
Santa Barbara Papers in Linguistics

VOLUME 11

Proceedings from the fourth
Workshop on American Indigenous Languages

July 6-8, 2001

Jeanie Castillo, Editor

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

SANTA BARBARA PAPERS IN LINGUISTICS

may be ordered from:

Papers in Linguistics
Linguistics Department
University of California, Santa Barbara
Santa Barbara, California 93106-3100
U.S.A.

Checks in U.S. dollars should be made out to UC Regents with \$5.00 added for overseas postage.

If your institution is interested in an exchange agreement, please write the above address for information.

Volume 1:	<u>Korean: Papers and Discourse Date</u>	\$13.00
Volume 2:	<u>Discourse and Grammar</u>	\$10.00
Volume 3:	<u>Asian Discourse and Grammar</u>	\$10.00
Volume 4:	<u>Discourse Transcription</u>	\$15.00
Volume 5:	<u>East Asian Linguistics</u>	\$15.00
Volume 6:	<u>Aspects of Nepali Grammar</u>	\$15.00
Volume 7:	<u>Prosody, Grammar, and Discourse in Central Alaskan Yup'ik</u>	\$15.00
Volume 8:	<u>Proceedings from the first Workshop on American Indigenous Languages</u>	\$20.00
Volume 9:	<u>Proceedings from the second Workshop on American Indigenous Languages</u>	\$15.00
Volume 10:	<u>Proceedings from the third Workshop on American Indigenous Languages</u>	\$15.00
Volume 11:	<u>Proceedings from the fourth Workshop on American Indigenous Languages</u>	\$15.00

FOREWARD

It is with great pleasure that we present the proceedings of the fourth Workshop on American Indigenous Languages (WAIL 2001). In continuing a tradition begun with the student discussion group on North American Indigenous Languages (NAIL), the evolving membership wishes to pay tribute to Marianne Mithun and Wallace Chafe for their consistent encouragement and support. We hope that this fourth volume of the Working Papers represents another step in the development of WAIL as a forum where we may all share our discoveries, both descriptive and theoretical, concerning these increasingly endangered languages.

Paul Barthmaier

Joe Holmberg

Greg Brown

Kirk Miller

Kelsi Camp

Jennifer Van Vorst

Jeanie Castillo

CONTENTS

FOREWARD	1
CONJUNCT/DISJUNCT SYSTEMS IN BARBACOAN LANGUAGES <i>Timothy Jowan Curnow</i>	3
TOPIC, FOCUS, DEFINITENESS, SPECIFICITY, AND REFERENTIAL PROMINENCE IN WESTERN APACHE NOUN PHRASES AND RELATIVE CLAUSES <i>Willem J. de Reuse</i>	13
SECONDARY PALATALIZATION AND CHANGES IN VOWEL FORMANTS IN ISTHMUS MIXE <i>Julia Dieterman</i>	20
DICCIONARIO YAQUI-ESPAÑOL: A LEXICOGRAPHIC PROJECT <i>Zarina Estrada Fernández</i>	34
INDERIVATION IN WEST GREENLANDIC – IN PASSIVE AND OTHER VALENCE DECREASING STRUCTURES <i>Karen Langgård</i>	47
POSITIONAL VERBS AND RELATIONAL NOUNS IN ZANIZA ZAPOTEC <i>Natalie Operstein</i>	60
CROSS-REFERENCING IN EMÉRILLON (TUPI-GUARANI): A HIERARCHICAL AGREEMENT SYSTEM <i>Françoise Rose</i>	71
NOMINAL CLASSIFICATION IN MIRAÑA <i>Frank Seifart</i>	85
TOWARD A TYPOLOGY OF POSITION CLASS: COMPARING NAVAJO AND KET VERB MORPHOLOGY <i>Edward Vajda</i>	99
THE SEMANTICS OF YUROK INTENSIVE INFIXATION <i>Esther Wood & Andrew Garrett</i>	112

Conjunct/disjunct systems in Barbacoan languages

Timothy Jowan Curnow
RCLT, La Trobe University

1. Introduction

The Barbacoan language family consists of at least four languages spoken in southwestern Colombia and northwestern Ecuador (Curnow and Liddicoat, 1998): Guambiano, Awa Pit or Cuaiquer, Cha'palaachi or Cayapa, and Tsafiki or Colorado.¹ Totoró is also clearly in this family; however it is not clear from available materials if it is better treated as a separate language closely related to Guambiano or as a dialect of Guambiano.

All languages of this family are typologically similar. For example, all are verb final (with the usual correlates), primarily suffixing, and with nominative-accusative case marking. However all also share a typologically unusual system of marking, which will be called conjunct/disjunct marking here, following the terminology used by Hale (1980) in his description of a similar phenomenon in Kathmandu Newari. While this is not necessarily the most appropriate terminology, especially given that in the Barbacoan languages there are three terms in the system, it is the most generally accepted label used in the literature for this phenomenon.

At first glance, conjunct/disjunct marking appears to be a type of person marking system; however the details are much more complex. This paper explores some of the intricacies of the conjunct/disjunct system of the Barbacoan languages.

2. Basic conjunct/disjunct system

In statements, the conjunct/disjunct system of Barbacoan languages initially appears to be a system of first versus non-first person marking, with the conjunct form occurring in first person statements while the disjunct occurs in second and third person statements, as in the following examples from Awa Pit:²

- (1) *(na=na)* *pala* *ku-mtu-s*
(1SG.(NOM)=TOP) plantain eat-IMPF-CONJ[C]
'I am eating plantains.'
- (2) *(nu=na)* *pala* *ku-mtu-y*
(2SG.(NOM)=TOP) plantain eat-IMPF-DISJ[D]
'You are eating plantains.'
- (3) *(us=na)* *atal* *ayna-mtu-y*
(3SG.(NOM)=TOP) chicken cook-IMPF-DISJ[D]
'He/she is cooking chicken.' (Curnow, In press)

However a simple glance at how the conjunct and disjunct forms are used in questions is enough to dispel the idea that the system is one of first/non-first person marking, since in questions the conjunct form is used with second person and, in Awa Pit at least, the disjunct is used with first and third person:

- (4) *m̄in-a=ma* *ashap-tu-y?*
 who-ACC=INTER annoy-IMPF-DISJ[D]
 ‘Whom am I annoying?’
- (5) *shi=ma* *ki-mtu-s?*
 what=INTER do-IMPF-CONJ[C]
 ‘What are you doing?’
- (6) *m̄in=ta-s* *a-mtu-y?*
 where=LOC-ABL come-IMPF-DISJ[D]
 ‘Where is he coming from?’ (Curnow, In press)

There is thus a binary distinction here between what will be termed locutor and nonlocutor participants: locutor covers first person in statements and second person in questions, while nonlocutor covers second and third person in statements and first and third person in questions. In the examples above (although not always, as will be seen below) conjunct marking is used when there is a locutor subject, and disjunct is used with nonlocutor subjects.

This unification of first person in statements with second person in questions appears odd, at first glance. But as has been pointed out by Hargreaves (1990; 1991), it can be understood in terms of epistemic source or epistemic authority. In making a statement, a speaker claims that he or she knows the information in the proposition, and thus first person has a privileged position. In asking a question, the speaker is claiming that the addressee has knowledge of the proposition; and thus in questions, the second person occupies a privileged position with respect to knowledge. This idea of a claim to knowledge thus unifies first person in statements and second person in questions, the locutor participants.

A conjunct/disjunct system is defined here as a system of marking where, at least in agentive, volitional contexts (or in neutral contexts with a predicate which is normally interpreted as being agentive and volitional), there is a distinction between the marking in statements which contain a first person subject and those which contain no first person participant; and the marking used in statements when there is a first person subject (conjunct marking) is also used in questions when there is a second person subject. As will be seen below, the marking used in statements when there is no first person argument (disjunct marking) may be used in questions which contain no second person subject; conjunct marking or other marking may also be used in other contexts; and different marking may be used in nonvolitional contexts, depending on the language.

The basic conjunct/disjunct system of Awa Pit was seen above, where the conjunct marking was used in locutor subject contexts, the disjunct marking in

nonlocutor subject contexts. This pattern can also be seen in the other Barbacoan languages for which data is available.

In Tsafiki, the Congruent suffix (conjunct) is used in locutor subject contexts (both statements, as in example (7), and questions, as in example (8)), while an unmarked verb (disjunct) is used for nonlocutor subject contexts in statements, as in example (9):

- (7) *tse Tsachi jo-yo-e*
 1F Tsachi be-CONGR-DCL[C]
 'I am a Tsachi'
- (8) *nu seke tera ki-yo-n*
 2 good dance do-CONGR-INT[C]
 'Did you dance well?'
- (9) *ya/nu Tsachi jo-e*
 3/2 Tsachi be-DCL[D]
 'He/you are a Tsachi' (Dickinson, 1999:30)

In questions, conjunct marking is used if there is a second person subject, as in example (8), and disjunct marking is used if there is a third person subject, as in example (10):

- (10) *ya seke tera ki-n*
 3 good dance do-INT[D]
 'Did he/she dance well?' (Dickinson, 1999:30)

However questions with a first person subject have a separate third form in Tsafiki, the Noncongruent (glossed with [C*] here), as in sentence (11):

- (11) *la seke tera ki-i-n*
 1M good dance do-NCONGR-INT[C*]
 'Did I dance well?' (Dickinson, 1999:30)

Cha'palaachi likewise uses conjunct versus disjunct in statements to indicate first versus non-first (examples (12) versus (13)), and uses the conjunct in second person questions (as in example (14)).

- (12) *de-ca-yu*
 PL-take-CONJ[C]
 'We took'
- (13) *ca-ve*
 take-DISJ[D]
 'He/she took, you took'

- (14) *naaju chu-yu?*
 how be-CONJ[C]
 ‘How are you?’ (Vittadello, 1988:60,62)

While no examples are given, Vittadello (1988) also notes that Cha’palaachi has a special third form used for first person in questions, like the Tsafiki Noncongruent form in example (11).

In Guambiano, statements are marked with Locutor singular subject or Locutor plural subject marking (see examples (15) and (16)) if they have a first person subject, and are marked with a Nonlocutor suffix if they have a non-first person subject (see examples (17) and (18)), the standard conjunct versus disjunct pattern with the additional complication of number marking in conjunct forms.

- (15) *srusró ku-r*
 young be-LOCT:SING[C]
 ‘I am young’
- (16) *nu pisu-yu pir-ér*
 big pool-LOC bathe-LOCT:PL[C]
 ‘We bathed in the lake’
- (17) *ye-wan ma-n*
 potato-AC eat-NO:LOCT[D]
 ‘You (plural) ate the potato’
- (18) *machik misak ka-n*
 Paez person be-NO:LOCT[D]
 ‘He is a Paez’ (Vásquez de Ruiz, 1988:99)

Unfortunately at this stage it is not known how questions are formed in Guambiano; thus it is not certain that the Locutor suffixes (conjunct) are used in questions with second person subjects, as would be expected in a conjunct/disjunct system.

Thus it can be seen that all Barbacoan languages make a distinction between forms used (in certain contexts) for first person subjects versus non-first person subjects, and in at least all languages except Guambiano, for which no data is available, the form used in a statement with a first person subject is used in a question with a second person subject — that is, there is a conjunct/disjunct system in all these languages.

3. No contrastive marking

Conjunct/disjunct marking is restricted to finite, nondirective clauses. Nonfinite forms never have a conjunct/disjunct suffix, and imperative and hortative suffixes are inherently marked for person (second and first person respectively) and do not participate in the conjunct/disjunct system.

While the basic marking of finite nonimperative clauses in all Barbacoan languages is a conjunct/disjunct contrast, each language has certain tense, aspect or mood forms which do not make this distinction.

For example, in Awa Pit there is a verbal aspect suffix, the Terminative, which may appear without tense marking. In this case, no conjunct or disjunct suffix occurs on the verb:

- (19) *pwitaylap-ti*
 tell.a.lie-TERM
 'I, you, he/she has lied' (Curnow, 1997:234)

Another example can be seen from Guambiano, where there are certain aspectual forms indicated through nominalizations and auxiliary verbs which do not show the conjunct/disjunct distinction. In contrast to the Awa Pit example above, in the Guambiano examples the disjunct suffix, the Nonlocutor, is used in these cases regardless of person, but with no possibility of contrast with the conjunct Locutor form:

- (20) *na ye-wan ma-ik kə-n*
 I potato-AC eat-NOMI be-NO:LOCT[D]
 'I ate the potato' (Vásquez de Ruiz, 1988:122)

Each of the Barbacoan languages has such contexts where no distinctive conjunct/disjunct marking is employed even in finite nonimperative clauses, although the contexts are different in each language.

4. Locutor object contexts

To this point, only locutor versus nonlocutor subject contexts have been considered. In some of the Barbacoan languages, however, a locutor object is also marked on the verb.

In Awa Pit, the marking associated with locutor objects depends upon the tense marking used. If there is no Past tense suffix (regardless of time reference), then in locutor subject and locutor object contexts alike the Conjunct suffix *-s* is used to indicate that there is a locutor argument, as in examples (21) and (22), in contrast to the Disjunct suffix *-y* which was seen in sentence (2). Thus both sentences (21) and (22) contain a verb-final Conjunct suffix, although in the first there is a first person subject while in the second there is a first person object.

