbb{J} \mathbb{I} \) does not perturb \( \mathbb{I} \mathbb{J} \) following it as do other class III couplets. \( \mathbb{I} \mathbb{J} \) following \( \mathbb{J} \mathbb{I} \), is perturbed to \( \mathbb{I} \mathbb{J} \) instead of \( \mathbb{I} \mathbb{I} \). **Pattern reversal** may not account for the perturbatory effect of \( \mathbb{I} \mathbb{J} \) on the couplets to which it typically applies (\( \mathbb{I} \mathbb{J} \) and \( \mathbb{J} \mathbb{I} \)) since \( \mathbb{I} \mathbb{J} \) remains \( \mathbb{I} \mathbb{I} \), and while the pattern of \( \mathbb{J} \mathbb{I} \) is reversed to \( \mathbb{I} \mathbb{J} \), this seems to be due more to conformity to the overall \( \mathbb{J} \mathbb{I} \mathbb{I} \mathbb{J} \) sequence pattern than any other factor.
In any case, class III couplets distinguish themselves from those of other classes in that syllables of morphemes following class III couplets and perturbed by them can have a toneme at a lower level associated with them in their derived form than the level of the toneme associated with them in their citation form. Whether we want to interpret this phenomenon as actual tone lowering or, as I prefer, as a type of pattern reversal or pattern mirroring is an important question.

The pattern reversal analysis is favored for several reasons. In the first place, phenomena involving 'lowering' are never as simple as one tone raising, but always are included in patterns of perturbation involving more than one toneme. Secondly, in the neighboring dialect of San Miguel El Grande (Pike, 1948:80) perturbation always involves the process of one tone raising except in very isolated cases (Mak, 1950: 82-86). The extent of 'lowering' in Chalcatongo is extremely limited in comparison to other dialects in which lowering has been reported (Mak, 1953:82). Notice that only three out of the nine couplet types cause 'lowering,' and only one does so exclusively.*

A classification of couplet types on the basis of their perturbability is not as easy to devise as one based on their perturbatory effect. 77, 74, and 71 are never perturbed. As mentioned above 71 has a peculiar but regular pattern of perturbability. 47 can only be perturbed to 77. Accounts of the perturbability patterns of the other couplets would amount to case-by-case descriptions of the data in Table #8.

* It should also be noted that the perturbed second couplet of a couplet pair has a falling pattern much more often than a rising one.
Certain generalizations are possible concerning the correlation of couplet makeup with tonomechanical behavior. It appears that couplets ending in a hi toneme will be most likely to cause perturbation, while those ending in a lo toneme are least likely to do so. However, lo tonemes in couplet initial position seem to increase the susceptibility of the couplet to perturbation, while hi tonemes in initial position block perturbability.

The tonomechanical constraints cited thus far conform to the following applicational hierarchy:

1) Hi tone perturbability blocking.
2) Four toneme sequence at the same tonemic level avoided.
3) JJ is only perturbed to 71.

Hi tone perturbability blocking must be ordered before four same level toneme avoidance because the sequence 7777 remained unperturbed. Four same level toneme avoidance must be ordered before JJ + 71 because JJJJ + JJ+JJ, following the perturbation pattern typical of the avoidance constraint.

The above account of the tonomechanics of the noun-noun construction is by no means exhaustive. In other Mixtec dialects, couplets with the same basic toneme patterns behave differently in the same environment, grouping themselves into subclasses (Mak, 1953:89). In the San Esteban dialect certain syntactic units involve one type of perturbation while other constructions involve another (Mak, 1953:88).

Such may very well be the case in Chalcatonogo, since body part-noun locative constructions and potential verb-noun imperative constructions, while in some cases having similar or identical (hii-with, e.g.) tonomechanics to noun-noun genitive constructions, in other cases show
tonomechanical patterns particular to themselves. Noun-adjective constructions, on the whole, seem to pattern themselves after noun-noun constructions.

Verb-noun (or pronoun) direct object constructions involve almost no perturbation. Only 77 verbs appear to perturb in this instance, and then only two couplet types are affected (77+/+ → 777+ and 777+ → 777+).

In Chalcatongo, as in other Mixtec dialects, disyllabic morphemes may be cliticized to other morphemes, a process which involves the reduction of the cliticized morpheme to one syllable. Although in other dialects the toneme associated with the cliticized morpheme is that which was associated with the initial syllable of its non-cliticized counterpart, such a relationship does not seem to hold in Chalcatongo.

As illustrated in Table 89*, tonomechanical patterns for pronouns before or after verbs or after nouns (in the possessive construction) are distinct from those considered above in that pronouns may be perturbed by a following morpheme. This anticipatory perturbation pattern occurs in the San Esteban dialect as well (Mak, 1953:88).