- (21) *na=na pala ku-mtu-s*
 1SG.(NOM)=TOP plantain eat-IMPV-CONJ[C]
 'I am eating plantains.'
- (22) *na-wa=na Santos tittu-mtu-s*
 1SG-ACC=TOP Santos spy:on-IMPV-CONJ[C(*)]
 'Santos is spying on me.' (Curnow, In press)

If there is a Past tense morpheme in an Awa Pit sentence, then different suffixes are used, either the Conjunct Subject *-w* (glossed [C]) when there is a locutor subject, as in example (23), or the Conjunct Undergoer *-s* (glossed [C*]) when there is a locutor object, as in sentence (24).³ These two forms contrast with the Past Disjunct *-zi*. Thus unlike in the preceding two sentences, where a first person subject and a first person object were not differentiated through verbal marking, the difference in the role of the first person arguments in the following pair of sentences is marked by verb-final suffixes:

- (23) *nash-na=kima na=na kal ki-mtu-ata-w*
 afternoon-INF=until 1SG.(NOM)=TOP work do-IMPF-PAST-CONJ:SUBJ[C]
 ‘I was working till late.’
- (24) *Juan=na (na-wa) izh-ti-s*
 Juan=TOP (1SG-ACC) see-PAST-CONJ:UNDER[C*]
 ‘Juan saw me.’ (Curnow, In press)

Marking of locutor objects is thus obligatory in Awa Pit — if there is a locutor object, the verb must be marked either with Conjunct or Conjunct Undergoer suffixes, depending on tense. In contrast, such marking of locutor objects is not found in Tsafiki (Connie Dickinson, personal communication). The situation in the other two languages is not clear, however there is a suggestion that such marking is found in Guambiano. As was seen above in examples (15)-(18), the presence of a locutor subject is marked on the verb with the Locutor suffix (singular or plural) in Guambiano, while sentences without locutor elements are marked with the Nonlocutor suffix. While never explicitly addressed in descriptions of Guambiano, in the material published about this language any sentence which contains a locutor object is marked with what is termed the Applicative suffix (glossed with [C*] here), as in the following example.⁴

- (25) *Maria-i wera-pe tsa-tan*
 María-GEN dog-TOP bite-APL[C*]
 ‘María’s dog bit me’ (Triviño Garzón, 1994:615)

Thus in at least some Barbacoan languages, though not in others, special morphology is used in locutor object contexts, not just in locutor subject contexts.

5. Inclusion of affected locutor participants

We have seen that locutor subjects require special conjunct marking in all Barbacoan languages, and the presence of a locutor object is shown with special verb marking in at least two of the languages through the (non-Past) Conjunct or the (Past) Conjunct Undergoer suffixes in Awa Pit and the Applicative suffix in Guambiano. These suffixes can actually be used in more extended contexts, when a locutor entity is somehow affected by the action or state expressed in the clause. This occurs in a variety of

contexts, particularly related to physical or psychological states, and also with weather verbs, as can be seen in the following Awa Pit and Guambiano examples:

- (26) *kerosín way-a-s*
 petrol lack-PAST-CONJ:UNDER[C*]
 ‘Petrol was lacking to me.’
- (27) *alu ki-ma-ti-s*
 rain do-COMP-PAST-CONJ:UNDER[C*]
 ‘(I was on my way to bathe,) it rained on me.’
- (28) *pina us a-ti-s*
 very heavy be-PAST-CONJ:UNDER[C*]
 ‘I found [the bag] very heavy.’
- (29) *pina ii ki-mtu-s*
 very be:hot do-IMPF-CONJ[C(*)]
 ‘I feel it’s very hot.’ (Curnow, In press)
- (30) *sre mur pu-tan*
 rain hard arrive-APL[C*]
 ‘It’s raining hard’ (Triviño Garzón, 1994:615)

This use of special marking to indicate that a locutor entity was affected, even if not an argument, is also possible in Tsafiki, at least with weather verbs. These verbs may optionally be marked with the third term in the Tsafiki conjunct/disjunct system, the Noncongruent suffix (Connie Dickinson, personal communication).

It is important to point out that the use of marking indicating that a locutor participant was affected, even though not an argument of the verb, is entirely optional — all of these sentences could simply be marked with disjunct morphology, and report the same event, although this would of course affect the interpretation of the sentence, giving a neutral reading rather than indicating that the event or state affected a locutor element. This contrasts with the use of the same marking when a locutor element is an object, when the disjunct marking cannot be used in Awa Pit nor apparently in Guambiano.

6. Contrasts in locutor subject contexts

Thus far the description of the conjunct/disjunct system in Barbacoan languages has implied that if there is a locutor subject, the conjunct suffix is necessarily used. In fact this is not entirely the case, at least in Tsafiki and Awa Pit, the best described languages.

In Tsafiki, a locutor subject is normally indicated with the use of the Congruent suffix (conjunct), as was seen in examples (7) and (8). However in certain situations a Noncongruent form may be used instead; this form is otherwise used for a first person subject in a question, as in example (11), or with weather verbs, as noted in section 5.

The precise usage of the Noncongruent rather than the Congruent is complex (see Dickinson, 1999; 2000 for details), but it is used in ‘unusual’ situations which go against speaker expectations, such as to indicate nonvolitionality, surprise, irony, and so on.

- (31) *la ya=ka machite=chi pore-yo-e*
 1M 3=ACC machete=INSTR cut-CONGR-DCL[C]
 ‘I cut him (intentionally) with the machete’
- (32) *la ya=ka machite=chi pore-i-e*
 1M 3=ACC machete=INSTR cut-NCONGR-DCL[C*]
 ‘I cut him (unintentionally) with the machete’ (Dickinson, 1999:32)
- (33) *tse jida jo-yo-e*
 1F cold be-CONGR-DCL[C]
 ‘I’m cold’
- (34) *la jida jo-i-e*
 1M cold be-NCONGR-DCL[C*]
 ‘I’m cold! (e.g., skin cold to touch, but I don’t feel it)’ (Dickinson, 1999:34)

In Awa Pit, intransitive verbs with locutor subjects where the subject is a patient rather than an agent may be marked with either the Conjunct Subject or the Conjunct Undergoer forms if the sentence contains a Past tense suffix.⁵ The difference in interpretation between the two types of marking, seen in (35) and (36), is unknown.

- (35) *mayŋ-ma-ta-w*
 lose:consciousness-COMP-PAST-CONJ:SUBJ[C]
 ‘I lost consciousness.’
- (36) *mayŋ-ma-ti-s*
 lose:consciousness-COMP-PAST-CONJ:UNDER[C*]
 ‘I lost consciousness.’ (Curnow, In press)

7. Summary

This paper has explored the verbal marking systems of the Barbacoan languages. All can be seen to have a conjunct/disjunct system of marking, with a contrast between the marking of agentive, volitional statements with first person subjects (conjunct) versus those which have no first person participant (disjunct); the conjunct form is also used to mark questions with a second person subject. However all Barbacoan languages also have a third term in the system, various called the Conjunct Undergoer (Awa Pit), Applicative (Guambiano) or Noncongruent (Tsafiki).

The three terms in each of the three best described Barbacoan languages have slightly different ranges of use.

Awa Pit: The conjunct forms (Conjunct Subject in the Past tense, corresponding uses of the Conjunct in non-Past tenses) are used with verbs which have locutor subjects, including psychological state verbs. The third term in the system, glossed [C*] (Conjunct Undergoer in the Past, corresponding uses of the Conjunct in non-Past tenses), is used with verbs which have locutor objects, and may also be used to indicate that there is an affected locutor participant with verbs such as weather verbs and psychological or physical state verbs. The disjunct forms (Disjunct) are used to indicate that there are no locutor participants in an utterance.

Tsafiki: The conjunct forms (Congruent) are used with verbs which have locutor subjects, in 'normal' situations. The third term (Noncongruent) is used with verbs which have locutor subjects in 'abnormal' situations, such as when the action was accidental, surprising to the speaker, or the speaker is indicating irony; they may also be used on weather verbs; and they are used in questions which have first person subjects. The disjunct forms (unmarked) are used to indicate that there are no locutor participants in an utterance.

Guambiano: The conjunct forms (Locutor) are used with verbs which have locutor subjects. (In fact, the use of conjunct forms in questions with second person subjects has not yet been established for Guambiano.) The third term (Applicative) is used with verbs which have locutor objects, and may be used with psychological or physical state verbs and weather verbs. The disjunct forms (Nonlocutor) are used to indicate that there are no locutor participants in an utterance.

Thus there are somewhat different contexts of use for each of the three forms in the Barbacoan languages. In all the languages, however, it is clear that there is an underlying opposition between locutor and nonlocutor, that is between first person in statements and second person in questions on the one hand, and other participants on the other. Of course, whether this distinction is strictly one of person (as is assumed, for example, in Curnow, 1997), or is perhaps better treated as a side-effect of an opposition between other speaker- or hearer-centered phenomena (such as the mirativity distinction discussed in Dickinson, 2000) may depend on the individual language, and is definitely a subject for further intense work in each of the Barbacoan languages.

Notes

¹ This paper was written while the author was on an Australian Research Council Postdoctoral Fellowship.

² Throughout this paper, the original glosses from the various sources are used, with the exception of Cha'palaachi, where glosses have been added. All translations from Spanish in examples are my own. As different labels have been used in different descriptions for the phenomenon under consideration, a [C], [D] or [C*] has been added at the end of glosses to indicate whether the sentence contains a conjunct form, a disjunct form or the third term in the system, to simplify comparisons and assist the reader.

³ The choice of the term Conjunct Undergoer, a semantic role, contrasting with Conjunct Subject, a grammatical relation, will become clear below, where it will be seen that the Conjunct Undergoer can indicate simply that a locutor participant was affected by the action, even though it was not an argument of any sort. The Conjunct Subject, in

contrast, only occurs when there is a locutor participant in the subject grammatical relation.

⁴ In some sources (e.g., Vásquez de Ruiz, 1988), the Applicative suffix is treated as a unit, *-tan* and is called the Patientive suffix; in more recent works (e.g., Triviño Garzón, 1994; Vásquez de Ruiz, 1992), the Applicative is simply *-t*, and is followed by the Nonlocutor (disjunct) *-an* (and is glossed either as Patientive or Applicative, depending on the author). The example glosses here have been unified in favor of the first analysis, although the choice is immaterial to the point made here.

⁵ The possibility of such a contrast is not available in non-Past sentences, as the Conjunct form *-s* is used in the non-Past corresponding to both the Conjunct Subject and the Conjunct Undergoer.

References

- Curnow, Timothy Jowan. 1997. A grammar of Awa Pit (Cuaiquer): An indigenous language of south-western Colombia. PhD, Australian National University.
- Curnow, Timothy Jowan. In press. Conjunct/disjunct marking in Awa Pit. *Linguistics*.
- Curnow, Timothy Jowan and Liddicoat, Anthony J. 1998. The Barbacoan languages of Colombia and Ecuador. *Anthropological Linguistics* 40:384-408.
- Dickinson, Connie. 1999. Semantic and pragmatic dimensions of Tsafiki evidential and mirative markers. *Chicago Linguistic Society* 35:29-44.
- Dickinson, Connie. 2000. Mirativity in Tsafiki. *Studies in Language* 24:379-421.
- Hale, Austin. 1980. Person markers: Finite conjunct and disjunct verb forms in Newari. *Papers in South-East Asian Linguistics* 7, ed. by Ronald Trail, 95-106. (Pacific Linguistics Series A-53.) Canberra: Pacific Linguistics, Australian National University.
- Hargreaves, David. 1990. Indexical functions and grammatical sub-systems in Kathmandu Newari. *Chicago Linguistic Society* 26:179-93.
- Hargreaves, David. 1991. The conceptual structure of intentional action: Data from Kathmandu Newari. *Proceedings of the Berkeley Linguistics Society* 17:379-89.
- Triviño Garzón, Lilia. 1994. Hacia una tipología de la predicación de la oración simple en la lengua gumabiana. *Bulletin de l'Institut Français d'Études Andines (Lima, Peru)* 23:601-18.
- Vásquez de Ruiz, Beatriz. 1988. La Predicación en Guambiano. (Lenguas Aborígenes de Colombia, Descripciones 2.) Bogotá: Centro Colombiano de Estudios en Lenguas Aborígenes, Universidad de los Andes.
- Vásquez de Ruiz, Beatriz. 1992. La modalidad epistémica en guambiano (Guambía). II Congreso del CCELA, ed. by Tulio Rojas Curieux, 209-18. (Lenguas Aborígenes de Colombia, Memorias 2.) Bogotá: Centro Colombiano de Estudios en Lenguas Aborígenes, Universidad de los Andes.
- Vittadello, P. Alberto. 1988. Cha'palaachi: El Idioma Cayapa, Tomo 1: Gramática. (Miscelánea Antropológica Ecuatoriana, Serie Monográfica 5.) Esmeraldas, Ecuador: Pontificia Universidad Católica del Ecuador and Museos del Banco Central del Ecuador.

Topic, focus, definiteness, specificity, and referential prominence in Western Apache noun phrases and relative clauses

Willem J. de Reuse
University of North Texas

Western Apache (henceforth WA) is a Southern Athabaskan or Apachean language spoken by approximately 13,000 people on and around five reservations in central and east central Arizona.¹

WA has a well developed system of enclitics marking a variety of pragmatic or semantic distinctions on noun phrases. This system is rather different from the determiner system of Navajo, a well-studied language closely related to WA (Willie 1991). It is possible to find cognates for the WA forms in Navajo, but they seem to be less common in texts, and little discussed in the literature. The four forms making up the WA system of noun phrase marking enclitics (San Carlos dialect) are listed, with glosses, in (1):

- | | | |
|------|--------------------|---|
| (1a) | =(h)i ² | specific |
| (1b) | =(h)í | topic/definite |
| (1c) | =(h)ihí | specific and definite (a sequence of (a) and (b)) |
| (1d) | =go | contrastive focus |

In (1a-c) the (h) is generally present only when the word the enclitic is attached to ends in a vowel or in a nasal. It will be noted that (1c) is really in sequence of (1a) and (1b), and I will argue that its meaning is compositional as well.

These four forms can be added to a noun phrase such as shiye' 'my son (man speaking)' as in (2):

- | | | |
|------|-----------|--|
| (2a) | shiye'i | 'a son of mine, a certain son of mine' |
| (2b) | shiye'í | '(regarding) my son' |
| (2c) | shiye'ihí | 'a particular son of mine, one of my sons' |
| (2d) | shiye'go | 'my <u>son</u> (as opposed to other people)' |

The implication of (2a) is that the speaker has more than one son, and it is clear to the speaker, but not to the hearer, which one of the sons is referred to. In other words, the enclitic =(h)i makes shiye' specific. It can often be translated as 'a certain'. A noun is specific if it refers to someone or something in particular in the speaker's mind. In the literature, other terms for what I call "specific" are "extensional", "de re", or "referential". "Specific" is opposed to "non-specific", also called "intensional", "de dicto" and "non-referential" (Lyons 1999:166-78). From Lyons' thorough account, it also appears that there is no correlation between definiteness and specificity, i.e. specific nouns can be definite or indefinite, and non-specific nouns can also be definite or indefinite. Used by itself, the enclitic =(h)i generally marks indefinite specific nouns, but when it combines with (1b) into (1c), it marks definite specific nouns.

The implication of (2b) is that the speaker has one son only, and therefore it is clear to the hearer which son is referred to. In other words, the enclitic =(h)í makes shiye' definite. A noun is definite if it refers to something or someone identifiable by the hearer. From Lyons (1999), one learns that there is more to definiteness than the notion of identifiability, but since the above definition seems to work for WA, and is used in other reference works such as Payne (1997:261-4), I will adopt it here. Definite nouns are often specific, but they can be non-specific, as we will see under (8c-d). It should also be noted that the most common function of =(h)í is not definiteness marking, but topicalizing. Indeed, =(h)í is often better translated as a topicalizing element. As a

topicalizing element, it can mark a change of topic as well as simply mark the topic. An example of a change of topic = (h)í occurs in the first line of the question and answer pair in (3):

- (3) Shinabil gólǫ́. Nihí?
 my.car it.exists you = TOP³
 'I have a car. What about you?'

Shī aídó', shinabil gólǫ́.
 I also my.car it.exists
 'Me too, I have a car.'

The contrasting sentences under (4) illustrate = (h)í as a topic marker:

- (4a) Góséhí ápos yíyāā.
 dog = TOP apple it.ate.it
 'The dog ate the apple.' or: 'Speaking about the dog, it ate the apple.'

- (4b) Gósé áposí yíyāā.
 dog apple = TOP it.ate.it
 'The dog ate the apple.' or: 'Speaking about the apple, the dog ate it.'

It is impossible however, to have the enclitic on both nouns, as in (5):

- (5) *Góséhí áposí yíyāā.
 dog = TOP apple = TOP it.ate.it

The ungrammaticality of (5) makes sense, since it is infelicitous to have two topics in one clause, whereas there are no restrictions against two definite nouns in one clause. As in Navajo (Willie 1991), a definiteness interpretation is preferred for nouns without any enclitics at all. This interpretation can be cancelled in several ways. I assume that the definiteness usage of = (h)í is derived from the topicalizing usage, since it would not be possible to explain the restriction against two = (h)í in one clause if = (h)í was primarily a definiteness marker.

I now turn to (2c). The implication of (2c) is that the speaker has more than one son, and that one of these sons has been picked out already as the topic of conversation, so that is clear to the speaker, as well as to the hearer, which one of the sons is referred to. In other words, the enclitic = (h)ihí marks shiye' as specific as well as definite. Thus, the meaning of = (h)ihí is compositional, since it contains specific = (h)i followed by topic/definite = (h)í.⁴

Finally, let us look at the marking with enclitic =go in (2d). The enclitic =go has nothing to do with specificity, topicalization, or definiteness. It contrasts the noun phrase my son with other people in discourse. Thus, =go indicates contrastive focus, as first noted by Potter (1997).⁵ The enclitic =go has several other more important functions, which I will return to later. The examples in (6) contrast shiye' and shiye'go in sentences.

- (6a) Shiye' hish'ǫ́ ni'.
 my.son I.see.him PAST
 'I saw my son.'

- (6b) Shiye'go hish'ḡḡ ni'.
 my.son=FOC I.see.him PAST
 'It is my son I saw (and not someone else).'

The four enclitics discussed so far are noun phrase markers. Since relative clauses and their heads are also noun phrases, it is not surprising that they occur on relative clauses. They are then attached at the right boundary of relative clauses, which, following the extensive literature on the subject in Navajo, I will call internally headed. The syntax of WA relative clauses is very similar to that of Navajo. It is the profusion of the enclitics on WA relative clauses which is remarkable. It is worth emphasizing that the enclitics, when occurring on relative clauses, look like relative clause markers, but since they can also occur on any noun phrase, they are not relativizers.

I will now set up a scenario for discussing examples of relative clauses. Suppose that speaker A, a car salesman, sees B, a potential buyer, walking along on his car lot.⁶ Then A might ask B something like (7):

- (7) Hant'é biká hañtāā?
 what for.it you.look
 'What are you looking for?'

B might answer with any of the grammatical utterances under (8):

- (8a) Nabil ła' nahishñiih hásht'íni⁷ biká hanshtāā.
 car INDEF I.buy.it I.want=SPEC for.it I.look
 'I am looking for a car I want to buy.'

(8a) is used when B already has a certain car in mind, and would be looking for that car on the car lot, and A would not know which car that is. Note the element ła', translatable as 'a' or 'some' which marks the noun nabil 'car' as indefinite.

- (8b) Nabil nahishñiih hásht'íni biká hanshtāā.
 car I.buy.it I.want=SPEC for.it I.look
 'I am looking for this car I want to buy.'

(8b) is similar to (8a); here the noun nabil is not marked as indefinite by ła'. The translation includes the colloquial use of this, which here has an indefinite specific reading (Lyons 1999:176-7). Another, even more colloquial translation provided by one of my consultants is: 'There is this car that I want to buy ...' I assume that the absence of ła' results in some sort of "referential prominence" (Lyons 1999:176) or "discourse referentiality" (Givon 1979, Du Bois 1980), but without making the noun phrase definite.

There are problems with the interpretation of (8c) and (8d), which I will discuss together.

- (8c) Nabil ła' nahishñiih hásht'ḡḡhí biká hanshtāā.
 car INDEF I.buy.it I.want=TOP for.it I.look
 'I am looking for a car I want to buy.'

- (8d) Nabil nahishñiih hásht'ḡḡhí biká hanshtāā.
 car I.buy.it I.want=TOP for.it I look
 'I am looking for the car I want to buy.'

In (8c), the head noun phrase is non-specific, and apparently indefinite, because of *ła'*. Therefore, there is an apparent conflict between the *=(h)í* at the end of the relative clause, which I assume marks the head as definite, and this *ła'* on the head noun phrase, which marks it as indefinite. It is unlikely that this conflict can be resolved along the lines of Williamson's (1987:185-7) account of an indefinite restriction for relative clauses in Lakota, a genetically unrelated but typologically similar language. In the corresponding Lakota relative clauses, the internal head, although marked with an indefinite article, is always translated as an English definite. However, the internal head in (8c) is translated as an indefinite. Furthermore, as seen in (8d), WA relative clauses with *=(h)í* and without the indefinite *ła'* on the internal head are also possible, and here the head is translated as a definite. This is evidence that potential parallels between Lakota relative clauses and (8b) are not helpful. The problem with the heads in (8c) and (8d) is that they both appear to be formally non-specific as well as definite, since they are marked by *=(h)í*, but it at the same time they have different translations. I suggest that in both (8c) and (8d), the speaker does not have a particular car in mind, which makes the head non-specific, and that in both cases, the speaker assumes that A has one in mind, which makes the head definite. After all, car salespeople always have a particular car in mind that is just right for you! My hypothesis about the difference between (8c) and (8d), is that in (8c) there is just an assumption on the part of A, whereas in (8d), B is quite sure that A has a particular car in mind for him, possibly because B remembers that A discussed a particular car with him earlier, but B can't now remember which car that was. This difference might well trigger the difference in English translations.

I assume then that the presence of *ła'* in (8c) does not result in making the head indefinite, but rather marks the absence of the referential prominence discussed with regard to (8b). The presence of *ła'* in (8c) is thus the converse of the absence of *ła'* in (8b), which does not result in making the head definite, but marks referential prominence. If this hypothesis is correct, the gloss of *ła'* as *indefinite* is inaccurate since it actually marks something like "lack of referential prominence". I will retain the gloss as a convenient label.

(8e-f) are much less problematic. For (8e-f) one has to assume a situation where A and B have met before, and where both A and B know which car in the lot B wants to buy. If B comes to see A again, B might utter (8f). The head noun phrase is both definite and specific. Since the combination of specificity and definiteness on this head naturally implies referential prominence, *ła'* is impossible, and (8e) is ungrammatical.

(8e) *Nabil *ła'* nahishñiih hásh't'ínihí biká n'íyāā.
 car INDEF I.buy.it I.want=SPEC=TOP for.it I.came

(8f) Nabil nahishñiih hásh't'ínihí biká n'íyāā.
 car I.buy.it I.want=SPEC=TOP for.it I.came
 'I came for the car I want to buy.'

Let me comment at this point on the frequency of sentences such as (8a-f). These sentences are not common in texts, and the examples in (8a-f) were elicited.⁸ They were verified by three different speakers, but the fact remains that relative clauses utilizing the three enclitics illustrated so far are rare in texts. Apparently, the WA relativization strategies illustrated with (8a-f) are not the favorite and most common way of relativizing. The most common way is to use the enclitic *=go*, which here does not have a focusing function. Indeed, *=go* is most often, following the analysis by Schaubert (1979:226) for Navajo, an almost semantically empty subordinator. An example with *=go* in a relative clause is (8g).

- (8g) Nabil ɭa' nahishɲiih hásht'ɲgo biká hanshtāā.
 car INDEF I.buy.it I want=SUB for it I.look
 'I am looking for a car I want to buy.' (or:)
 'Wanting to buy a car, I am looking for one.' (or:)
 'I want to buy a car and I am looking for one.'

I am not sure whether it is possible to remove ɭa' in such sentences, since my consultants have contradictory opinions about a sentence like (8h).

- (8h) *?Nabil nahishɲiih hásht'ɲgo biká hanshtāā.
 car I.buy.it I want=SUB for it I.look

The strategy illustrated by (8g) is the most common way of relativizing in WA. In his account of relativization in the White Mountain dialect of WA, Greenfeld provides this structure only, implying that it is the normal way of relativizing (1999:362-3). However, as one can see from the translations under (8g), such sentences are not necessarily translatable as relative clauses, and might not have embedded relative clause structures either. There are other languages in the world that appear to have similar subordinating structures, sometimes translatable as relative clauses, and sometimes not. Hale (1976) discovered a type of clause in Walbiri, an Australian language, which he called the "adjoined relative clause". (9) is a Walbiri example.

- (9) ɲatjulu-ɭu ø-ɲa yankiri pantu-ɲu,
 I-ERG AUX emu spear-past

 kutja-lpa ɲapa ɲa-ɲu.
 COMP-AUX water drink-past
 'I speared the emu which was/while it was drinking water.' (Hale 1976:78)

In Walbiri, it appears that the relative clause interpretation (i.e. 'which was drinking water' in (9)) is available when the main and subordinate clauses share an identical argument, and that the adverbial clause interpretation (i.e. 'while it was drinking water' in (9)) is available when the time reference is the same. In (9), both interpretations are possible, because the clause satisfies both requirements (Hale 1976:79). These facts are also true for the WA =go construction in (8g). Because of the similarities between the adjoined relative clause of Walbiri and =go clauses, I suggest that the WA =go clauses are adjoined rather than embedded.⁹

To conclude, I have tried to give a semantic and pragmatic description of four enclitics that are common on WA noun phrases. I then pointed out that the three enclitics (=h)i, =(h)í, and =(h)ihí also occur at the right edge of relative clauses. In this position, they have the same pragmatic and semantic functions, and thus are not relative clause markers per se. Finally, I observed that the three enclitics are not common on relative clauses. The embedded types of relative clauses marked with these three enclitics are apparently dispreferred, and WA prefers the adjoined relative clause construction with the enclitic =go. In the course of my treatment of relative clauses, I also suggested a possible function for the marker ɭa' in relative clause heads.