Pronoun enclitics may be classified loosely into four groups:

a) Group 1 - mainly hi
   A) [na],[ni] - mainly stay 1 after 1
   B) [ro],[ri] - go to 1 or 1 after 1

---

* Abbreviations used in Table 89: 1s - first person singular, etc., P - plural, I - informal, R - respected or formal, F - feminine, M - masculine, REP - repetitive particle, FUT - future, VEN - ventive.
b) Group 2 - mainly 1d

A) [na] - ̃ goes to 7 after 1
    3s
B) [de] - ̃ remains ̃ after 1
    3s
c) Group 3 - [30]
d) Group 4 - [to] - always lo, but with 7 variant.

The preverbal particles tested followed two main patterns:

A) [ka]
   [kiʔi]
   [ni]
   Perturb following morpheme downward.
B) [na]
   [ha]
   Perturbed by following morpheme variously.

The great variety of perturbation by following morphemes on [na] and especially [ha] may indicate that these monosyllables are in reality neutral in tone. Such encliticized elements with neutral tone are called 'unstable neutral syllables' by Pike (1948:25). [ha] appears, however, to be basically lo when used preverbally to form relative clauses, acting as an unstable neutral syllable mainly in pre-nominal or pre-adjectival position. Relativizers before verbs in general, such as [nda sa] 'how', [nɔʔɔ] 'that', and [ndeu] 'which' also neither perturb nor are perturbed.

Conclusion

It is obvious that we have only begun to account for sandhi phenomena in the Chalcatongo dialect. It is our hope that this study will serve as the point of departure for deeper probes into the tono-mechanics of Chalcatongo morphophonology.
Acknowledgements

I would like to take this opportunity to thank Leanne Hinton, John Ohala, Mariscela Amador, Haruko Kawasaki, and John Kingston for their guidance and encouragement, and the National Science Foundation for its financial support.

References


Mak, Cornelia. "A unique tone perturbation in Mixtec." IJAL, Vol. 16, p. 82-86.


### Overall Mean Levels of Tones in Hertz ($F_0$)

<table>
<thead>
<tr>
<th></th>
<th>Sample 1</th>
<th></th>
<th>Sample 2</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td>Mid</td>
<td>Low</td>
<td></td>
<td>High</td>
<td>Mid</td>
<td>Low</td>
</tr>
<tr>
<td>7.6</td>
<td>8.3</td>
<td>2.5</td>
<td>9.9</td>
<td>15.1</td>
<td>10.8</td>
<td>155</td>
<td>83</td>
</tr>
<tr>
<td>18</td>
<td>28</td>
<td>15</td>
<td>58</td>
<td>155</td>
<td>83</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1**

323
Overall Mean Levels of Tones in Hertz by Syllable ($F_o$)

<table>
<thead>
<tr>
<th>Syllable 1</th>
<th>Syllable 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Mid</td>
<td>Mid</td>
</tr>
<tr>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Mid</th>
<th>Low</th>
<th>High</th>
<th>Mid</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.6</td>
<td>7.9</td>
<td>6.4</td>
<td>8.0</td>
<td>21.6</td>
<td>8.7</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>93</td>
<td>20</td>
<td>17</td>
<td>62</td>
<td>63</td>
</tr>
</tbody>
</table>

Table 2

324
### High Tone

<table>
<thead>
<tr>
<th>Syllable 1</th>
<th>Syllable 2</th>
<th>Syllable 1</th>
<th>Syllable 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVV</td>
<td>CVCV</td>
<td>CVV</td>
<td>CVCV</td>
</tr>
<tr>
<td>201</td>
<td>202</td>
<td>196</td>
<td>196</td>
</tr>
<tr>
<td>187</td>
<td>x</td>
<td>191</td>
<td>193</td>
</tr>
<tr>
<td>198</td>
<td>197</td>
<td>184</td>
<td></td>
</tr>
</tbody>
</table>

Mean Level in Hertz by Couplet Type ($F_o$)

### Mean Level in Hertz by Vowel of Syllable Nucleus or by Syllable Initial Consonant ($F_o$)

<table>
<thead>
<tr>
<th>Syllable 1</th>
<th>Syllable 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>a e i o u</td>
<td>c</td>
</tr>
<tr>
<td>[+nas]</td>
<td>[+voice]</td>
</tr>
<tr>
<td>188</td>
<td>190</td>
</tr>
<tr>
<td>192</td>
<td>193</td>
</tr>
<tr>
<td>206</td>
<td>196</td>
</tr>
<tr>
<td>194</td>
<td>202</td>
</tr>
<tr>
<td>204</td>
<td>188</td>
</tr>
<tr>
<td>194</td>
<td>198</td>
</tr>
</tbody>
</table>