A topic for further study would be the diachronic paths of functional change that some of the enclitics have undergone. I will not comment on =(h)i, which only seems to have one function. =(h)í appears to be primarily topicalizing, and to have acquired an additional definiteness marking function. =go seems to be primarily a subordinator, and it retains this function in adjoined relative clause constructions. I strongly suspect that its function as a contrastive focus marker on noun phrases is a derived one. If that is the case, we would then be observing the functional evolution of

one WA enclitic =(h)í from discourse marker (as a topicalizer) to a grammatical element (a definiteness marker), and at the same time the functional evolution of another WA enclitic =go from a grammatical element (a subordinator) to a discourse marker (of contrastive focus).

Notes

¹The other Apachean languages are Navajo, Chiricahua, Mescalero, Jicarilla, Lipan, and Plains Apache. These languages form a chain, with some mutual intelligibility between geographically contiguous languages, such as between WA and Navajo, and between WA and Chiricahua. There are four major dialects or variants of WA: San Carlos, Cibecue, White Mountain, and Tonto. The WA forms in this paper are from my own fieldwork (1992-present) on the San Carlos dialect. The White Mountain and Cibecue dialects appear to have a more complicated system with an additional human/non-human dimension. The facts regarding topic, focus, definiteness, and specificity are the same in all four dialects. The research leading to this paper was funded by a grant from the National Science Foundation to the University of Arizona (Nr. SBR-9408543), and to the University of North Texas (Nr. SBR-9896227). This support is hereby gratefully acknowledged. I am very grateful to the following San Carlos consultants for their help and patience: Phillip Goode (deceased), Joycelene Johnson, and Kathleen Kitcheyan. I thank the participants at the WAIL/SSILA Summer 2001 Conference in Santa Barbara, in particular Sasha Aikhenvald, Wally Chafe, Bob Dixon, and Michael Krauss, for their comments.

²The spelling of WA is the standard system used and described in White Mountain Apache Culture Center (1972:vii-xii, 107-110) and in Bray (1998:xii-xviii). This standard system only distinguishes high tone (acute accent) and low tone (no accent); the mid tone is usually marked as a low tone, but occasionally it is marked as high tone (acute accent), particularly on a short i or a long ii, or on a long oo (then often written ú). Because of the unsatisfactory marking of the mid tone in the standard system, the following addition to the system is used in this paper: Low tones are unmarked *v*, *vv*, mid tones are marked *v̄*, *v̄v̄*, and high tones are marked *ú*, *úú*. Underlined n stands for a sound varying between [n] (most San Carlos subdialects) and prenasalized voiced [ʎd] (some San Carlos subdialects). See de Reuse, Goode et al. (2001) for a grammar consistently using these additional conventions.

³Abbreviations used in glosses and morphological analyses are:

AUX	auxiliary
COMP	complementizer
ERG	ergative case
=FOC	contrastive focus enclitic
=SUB	subordinating enclitic, homonymous with =FOC
INDEF	indefinite marker
PAST	asserted past particle
=SPEC	specific enclitic
=SPEC=TOP	specific enclitic followed by topic/definite enclitic
=TOP	topic/definite enclitic

⁴Navajo scholars will note that the sequence =(h)ihí looks like the Navajo enclitic =íggí, which might well be cognate with =(h)ihí. I do not know whether a similar compositional analysis is possible for Navajo.

⁵This enclitic marks contrastive focus, as defined by Chafe (1976), only. I am aware of the fact that the term "focus" is used in many other ways, which are irrelevant here.

⁶One also has to suppose, of course, that they are both WA speakers, and that they expect each other to speak and be spoken to in WA.

⁷There is a morphophonological rule by which *hasht'ḡ* plus *=(h)i* becomes *hásht'íni*. That is, *=(h)i* causes a preceding long nasalized mid-toned vowel to become a short oral high-toned vowel followed by *-n-*. The enclitic *=(h)í* does not cause this change.

⁸One reason why such sentences are rare might be related to processing problems due to embedding. One way of avoiding embedding is to right-dislocate the relative clause, as in the sentence below, which some consultants prefer to the embedded (8f):

Nabil biká nfyāā, nahishḡiih hásht'ínihí
 car for.it I.came I.buy.it I.want = SPEC = TOP
 'I came for the car, the one I want to buy.'

⁹I do not know whether there are in Navajo any clauses with *=go* that can be interpreted as relative clauses. The literature on Navajo relative clauses does not mention such *=go* clauses.

References

- Bray, Dorothy, ed., in collaboration with the White Mountain Apache Tribe 1998. *Western Apache-English Dictionary. A Community-Generated Bilingual Dictionary*. Tempe, Az.: Bilingual press/Editorial Bilingue.
- Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. *Subject and Topic*, ed. by Charles N. Li, 25-55. New York: Academic Press.
- Du Bois, John W. 1980. Beyond definiteness: the trace of identity in discourse. *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*, ed. by Wallace L. Chafe, 203-74. Norwood, New Jersey: Ablex.
- Givon, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.
- Greenfeld, Philip J. 1999. Some Special Morphological Characteristics of the White Mountain Dialect of Apachean. *Anthropological Linguistics* 41.356-381.
- Hale, Kenneth L. 1976. The adjoined relative clause in Australia. *Grammatical Categories in Australian Languages*, ed. by R. M. W. Dixon, 78-105. (Linguistic Series No. 22, Australian Institute of Aboriginal Studies.) New Jersey: Humanities Press.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Payne, Thomas E. 1997. *Describing morphosyntax. A guide for field linguists*. Cambridge: Cambridge University Press.
- Potter, Brian C. 1997. *Wh/Indefinites and the Structure of the Clause in Western Apache*. Ph. D. dissertation, Department of Linguistics, University of California, Los Angeles.
- de Reuse, Willem J., with Phillip Goode et al. 2001. *A Practical Grammar of the San Carlos Apache Language*, viii + 674 pp. Unpublished manuscript, Denton, Texas.
- Schauber, Ellen. 1979. *The Syntax and Semantics of Questions in Navajo*. New York and London: Garland Publishing.
- White Mountain Apache Culture Center. 1972. *Western Apache Dictionary*. Fort Apache, Arizona: White Mountain Apache Culture Center.
- Williamson, Janis S. 1987. An Indefiniteness Restriction for Relative Clauses in Lakota. *The Representation of (In)definiteness*, ed. by Eric J. Reuland and Alice G.B. ter meulen, 168-190. (Current Studies in Linguistics Series 14.) Cambridge: MIT Press.
- Willie, MaryAnn. 1991. *Navajo Pronouns and Obviation*. Unpublished Ph. D. dissertation, Department of Linguistics, University of Arizona.

Secondary Palatalization and Changes in Vowel Formants in Isthmus Mixe*

Julia Dieterman

The University of Texas at Arlington

SIL International

1.0 Introduction

Although the effects of palatalization on consonants have been documented in various languages, discussion of the effects of palatalized consonants on adjacent vowel qualities is usually limited to the description of a transitional [i]-like on-glide or off-glide. As such, the question this study endeavors to answer is this: do the palatalizing effects of palatalized consonants change the steady-state formants of adjacent vowels in comparison to the steady-state formants of vowels adjacent to plain consonants? If there are changes, which formants are affected and how much are they affected?

Isthmus Mixe is an ideal language to research for information about palatalization. In Isthmus Mixe every consonant phoneme may be palatalized by the morpheme marking either third person possession on nouns or third person subject on verbs. This morpheme affixes to the initial consonant of the word causing it to become palatalized: for example, /pa²m/ 'illness' and /p¹a²m/ 'her illness'; /hot/ 'liver/center' and /h¹ot/ 'his liver/center'. This type of palatalization manifests itself most clearly as the addition of a high front tongue position that occurs simultaneously with the primary consonantal articulation. As shall be demonstrated, however, this morphemic palatalization has a measurable fronting effect on the vowel immediately following the palatalized consonant.

1.1 Isthmus Mixe language and people

Isthmus Mixe, also known as Eastern Mixe or Guichicovi Mixe,¹ is a Mesoamerican language of the Mixe-Zoque family spoken in southern Mexico in the state of Oaxaca. It is primarily an oral language, with some religious and governmental documents having been translated with the assistance of bilingual speakers. In 1990 there were reported to be 20,000 speakers (Grimes 2000).

The Isthmus Mixe people live in the region the Isthmus of Tehuantepec in southern Mexico (Figure 1). Their more numerous and well-known neighbors in the Isthmus are Zapotec speakers. Several authors have suggested that the Mixe-Zoque languages may be traced back to the Olmecs, 1500 BCE, an ancient civilization in southern Mexico (Campbell and Kaufman 1976, Hiltz 1999, Stross 1985). However, Wichmann (1995), who has done extensive studies in Proto-Mixe-Zoque, does not believe there is sufficient proof to state that the Mixe-Zoque languages are of Olmec origin.

Although the consonantal palatalization feature as described for Isthmus Mixe is attested in all Mixe languages, there are no published studies of any Mixe language which analyzes and describes vowel formants in relationship to palatalization. Thus the research documented in this paper is the first of its kind.



Figure 1: Map of Mexico
(EnchantedLearning.com: 2001)

2.0 Primary and secondary palatalization defined

Palatalization is a common phonological process in the languages of the world. What is less well known is that there are two major types of palatalization which have been documented that produce different results: 1) palatalization which modifies the main articulation of the consonant itself, called primary palatalization, and 2) the addition of a high front tongue position occurring simultaneously with the primary consonantal articulation, usually referred to as secondary palatalization (Bhat 1974, Ladefoged 1993). It is essential to distinguish between these two major types of palatalization to avoid ambiguity and confusion. Bhat (1974:19) states that palatalization was considered to be a single diachronic or morphophonemic process by linguists and the cover term *palatalization* was commonly used without making important and necessary distinctions as to the type of palatalization process. The terms *secondary palatalization* and *secondary articulation* have also been used by some authors to refer to what is defined as primary palatalization in this paper.

2.1 Primary palatalization

In primary palatalization, the point of articulation of the affected consonant moves toward the palatal region, usually in the presence of front vowels or the palatal consonant/semi-vowel /j/. For example, in English, /k/ in *keep* is more palatal than /k/ in *karma*. English orthography often reflects this difference in using *k* in more palatal environments, e.g. *key*, *kiss*, *keg*, and *c* preceding vowels which do not have a palatalizing effect, e.g. *car*, *cold*, *cut*. Many uses of the letter *k* in nonpalatalizing environments are in words of non-English origin, e.g. *karma* (Sanskrit), *karate* (Japanese). Primary palatalization mainly affects alveolar and velar articulations and consonants located between them.

2.2 Secondary palatalization

In contrast to primary palatalization, secondary palatalization does not change the place of articulation of a segment. Rather, it superimposes a secondary articulation on the main articulation.² In Russian, for example, there is a phonemic contrast between /bratʲ/ 'to take' and /brat/ 'brother'. When palatalization is added to /t/ it is a secondary [i]-like articulation added to the primary dental articulation of the Russian /t/, i.e. the /t/ remains a dental stop. Another Russian example is /mʲatʲ/ 'to crumple' and /matʲ/ 'mother' in which the primary bilabial articulation of the palatalized nasal is still bilabial.

According to Bhat, secondary palatalization usually affects all of the consonants of a language, resulting in contrasting sets of plain and palatalized consonants. Secondary palatalization is much less common than primary palatalization and is areal in occurrence. Languages in which secondary palatalization occurs may have some type of primary palatalization as well (Bhat 1974:36).

In a survey of 317 languages, Maddieson (1984) found that secondary palatalization occurs in about 10% of the languages in the survey. Two Mixe-Zoque languages, Totontepec Mixe and Copainalá Zoque, are included in his survey. According to Hombert and Maddieson (1998) secondary palatalization is one of the distinctive characteristics occurring in few languages which can be used in language recognition classification. Russian is mentioned as a well-known example.

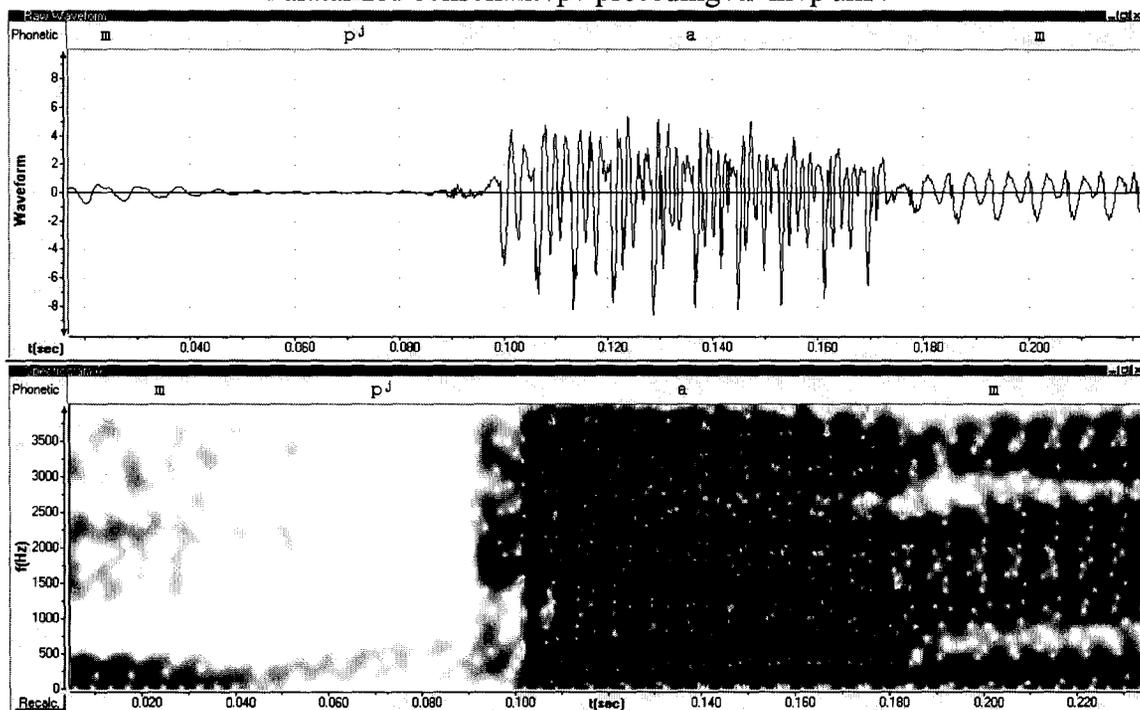
Secondary palatalization is a noted feature in such languages as Irish (Ní Chiosáin 1994), Nenets (Yurak) (Salminen 1999), and the Slavic languages, among which Russian is perhaps the most well documented (Comrie 1981, Fant 1970, Ladefoged 1993).

In a discussion of secondary palatalization in Russian, Ladefoged and Maddieson (1996:364) show the differences among /potʲ/, /pʲotr/ and /pjotʲ/ in comparative spectrograms which reveal:

the distinction between palatalization and a sequence with j. In pʲotr 'Peter' the transition away from the palatal position, indicated by a falling F2, begins immediately on consonant release. In contrast, in pjotʲ 'drinks' there is a short steady state before the transition begins.

Thus we have evidence that there are measurable acoustic correlates to secondary palatalization.

Palatalized consonant /pʲ/ preceding /a/ in /pʲam/.



Plain consonant /p/ preceding /a/ in /pak/.

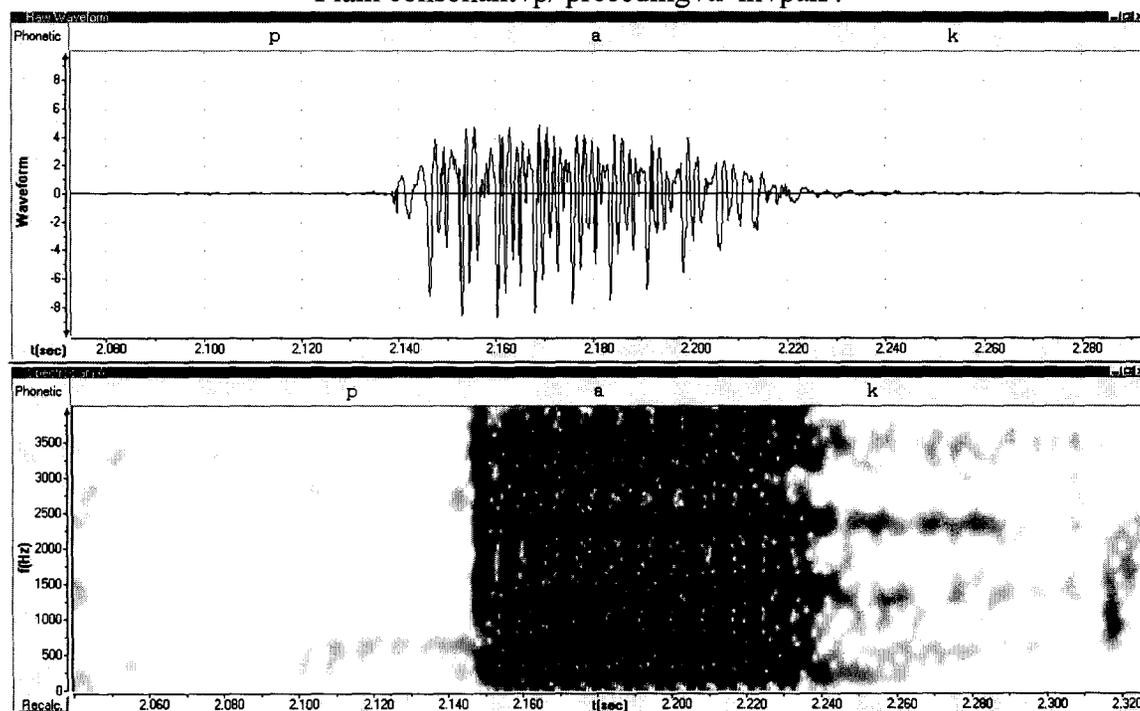


Figure 2: Isthmus Mixe palatalized and plain /p/

2.2.1 Secondary palatalization in Isthmus Mixe

Secondary palatalization is widespread in Isthmus Mixe since the morpheme that indicates third person is palatalization of the initial consonant of the word. This morpheme affixes to the nouns and verbs, which all begin with one consonant followed by one vowel. Isthmus Mixe consonant phonemes are: /p, t, k, ʔ, t̃s, ʃ, h, m, n, w, j/ and the vowel quality phonemes are: /i, i, u, e, o, a/.³ All the consonant phonemes in the language may occur word initial and they may be followed by any of the six vowel qualities.⁴ Therefore, all the consonants and all the vowels may be affected by palatalization in word initial syllables. This palatalization is caused by the addition of the third person morpheme, rather than any phonetic influence, whereas, palatalization in other languages is usually caused by the phonetic influence of an adjacent high front vowel.

In the waveforms and spectrograms shown in Figure 2, it can be seen that there is no phonetic influence causing palatalization of the consonant /p/. There is nothing preceding the palatalized consonant /p^j/ in /p^jam/ to trigger palatalization, nor is the low vowel /a/ which follows /p^j/ a trigger. Palatalization is a manifestation of the third person morpheme affixed to the consonant.

High F2 transitions are seen adjacent to the palatalized consonant. The palatalization is heard as occurring simultaneously with the primary consonantal articulation, and as a short glide transition on the vowel. In contrast, the F2 transition is low following the plain consonant /p/ in /pak/. The steady-state portion of the vowel formants is the level area, where the formants are neither rising nor falling. It is the steady-state portion which is measured and discussed in this study.

3.0 Acoustic analysis of Isthmus Mixe vowels

For this particular acoustic analysis, a selection was extracted from a text read by a native speaker of Isthmus Mixe. The fuller context of the narrative text encompasses the study of the individual elements under consideration. Analysis of this type of data is based on more natural speaking than would be possible using elicited sentences.

3.1 Isthmus Mixe data corpus

The speaker of the data analyzed in this study was an Isthmus Mixe male, age 42, of Mogoñé, Oaxaca, Mexico, tape-recorded in July of 1995. The text was read in normal (unemphatic) narrative style, well articulated but not overly formal. A selection of fluent continuous speech, 100 seconds long, was digitized for analysis. There are a total of 356 syllables which averages 3.56 syllables per second.⁵ All of the vowels were analyzed (except those found in non-Mixe words, such as Spanish names), yielding a total of 300 vowels.

3.2 Frequency counts of Isthmus Mixe vowels

The first analysis shown is the frequency counts of the 300 vowels in the selection of text. The frequency counts of the six vowel qualities in the language, namely /i, i, u, e, o, a/, vary considerably (Table 1 and Figure 3). Each vowel quality is divided by those adjacent to plain consonants, the upper section of the bar (Figure 3), and those adjacent to

palatalized consonants, the lower section of the bar. In referring to vowels adjacent to palatalized consonants, there may be a palatalized consonant preceding the vowel, following the vowel, or both.⁶

Of 300 vowels analyzed, the numbers from highest to lowest frequencies are (Table 1): /i/ with 115, of which 66 are adjacent to plain consonants and 49 are adjacent to palatalized consonants; /a/ with 101, of which 29 are adjacent to plain consonants and 72 are adjacent to palatalized consonants; /ɪ/ with 37, of which 24 are plain and 13 are palatalized; /u/ with 19, of which 14 are plain and 5 are palatalized; /e/ with 14, of which 6 are plain and 8 are palatalized; and /o/ with 14, of which 8 are plain and 6 are palatalized. Two vowels, /i/ and /a/, make up 72% of the total number. Of the 300 vowels analyzed, just over half of the vowels are adjacent to palatalized consonants (51%). Given that all of the vowels in the 100 second sample of continuous narrative text⁵ were included in the analysis, it appears that the effects of palatalization are a major phonetic feature of the language. In contrast to using elicited data, an analysis of continuous speech shows such frequencies as displayed in Figure 3 and gives an interesting overview of the language.

3.3 Methodology

The vowels were analyzed using the Speech Analyzer software developed by the Summer Institute of Linguistics, version 1.5 (10.6), copyright 1996-1998, by JAARS - ITCS, Waxhaw, NC. The descriptions of basic acoustic analysis principles by Baart 1999, Fant 1973, Stevens 1998, Van Summers 1987, and others were used as guides in the analysis. For measuring the Isthmus Mixe vowel formants, the spectrum was calculated over the steady-state portion of the vowel; transition areas from the preceding or following consonants were avoided. Vowel formants F1, F2, and F3 were measured.

Using the Statistical Package for the Social Sciences (SPSS) version 10, Analysis of Variance (ANOVA) and t-tests were performed on the six sets of vowels in the language, i.e. those adjacent to palatalized consonants and those adjacent to plain consonants. Because some vowel qualities have few occurrences in the 300 vowel text used to show the frequencies, additional tokens were analyzed from the text immediately following the 100 second text to bring the total of each group to 30, the number minimally suitable for reliable t-test results. The results are presented in Section 4.0.

TABLE 1: ISTHMUS MIXE VOWEL FREQUENCIES

Vowel	N	Plain	Palatalized
/i/	115	66	49
/a/	101	29	72
/ɪ/	37	24	13
/u/	19	14	5
/e/	14	6	8
/o/	14	8	6
Totals:	300	147 (49%)	153 (51%)

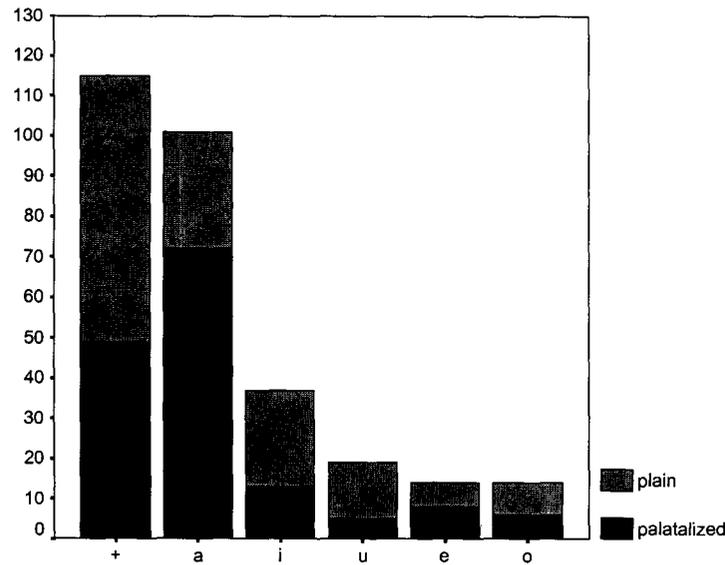


Figure 3: Isthmus Mixe Vowel Frequencies
Raw frequencies of each vowel type in the 100 second sample.

4.0 Results and Discussion

4.1 ANOVA Results

A preliminary series of ANOVAs showed that the effects of the palatalization of either the previous or following vowels were generally independent of the effects of the place of articulation of the previous segment. In Table 2, “PALATALIZATION” refers to a dependent variable with four levels: 1) no adjacent palatalized consonant, 2) a preceding palatalized consonant, 3) a following palatalized consonant, and 4) both preceding and following palatalized consonants. “PLACE OF ARTICULATION” encompasses four levels: 1) labial, 2) alveolar, 3) velar, and 4) glottal.