* = < 3 realizations in sample
x = no data

Mean Level in Hertz by Vowel of Syllable Nucleus or by Syllable Initial Consonant ($F_o$)

Table 3
### Mid Tone

<table>
<thead>
<tr>
<th></th>
<th>CVV</th>
<th></th>
<th>CVCV</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>171</td>
<td>171</td>
<td>169</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>164</td>
<td>x</td>
<td>165</td>
<td>163</td>
</tr>
<tr>
<td></td>
<td>157</td>
<td>155</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

Syllable 1 | Syllable 2 | Syllable 1 | Syllable 2

Mean Level in Hertz by Couplet Type ($F_o$)

<table>
<thead>
<tr>
<th>Syllable 1</th>
<th></th>
<th>Syllable 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a e i ə o u</td>
<td>165</td>
<td>c [±nas]</td>
<td>158</td>
</tr>
<tr>
<td>c [±voice]</td>
<td>166</td>
<td>k 3</td>
<td>167</td>
</tr>
</tbody>
</table>

Mean Level in Hertz by Vowel of Syllable Nucleus or by Syllable Initial Consonant ($F_o$)

* = < 3 realizations in sample

x = no data

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4

326
Low Tone

<table>
<thead>
<tr>
<th>Syllable 1</th>
<th>Syllable 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVV</td>
<td>CVCV</td>
</tr>
<tr>
<td>110 115 115</td>
<td>110 115 116</td>
</tr>
<tr>
<td>130 134.5</td>
<td>130 134.5</td>
</tr>
<tr>
<td>JJ JJ JJ JJJ JJ JJJ</td>
<td>JJ JJ JJ JJJ JJJ</td>
</tr>
</tbody>
</table>

Mean Level in Hertz by Couplet Type ($F_o$)

Syllable 1

<table>
<thead>
<tr>
<th>Syllable 2*</th>
</tr>
</thead>
<tbody>
<tr>
<td>JJ HZ level discrepancy 'hidden'</td>
</tr>
<tr>
<td>110 118 120 116 115 117 117</td>
</tr>
<tr>
<td>a e i 4 o u c c [nas] [+nas] [+voice] k 3</td>
</tr>
</tbody>
</table>

* = < 3 realizations in sample

x = no data

Mean Level in Hertz by Vowel of Syllable Nucleus or by Syllable Initial Consonant ($F_o$)

Table 5

327
Fundamental Frequency Level for

[a] [u] and [i]

Number of Realizations at a Given f Level (in Hz.)

![Graphs showing F0 in Hz for [a], [u], and [i]]

### [a]

<table>
<thead>
<tr>
<th>Vowel</th>
<th>[a]</th>
<th>[u]</th>
<th>[i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean F0 in Hz.</td>
<td>148.4</td>
<td>153.8</td>
<td>157.1</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.70</td>
<td>3.65</td>
<td>3.99</td>
</tr>
<tr>
<td>Variance</td>
<td>13.73</td>
<td>13.33</td>
<td>15.80</td>
</tr>
</tbody>
</table>

### [u]

<table>
<thead>
<tr>
<th>Variables</th>
<th>t-value</th>
<th>df</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>[a] vs. [u]</td>
<td>4.52</td>
<td>38</td>
<td>&lt;.0005</td>
</tr>
<tr>
<td>[a] vs. [i]</td>
<td>6.97</td>
<td>38</td>
<td>&lt;.0005</td>
</tr>
<tr>
<td>[u] vs. [i]</td>
<td>2.66</td>
<td>38</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

Table 6
Percentage of Couplets where $V_1 = V_2$ by Couplet Type

<table>
<thead>
<tr>
<th>CVV</th>
<th>CVCV</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
</tr>
</tbody>
</table>

$x = < 3$ realizations in sample

Table 7
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8

Perturbations in Noun-Noun and Other Constructions

330
| Verbs     | 1PI | 1PR | 2PP | 2PR | 3SF | 3SM | 3SR | 1PP | 30 | 50 | 1ST | 1SF | 2PI | 2PP | 1PP | PL | (REP.) | na | PAST | FUT | VEN |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|----|----|-----|-----|-----|-----|-----|----|---|-----|---|-----|----|-----|
| nat₂ | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| pot. | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| t̃/inde | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| niso | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| pot. ku., | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| (pot. & real) | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| ken | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| hini | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| kuni | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |

Tonal Patterns of Pronouns and Particles
(if verb or noun is perturbed by the pronoun or particle, tones are indicated)

Table 9