Significance is set at the 95% confidence level ($p \leq 0.05$); all significant ‘p’ values are starred. In the ‘p’ column under the label PALATALIZATION, the ‘p’ values of every F2 formant value for every vowel quality are significant. Those F3 values⁷ which are significant serve to reinforce the F2 values. There are a few scattered significant ‘p’ values in the PLACE OF ARTICULATION column and the INTERACTION: PALATALIZATION/ARTICULATION column which show no consistent patterning. Generally speaking, the effects of place of articulation of the previous segment are not significant and neither are the interaction effects of palatalization of an adjacent segment and place of articulation of the previous segment together. However, the effects of palatalization are significant for all F2 values. Therefore, since place of articulation did not show a significant difference in formant values in relationship to palatalization, the data were

pooled across place of articulation to form six vowel sets, those adjacent to palatalized consonants and those adjacent to plain consonants. Thus the model was simplified, and further analyses were done exclusively on the effects of palatalization of the preceding or following segment (or both) on the formant values of the vowels.

TABLE 2: ANOVA SUMMARY

		PALATALIZATION		PLACE OF ARTICULATION		INTERACTION: PALATALIZATION AND ARTICULATION	
		f	p	f	p	f	p
/i/	F1	0.473	0.703	0.300	0.825	1.767	0.120
	F2	3.770	0.018*	1.848	0.154	0.644	0.717
	F3	0.896	0.451	0.017	0.997	0.295	0.952
/ɪ/	F1	1.146	0.343	1.834	0.157	2.178	0.066
	F2	5.066	0.005*	1.856	0.153	0.590	0.736
	F3	0.531	0.664	2.804	0.052	1.072	0.396
/u/	F1	1.021	0.392	1.163	0.333	0.436	0.649
	F2	4.379	0.008*	1.847	0.151	0.129	0.879
	F3	5.912	0.002*	4.955	0.004*	0.131	0.878
/e/	F1	2.470	0.073	3.384	0.025*	3.101	0.035*
	F2	3.691	0.018*	0.436	0.728	2.178	0.102
	F3	4.877	0.005*	0.701	0.556	4.186	0.010*
/o/	F1	0.161	0.922	2.390	0.082	3.294	0.019*
	F2	5.806	0.002*	1.842	0.154	3.001	0.028*
	F3	4.451	0.008*	2.197	0.102	3.723	0.011*
/a/	F1	2.476	0.075	0.198	0.897	0.815	0.546
	F2	13.742	0.000*	0.091	0.964	1.468	0.221
	F3	0.208	0.890	0.755	0.526	0.460	0.804

4.2 Results of t-tests

A series of independent t-tests were performed on each of the six vowel qualities, using the Grouping Variable of the two categories, palatalized and plain. These t-tests show statistically significant differences in F2 means for all vowels adjacent to palatalized consonants versus plain consonants. In Table 3, upper case C with superscript j (C^j) indicates a palatalized consonant adjacent to the vowel and plain upper case C indicates a plain consonant adjacent to the vowel. Values of 'p' which are starred are statistically significant.

TABLE 3: MEAN FORMANT VALUES (HZ) OF ISTHMUS MIXE VOWELS
ADJACENT TO PALATALIZED AND PLAIN CONSONANTS
Numbers in parentheses represent standard deviations.

		C ^j _	C_	t	p
/i/	F1	317.4 (30.9)	322.1 (32.6)	0.566	p = 0.574
	F2	2232.8 (121.9)	2132.9 (97.8)	-3.502	p = 0.001*
	F3	3582.9 (106.9)	3474.9 (231.4)	-2.320	p = 0.024*
/i/	F1	335.9 (44.5)	376.3 (44.2)	3.524	p = 0.001*
	F2	1763.8 (150.3)	1517.0 (179.6)	-5.770	p = 0.000*
	F3	2521.8 (108.3)	2482.0 (136.8)	-1.250	p = 0.216
/u/	F1	344.6 (30.7)	339.2 (34.1)	-0.648	p = 0.519
	F2	1079.5 (84.0)	1006.6 (48.2)	-4.126	p = 0.000*
	F3	2423.1 (151.2)	2424.6 (131.7)	-0.041	p = 0.967
/e/	F1	406.9 (32.8)	447.0 (51.3)	3.611	p = 0.001*
	F2	1920.8 (77.1)	1870.0 (70.2)	-2.672	p = 0.010*
	F3	2584.1 (90.6)	2562.5 (78.3)	-0.984	p = 0.329
/o/	F1	449.2 (56.2)	433.8 (56.4)	-1.064	p = 0.292
	F2	1184.7 (108.0)	1083.3 (82.5)	-4.086	p = 0.000*
	F3	2401.7 (105.0)	2476.3 (124.4)	2.507	p = 0.015*
/a/	F1	640.6 (65.7)	634.1 (45.8)	-0.448	p = 0.656
	F2	1617.1 (71.1)	1421.1 (108.0)	-8.304	p = 0.000*
	F3	2370.0 (95.1)	2341.8 (154.5)	-0.853	p = 0.397

Given that F2 values are related to vowel frontness and backness, and that the F2s of the steady states of these Isthmus Mixe vowels adjacent to palatalized consonants are significantly greater than the F2s of vowels not adjacent to palatalized consonants, the analyses of the acoustic data suggest that in Isthmus Mixe, the process of palatalization is best viewed as a case of fronting of the vowels. For example the F2 of /i/ adjacent to a palatalized consonant (C^ji) is 2232.8 Hz which indicates that the vowel is more fronted than /i/ adjacent to plain consonants (Ci) at 2132.9 Hz.

F1 means show a significant difference for two vowels, /i/ and /e/. Since F1 values are inversely related to the height of the vowel (the smaller the number, the higher the vowel), /C^ji/, with an F1 of 335.9 Hz is higher than /Ci/, 376.3 Hz, a difference of 40.4 Hz. The F1 of /C^je/ is 406.9 Hz and of /Ce/ 447.0 Hz, a difference of 40.1 Hz. In both cases the actual difference is just over 40 Hz, which slightly affects the height of these two vowels, while no significant difference is found in the other vowels.

4.3 Formant Plot of Isthmus Mixe Vowels

A vowel formant plot of the data in Table 3, shown in Figure 4, is constructed with the F1 means on the vertical axis, smaller numbers at the top descending to the larger. The F2 means are plotted on the horizontal axis with the smaller numbers beginning at the right so that it resembles the articulatory vowel space.

Although palatalization causes fronting of the vowels, the means of the palatalized vowels do not impinge upon other phonemic vowel spaces. This is not to say that a few individual tokens do not overlap the minimum and maximum ranges of other vowel

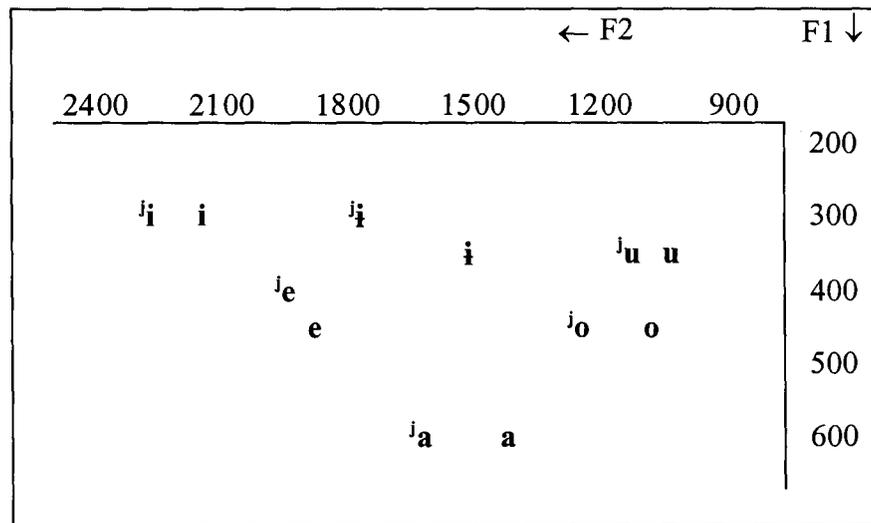


Figure 4: Isthmus Mixe Vowels Adjacent to Palatalized and Plain Consonants

The symbol ^j preceding the vowel indicates vowel means adjacent to palatalized consonants. A plain vowel indicates vowel means adjacent to plain consonants.

spaces. One might suppose that /i/ when adjacent to a palatalized consonant, would encroach upon the vowel space of /i/ adjacent to plain consonants, its closest neighbor. However, even though palatalization fronts /C^ji/, the F2 mean is still 369.1 Hz less than the F2 mean of the plain vowel /Ci/. A few individual tokens of /C^ji/ do shift into the range of the formants of /Ci/. Overall, this does not affect the F2 mean of /C^ji/. On the average then, no vowel quality shifts into the formants of another vowel quality, even when fronted by the effects of palatalization.

TABLE 4: MEANS AND RANGES OF ISTHMUS MIXE VOWELS IN HZ
Numbers in parentheses represent minimum-maximum.

		C ^j <u> </u>	C <u> </u>
/i/	F1	317.4 (252.1-378.2)	322.1 (240.7-378.2)
	F2	2232.8 (2005.6-2635.9)	2132.9 (1925.4-2280.6)
/ị/	F1	335.9 (275.1-424.0)	376.3 (263.6-469.9)
	F2	1763.8 (1478.4-2085.8)	1517.0 (1180.4- 1799.3)
/u/	F1	344.6 (286.5-389.7)	339.2 (263.6-389.7)
	F2	1079.5 (905.4-1237.7)	1006.7 (905.4-1088.7)
/e/	F1	406.9 (355.3-481.3)	447.0 (332.4-573.0)
	F2	1920.8 (1822.2-2085.8)	1870.0 (1696.2-1982.7)
/o/	F1	449.2 (378.2-595.9)	433.8 (355.3-584.5)
	F2	1184.7 (1042.9-1455.5)	1083.3 (928.3-1249.2)
/a/	F1	640.6 (481.3-744.9)	634.1 (550.1-722.0)
	F2	1617.1 (1489.9-1764.9)	1421.1 (1260.7-1615.9)

In Table 4 the means (from Table 3) are listed which are used in Figure 4 to display the phonemic vowel spaces, and below the means the minimum and maximum of each vowel provide evidence that some overlapping of individual tokens occurs. Overall, the averages show consistent fronting of vowels adjacent to palatalized consonants in comparison to vowels adjacent to plain consonants.

5.0 Conclusion

This research provides a preliminary quantitative overview of the frequency counts of the vowel qualities and the widespread effects of palatalization in this Isthmus Mixe text. It has been demonstrated that the steady-state formants of vowels adjacent to palatalized consonants are different from those adjacent to plain consonants, and the change is significant for F2 values for all vowel qualities. These data support the notion that the primary effect of palatalization in Isthmus Mixe should be regarded as fronting of the adjacent vowel.

Finally, as shown in Figure 4, regardless of the occurrence of palatalization, the phonemic vowel spaces are maintained. Each vowel quality demonstrates two allophones, one which occurs adjacent to plain consonants and a second, more fronted version, which occurs adjacent to palatalized consonants.

The research presented in this paper is a preliminary study of the phonetics of Isthmus Mixe, an endangered and under-described language. It gives a substantial phonetic basis for a more complete phonological description of Isthmus Mixe than has been done to date, as well as contributing to a better understanding of the phonological descriptions of other Mixe languages.

Notes

* I would like to thank my Mixe friends for their patience in helping me with the language, and who extended such excellent hospitality to me during my visits to the Isthmus. Much of the material of this article was presented in partial forms to audiences at the The University of Texas at Arlington Student Conference in Linguistics (February 2001) and the Workshop on American Indigenous Languages, in Santa Barbara (July 2001). I am grateful to the members of these audiences for their encouragement and suggestions. The trip to Santa Barbara was financed by the Linguistics program of The University of Texas at Arlington. A very special thanks to Norris McKinney who helped me with my earliest attempts at acoustic analysis and to David Silva for valuable comments which improved the analyses and the writing of the paper.

¹ The names reflect geographical or political relationships. *Isthmus* is the geographical area of the Isthmus of Tehuantepec. The label *Eastern* is in relationship to the other Mixe languages which are spoken in areas farther to the west. San Juan Guichicovi is the major city in this Mixe region.

² According to Ladefoged (1993:230), "the formal definition of a secondary articulation is that it is an articulation with a lesser degree of closure occurring at the same time as another (primary) articulation. ... Palatalization is the addition of a high front tongue position, like that in [i], to another articulation."

³ All the vowel quality phonemes may be mutated by the autosegmental features of length and creaky voice, and all the consonant phonemes may be mutated by the autosegmental feature of secondary palatalization.

⁴ One exception is that consonant /j/ is not followed by vowel /i/.

⁵ Ladefoged (2001:169) says that English-speaking network newscasters speak at about three syllables per second.

⁶ The consonant phoneme /j/ is considered to be a palatalized consonant regardless of the occurrence of the autosegmental morpheme for this particular study of the effects of palatalization on 300 vowels.

⁷ F1 and F2 values are adequate to distinguish the six vowel qualities. However, F3 values are shown in Table 2 and Table 3 to give a more complete description of the formants and as confirmation that F3 values have little or no function in distinguishing Isthmus Mixe vowels.

References

- Baart, Joan. 1999. A field manual of acoustic phonetics. Prepublication version included in *LinguaLinks Library*, Version 3.5 and higher (CD-ROM), SIL International.
- Bhat, D.N.S. 1974. A general study of palatalization. *Working Papers on Language Universals* (Stanford University) 14:17-58.
- Campbell, Lyle and Terrence Kaufman. 1976. A linguistic look at the Olmecs. *American Antiquity* 41.1:80-89.
- Clark, Ben T. 1983. *Russian*, third edition. New York: Harper & Row, Publishers.
- Comrie, Bernard. 1981. *The languages of the Soviet Union*. Cambridge: Cambridge University Press.
- Comrie, Bernard. 1987. *Russian*. *The world's major languages*, ed. by Bernard Comrie, 329-347. London/Sydney: Croom Helm.
- Fant, Gunnar. 1970. *Acoustic theory of speech production*. With calculations based on X-ray studies of Russian articulations. Second Edition. The Hague: Mouton.
- Fant, Gunnar. 1973. *Speech sounds and features*. (Current Studies in Linguistics 4.) Cambridge, MA: The MIT Press.
- Grimes, Barbara F (ed). 2000. *Ethnologue: Languages of the world*, 14th ed. Vol. 1. Dallas: SIL International.
- Hilts, Craig. 1999. Proto-Mixe-Zoquean: A case in linguistic and cultural reconstruction. *Historical linguistics* 235-248. (Working Papers in Linguistics No. 52.) Columbus, OH: The Ohio State University.
- Hombert, Jean-Marie and Ian Maddieson. 1998. A linguistic approach to automatic language recognition. *UCLA Working Papers in Phonetics* No. 96, 119-125.
- Ladefoged, Peter. 1993. *A course in phonetics*, Third edition. San Diego, CA: Harcourt Brace Jovanovich, Publishers.
- Ladefoged, Peter. 2001. *Vowels and consonants*. Malden, MA/Oxford: Blackwell Publishers, Inc.
- Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell Publishers.

- Maddieson, Ian. 1984. *Patterns of sounds*. (Cambridge Studies in Speech Science and Communication.) London/New York: Cambridge University Press.
- Ní Chiosáin, Máire. 1994. Irish palatalisation and the representation of place features. *Phonology* 11:89-106.
- Salminen, Tapani. 1999. Tundra Nenets. Internet: tasalmin@cc.helsinki.fi.
- Stevens, Kenneth N. 1998. *Acoustic phonetics*. (Current Studies in Linguistics 30.) Cambridge, MA: The MIT Press.
- Stross, Brian. 1985. Maya head variant numerals: the Olmec case. *Anthropological Linguistics*, 27.1:1-45.
- Van Summers, W. 1987. Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *Journal of the Acoustical Society of America*. 82(3), Sept. 1987:847-863.
- Wichmann, Soren. 1995. *The relationship among the Mixe-Zoquean languages of Mexico*. (Studies in Indigenous Languages of the Americas.) Salt Lake City: University of Utah Press.

Diccionario yaqui-español: a lexicographic project

Zarina Estrada Fernández
Universidad de Sonora

Yaqui is a Taracahitan language with almost twenty thousand speakers most of them settled in the eight “pueblos yaquis” along the Yaqui River: Belén o Pitahaya, Huirivis, Loma de Bâcum o Cócorit, Loma de Guamúchil, Pótam, Rahum, Tórim, and Vícam. There are, however, probably, another 5,000 Yaqui speakers in Pascua, Arizona, in the U.S.A. The Yaqui tribe from Sonora, Mexico, is one of the very few groups in this country who have developed their own bilingual program.¹

During the last five years, I have been involved in the *Diccionario yaqui-español* project, whose first goal was making a dictionary. In this presentation, I briefly discuss four issues: i. the function of the *Diccionario Yaqui-Español*, ii. the orthography adopted in this piece of work, iii. some of the analytical decisions required for the construction of the dictionary, and iv. the discussion of some interesting tasks which remain to be addressed in the future.

1. Introduction. The Yaqui-Spanish Dictionary project began in August 1995, when CONACYT (Consejo Nacional de Ciencia y Tecnología) approved a grant to the University of Sonora.² Three years before, in 1992, The University of Arizona, with Eloise Jelinek as the main scholar, had promoted the idea of building a binational dictionary for the Yaqui people. The motivation for the binational character of such a dictionary was the fact that the Yaqui from Arizona were losing their language. According to Fernando Escalante’s point of view, --a former PhD in Linguistics and also a Yaqui speaker from Arizona--, among the Yaqui from Arizona there was a lot of lexical loss, particularly within the semantic fields of plants and animals. In other words, the Yaquis from Arizona were losing all the lexicon which is not practically used in common day life, including words for naming animals, plants, and also ceremonies as well as some other behaviors of traditional way of life.

Unfortunately, the project proposed by the group of the University of Arizona never succeeded in obtaining financial support. Meanwhile, since Jelinek accepted for Mexican scholars to look for financial support, the project in Mexico was accepted by CONACYT. The Yaqui dictionary is now ready to go to press. In the following sections I will provide some information concerning the characterization of the Dictionary that we have developed.

2. Data. The data for the Yaqui-Spanish Dictionary was, primarily, provided by Crescencio Buitimea Valenzuela, a BA student of Linguistics, and by Anabela Carlón Flores, a BA student of Law, both at the University of Sonora. Data was also collected in Vícam and Huirivis from different native speakers. The Dictionary includes, at the end, a collection of texts, some of them written initially by Hilario Molina, a native speaker of Yaqui from Vícam, and a very active participant in the Bilingual Education Project of the Yaqui Tribe. *Ume’e Surem*, an additional narrative text of about 140 clauses, was elicited between 1990 and 1992 by Manuel Carlos Silva Encinas from Pablo Álvarez, a native speaker of Yaqui

born in Hermosillo, Sonora, in 1921.³

3. Function of the Dictionary. Very few bilingual dictionaries have as asymmetrical a structure as the Yaqui-español dictionary. (For example the *Hopi Dictionary* (1998)). The main function of the asymmetrical structure of the Yaqui-Spanish dictionary is the need to be useful for native speakers in revitalizing their own language, but also to be helpful to them in improving their knowledge of Spanish. In this way, the dictionary project recognizes the right of indigenous people to be educated in their own language, as well as to have the opportunity to develop a good understanding of Spanish, the national language of Mexico. Due to this reason, the main body and richest part of the dictionary is the Yaqui-Spanish section, and the simple one is the Spanish-Yaqui vocabulary, cf. examples in (1).

(1) *The Yaqui-Español data:*

- ta** *nexo*. Pero. Ta si uka aukamta bettesi a machiako junak ala ume goj naikitaka nau yaja. ‘Pero si lo tratado se ve complicado, entonces los ocho pueblos se reúnen’. Cf. bweta.
- taa'a** [taá'a]. *n.* Día, sol. U taa'a unna utte'ak. ‘El sol está muy fuerte’.
- taabu** [taábu]. *n.* Conejo. *Sylvilagus spp.* U taabu bwia gojo'oku kibakek. ‘El conejo entró al agujero’.
- taaka** [taáka]. *n.* Fruto, vaina. U ju'upa taaka si ousi kia. ‘El fruto del mezquite es muy sabroso’.
- tajo'ote** *v.* Vestirse. Maria chumti tajo'ote. ‘María se viste de prisa’. *Sin.* a'ana.

The Español-Yaqui vocabulary:

- padre, papá, señor respetable, persona mayor, Dios** *n.* achai.
- padre, papá de mujer, señor** *n.* japchi.
- padrino** *n.* bato'o achai.
- pagar** *v.* beje'etua.
- pájaro carpintero** *n.* cho'oroí, cholloi.
- paladar** *n.* kaapa.
- palo de Brasil** *n.* juchajko.
- palo dulce** *n.* ba'ego, baegotaate.
- palo fierro** *n.* ejea.
- palo que se raspa en la danza del venado** *n.* jirukia.

The Yaqui-Español Dictionary is based on the analysis of the user's competence, but it also takes advantage of the communicative environment from different kind of texts: pedagogical texts, ethnographic or folklore texts, and colloquial texts. The first kind of texts are taken originally from two educative textbooks for Preliminary School. Those texts, among others, were elaborated by a group of professors of the Yaqui Tribe. These

texts were initially included --without gloss-- in the two volumes set of the first grade text book entitled *Jiak Noki - Lengua Yaqui, Sonora*. In order to be published in the Dictionary, this collection of texts was revised and adapted by Estrada and Buitimea. The gloss of the texts, was carried out during 1997 as the main task of the *Seminario de análisis de textos yaquis*, which took place at the University of Sonora. As I mentioned previously, the ethnographic text *Ume'e Surem* was documented by Manuel Carlos Silva Encinas, and glossed with the kindly help of María de Jesús Muro. At the end of the Dictionary, some instructional texts are provided. These colloquial texts were obtained from Anabela Carlón as a part of a project whose main goal was to document the kind of foods the Yaqui people prepare and consume during their festivities.⁴ Much of this material is, up to certain level, included among the different examples of the dictionary.

4. Orthography. The orthography used all along in the Yaqui-Spanish Dictionary has been developed by the Yaqui Tribe Project on Bilingual Education. The complete standard alphabet has five vowels and seventeen consonants; all illustrated in (2):

- (2) a, e, i, o, u
' , b, bw, ch, d, g, j, k, l, m, n, p, r, s, t, w, and y

All the characters in (2) are given in the alphabetic order used in the dictionary. The phonetic chart of the consonant symbols of the Dictionary is provided in (3):

- (3)
- | | | | | |
|----|---|---|---|---|
| p | t | č | k | ' |
| b | d | | g | |
| bw | s | | | h |
| m | n | | | |
| | l | | | |
| | r | | | |
| w | | y | | |

Long vowels and geminate consonants are represented with two letters as in (4):

- (4) aa, ee, ii, oo, uu, and pp, tt, mm.

5. Structure of the Dictionary. The data in the dictionary are presented in the following way. Each entry begins in bold letters (see (5)). After that, within brackets, word stress is described, but only for those words not accented on the first vowel of the word. Next, the word category or word class is given written in italics. After that comes the equivalent item in Spanish. For names of plants or animals, the translation into Spanish is followed by the scientific name. After that, one, two or more sentences are provided to illustrate the word order, and also the language use as well as some of the morphophonological properties of the lexical item. Finally, at the end of a lexical entry, information about synonymous forms is sometimes provided. In (5) several lexical entries are illustrated:

(5)

aakame *n.* Víbora de cascabel. *Crotalus spp.* U aakame ousi joyok. ‘La víbora de cascabel tiene mucho veneno’; ¡Aakamta beba! ‘Pégale a la cascabel!’

aaki [aáki]. *n.* Pitahaya. *Stenocereus thuberi.* (Engelman) Buxbaum. Aaki takaam si ousi kaka. ‘El fruto de la pitahaya es muy dulce’; U aaki bakot jujariata betchi’ibo tu’i. ‘La pitahaya es buena para la picadura de víbora’; Aakim te juyau nu’uka am nenkibae. ‘Después de traer pitahayas del monte las venderemos’; Aakamem bitla, akim pupua intok bachita e’echa. ‘Ven víboras de cascabel, recogen pitahayas y siembran maíz’.

ama *adv.* Ahí. Ketun ama weyek u yoeme. ‘Todavía está ahí parado el hombre’. *Sin.* junum, junama.

ama butte *adv.* Demasiado. Empo nokpo ama butte. ‘Tú hablas demasiado’; Empo ama butte wasapo tekipanoa. ‘Trabajas demasiado en la parcela’.

amako [amáko]. *adv.* A veces. Batwepo ne u’uba amako. ‘A veces me baño en el río’.

aman, amani *adv.* Allá, por ahí. Aman wasuktiapo 27po. ‘Allá en el año 27’; Jaksa katek u yoeme? ... Amani. ‘¿Dónde está el hombre? ... Allá’; Batweta bo’okapo amani wakasim wam sajak. ‘Por dónde está el río, por ahí pasaron las vacas’. *Cf.* junamani.

As mentioned above, a collection of texts is given following the Yaqui-Spanish Dictionary, and the Spanish-Yaqui vocabulary. These texts are presented in Yaqui and Spanish with their respective lexical and morphosyntactic analyses or glosses. Lists of plants, animals and names of relatives are provided as well as some charts with pronominal and deictic forms, and also a list of suffixes. In (6), the text “Romina” is provided. All the textual material was included as an important part of the dictionary in order to be useful not only for Yaqui people, but also for linguists interested in the grammar of this language.

(6) Text “Romina” from Martín Zúñiga

¿Jakunsa e siika Romina, jakunsa e? ¿Jakunsa e siika Romina, ke kaa ne tejwak? Empo kia yeu sisime intok kaita e jikkaibae. ¡Romina banse’e! -¡Banse’e Romina em maala eu chaae! ¿Jaksa e kibakla Romina? Tukaariat naabujtia emo tatta’aru.

“¿A dónde vas Romina, a dónde? ¿A dónde vas Romina que no me dijiste? Tú nada más te vas y nada quieres entender. -¡Romina apúrate! -¡Apúrate Romina, te está gritando tu mamá! ¿En dónde estás metida Romina? Incluso hasta en la noche tú te pierdes”.

Glosa

1.	¿Jakunsa	e	siika	Romina,	jakunsa	e?
	jakúnsa	émpo	siika	romína	jakúnsa	émpo
	Interr	2Sg.Nom	ir.Pres	R.	Interr	2Sg.Nom

“¿A dónde vas Romina, a dónde?”

2. ¿Jakunsa e siika Romina ke kaa nee tejwak?
 jakúnsa émpo siíka romína ke kaá neé téwa-k
 Interr 2Sg.Nom ir.Pres R. que Neg 1Sg.Acu decir-Perf
 “¿A dónde vas Romina que no me dijiste?”
3. Empo kia yeu sisime intok kaita e jikkaibae.
 émpo kía yéu si-síme ínto-kéchia káita émpo jíkaja-bae
 2Sg.Nom solo Dir Rdp-ir.Sg.Pres y-también nada 2Sg.Nom oir.Des
 “Tú nada más te vas y nada quieres entender”.
4. -¡Romina banse’e!
 romína bamsé’e
 R. apurarse.Imp
 “-¡Romina apúrate!”
5. Banse’e Romina em maala eu chaae
 bamsé’e romína ém maála éu chaáe
 apurarse.Imp R. 2Sg.Gen madre 2Sg.Dat gritar.Pres
 “-¡Vente Romina, te está gritando tu mamá!”
6. ¿Jaksa e kibakla Romina?
 jáksa émpo kibáke-la romína
 Interr 2Sg.Nom meter-Adjv R.
 “¿En dónde estás metida Romina?”
7. Tukaariat naabujtia emo tatta’aru.
 tukáa-ria-t náabujtia emó tattá’aru
 noche-Apl-sobre incluso Rflx Rdp.perder.Pres
 “Incluso hasta en la noche tú te pierdes”.

6. Analytical decisions. Several problems had to be dealt with during the development of the dictionary. They relate to the following issues:

- i. What form should be selected as a lexical entry for verbs and nouns in the dictionary?
- ii. Should positionals and other stative notions be characterized as adjectives or stative verbs?
- iii. Should lexical entries be represented not only with words but also phrases?
- iv. What should be done with small discourse particles that lack lexical meaning?

- v. How should the discourse context be represented when a lexical entry coming from a text is given?

Some of these questions may seem trivial. Nevertheless, they were problems that kept recurring during the process of producing the dictionary.

Lexical entries. For most nouns, adverbs, pronouns and particles, there was no problem choosing a particular form of a word as a lexical entry of the dictionary. For nouns, the lexical entries were easy to choose since most of them are unmarked, except those which denote mass referents. Those nouns usually end with an *-m*. Examples are illustrated in (7):

- (7) mass nouns ending with *-m*
- | | |
|---------------|------------------|
| ba'am | 'water' |
| bajekam | 'wave' |
| bannaim | 'atole' |
| chiinim | 'cotton' |
| chookaraktiam | 'wrinkle' |
| kuchuwam | 'acne' |
| kujteerim | 'courage' |
| jaakam | 'phlegm' |
| ji'ikiam | 'needle' |
| mamyam | 'kind of greens' |

There are however, several unmarked mass nouns:

- (8) unmarked mass nouns
- | | |
|-----------|----------------|
| oona | 'salt' |
| bachí | 'corn' |
| bwia | 'soil' |
| bwichopia | 'soot', 'smut' |
| jeeka | 'wind' |
| muuni | 'beans' |

Several loans from Spanish are, unexpectedly, categorized in Yaqui as mass nouns (9a), while others that one might expect to be characterized as mass nouns, are categorized as non-plural, non-mass nouns. Examples are provided in (9b):

- (9) a. Borrowings from Spanish categorized as mass or plural nouns.

arado	<i>n.</i> araom	'plow', 'plough'
azadón	<i>n.</i> asaroonim	'hoe'
cuchillo	<i>n.</i> kuchi'im	'knife'

cuete	<i>n.</i> kuetem	‘firework’
hoz	<i>n.</i> oosam	‘sickle’
tijeras	<i>n.</i> chaptiam	‘scissors’

- b. Borrowings from Spanish not categorized as mass nouns:
 azúcar *n.* asuka ‘sugar’

For the verbs in general, the unmarked or simplest form, that is, the present tense form, was selected as the lexical entry for the dictionary. Consider the uninflected dictionary entries in (10):

(10)	a’abose	‘to offer food’
	a’ana	‘to dress’
	aamu	‘to hunt’
	ja’abwa	‘to stand up’ (trans.)
	ju’unea	‘to know’
	ke’e	‘to bite’
	sisiba	‘to scrap’
	sioka	‘to suffer’

However, in languages like Yaqui, many verbs are obligatorily inflected. For this reason, we needed to afford different kinds of decisions: one particular case was, for example, the aspectual inherent meaning of some verbs. For these verbs, their inherent aspectual meaning prohibits habitual interpretation, therefore, to choose as the lexical entry the unmarked form. For those verbs, the perfective form, with the suffix *-k* was chosen as the lexical entry, examples are provided in (11).

(11)	chittuak	‘to make drivel’
	essok	‘to hide’
	ja’abwek	‘to stand up’ (intr.)
	ju’unaktek	‘to create’
	siliktek	‘to twist’

Other analytical decisions were prompted by the irregular morphological derivational patterns of Yaqui verbs. These verbs are illustrated in (12). In (12a), I illustrate some verbs that have both the intransitive and transitive forms; in (12b), however, some verbs have no related transitive or intransitive form. Finally, in (12c) some examples of the class of verbs ending with the suffix *-te* are shown; all of which only have a transitive form.

(12)	intransitive	transitive	gloss
a.	biakte	biakta	‘to roll, to turn’
	butte	butta	‘to untie’

	chakte	chakta	'to drop'
	chamte	chamta	'to break'
	chepte	chepta	'to jump'
	etajte	etajta	'to cut'
	go'ote	go'ota	'to pour out'
	go'okte	go'okta	'to sink'
b.	-----	bilujta	'to smash with its hands'
	biute	-----	'to whistle'
c.	-----	a'ate	'to carry'
	-----	ba'ate	'to make something watery'
	-----	bwabwite	'to sharpen'

Moreover, there are other cases where the transitive and intransitive forms are semantically related, but their meaning are not predictable, for example:

- (13) be'okte 'to lighten, to pull out the tongue reiteratively'
 be'okta 'to lick, to lap'

Positionals verbs and stative adjectives. Commonly, verbs with telic interpretation in Yaqui end with the perfective suffix *-k/-ka*. Some examples are illustrated in (14):

- (14) aayuk 'to exist in a place'
 manek 'to be placed'
 bo'oka (sg.) to'oka (pl.) 'to be laid'
 katek (sg.) jooka (pl.) 'to be sit'
 weyek (sg.) ja'abwek (pl.) 'to be raised'

On the other hand, a relatively small group of forms referring to states, are also obtained with the suffix *-k/-ka*, sometimes realized as *-ko*, when the suffix echoes the vowel of the root. These words were categorized as adjectives since: they denote a property of an entity, not a position, occur within a noun phrase modifying a noun, and do not accept tense/aspect morphology.

- (15) ejoak 'stained'
 baseka 'light'
 beojko 'bright'
 cho'oko 'salty'
 jeoko 'viscose'
 wilojko 'thin'

This kind of adjectives, differ from other adjectives in their morphological characterization. In (16), I provide examples of adjectives ending with *-i*, *-la*, *-lai*, as well as those derived with the agentive relativizer (or nominalizer) *-me*:

- (16) Adjectives in *-i*:
- a. aw-i ‘fat’
 - bwas-i ‘cooked’
 - chakku-i ‘twisted’
 - etapo-i ‘open’
 - ilippan-i ‘short’
 - jiar-i ‘abandoned’
- b. Adjectives in *-la*:
- ba’ayejte-la ‘watered’
 - boyo-la ‘blistery’
 - bwichop-la ‘smoked’
 - chopo-la ‘haunchy’
 - jepe-la ‘aligned’
 - jupp-la ‘married’
- c. Adjectives in *-lai*:
- beche-lai ‘open’
 - beta-lai ‘thin’
 - chitabeta-lai ‘flat’
 - bwawisibu-lai ‘pointed’
 - lutu-lai ‘straight’
- d. Adjectives in *-me*:
- pajko-me ‘the one which participates very often in the pascola’

There are, however, plenty of non-derived adjectives. In (17) I provide some examples of adjectives ending in vowels *-a*, *-e* or *-u*:

- (17)
- alea ‘healthy’
 - allea ‘happy’
 - beje’e ‘expensive’
 - bwe’u ‘big’
 - chakku ‘twisted’
 - chiibu ‘acid’
 - ju’ena ‘bad’
 - juuba ‘stinky’

Expressions and phrases. In the elaboration of a dictionary it must be considered

that the relation between lexical form and meaning is not a one-to-one relation. Many examples from Yaqui, illustrate this fact. Words in Yaqui which need to be translated as descriptions or phrases in Spanish are illustrated in (18a), meanwhile the inverse, notions that in Yaqui constitute descriptive nouns, are provided in (18b):

- (18) a. a'abose to offer food in a particular date'
 baima 'to wash one's hands'
 bilujta 'to crush with one's hands', 'to smash with one's hands'
 chona 'to hit with one's hands', 'to hit with fist'
- b. ania au yoa 'to quake' *Lit.* the world quakes.
 ba'a nu'u 'water bottle', 'canteen'
 chabula suma 'to wreath', 'to be tight for'

Discourse particles. In general, the treatment of discourse particles in dictionaries of languages like Spanish for example, is often poor, non-systematic, and inaccurate. For languages as Yaqui, the documentation of discourse particles is still in the initial phase even though discourse particles play an important role in the language. Thus, we have documented some of the Yaqui discourse particles that occur in our corpora. For these particles it is not possible to provide a strict translation into Spanish. There is, however, for some cases, a descriptive definition and examples from Spanish which may be widely taken as equivalent particles. The discourse particles included in the dictionary are illustrated in (18):

- (19)
- a'e part.** Partícula discursiva o de la conversación cuyo uso es exclusivo del género femenino. ¡ A'e jana juni jiune u uusi! '¡ Mmh, qué tonterías dice el muchacho!'
- abwe part.** Partícula que se utiliza al inicio de una conversación. Equivale en español a: 'a poco', 'bueno', 'este', 'pues'. Abwe jiba a'abo noitek. 'Pues siempre vino'. *Cf.* bwe.
- ori part.** Partícula transicional común en el discurso. Equivale en español a: 'o sea', 'pues'. Intuchi Santa Roosa, into Guaalupe... junume'e aman papajkoriawa, ori, Loma Guamuchilpo. 'También Santa Rosa y Guadalupe. esas fiestas allá se celebran, o sea, en Loma de Guamúchil'; Jakun empo siika?... Ori... si mekka. '¿ A dónde vas?... Pues... muy lejos'.

Discourse context. Another kind of analytical decision relates to the ethnocultural meaning of some words. For this type of words, it is almost impossible to provide a translation without describing the discourse context. Most of these words have to do with the social organization of the Yaqui, words referring to dignity or ritual cargos of men during the festivities or to ceremonial objects. Examples are provided in (20):

(20)

alawasi *n.* Alguacil. Alawasi Joanta a jo'awau un'usek. 'El alguacil fue por Juan a su casa'; Alawasi kobanaota susua. 'El alguacil cuida al Gobernador'.

alpees [alpeés]. *n.* Alferez. Oficial que lleva la bandera en cualquier procesión religiosa. U alpees kaa lauti yepsak. 'El alférez llegó tarde'; Alpeesta te jo'owau un'une. 'Traeremos al alférez de su casa'.

jirukia [jirúkia]. *n.* Palo que se raspa en la danza del venado. Jirukia baapo kom wechek. 'El palo cayó al agua'; Ume'e yoemem maso bwibwikame jirukiata koptak. 'A los cantantes de venado se les olvidó el palo'.

maejto yo'owe [maejtó yo'owe]. *n.* Primer maestro, sacerdote mayor, encargado de la iglesia. Kat ne maejto yo'oweta teak. 'No encontré al primer maestro'.

teneboim *n.* Tenabaris. Sonajas elaboradas con capullos de mariposas que los danzantes de pascola portan en las piernas. In jaboi teneboim jooa. 'Mi abuelo está haciendo tenabaris'.

7. Remaining tasks. The Yaqui-Spanish Dictionary project has been developed assuming in general, that the meaning of words must be found mainly in their usage. This has motivated that along the dictionary, many naturalistic data were taken into account to illustrate different words. There are some interesting issues that need to be addressed in order to deepen our knowledge of the grammar of this language. Some of those issues are: the exploration of complex clauses, the documentation of derived morphology, the study of descriptive phrases. If we assume, that in general, the dictionaries have been an important tool for native speakers and linguists, the exploration of language focused in the use of some particular part of grammar will shed light on many aspects of language which do not necessarily emerge in our everyday use of language. It will be necessary, for example, to explore language argumentation techniques, to understand some aspects of grammar. A language learning or language-use dictionary like the Yaqui-Spanish dictionary we have developed will be of great help in exploring those different sections of grammar. The compilation of the Yaqui-Spanish dictionary has brought new insights in lexicographic research on Indigenous languages of Mexico and to endangered languages in general.

Notes

¹ I want to acknowledge Karen Dakin, Luis Barragán and Søren Wichmann for their helpful comments on earlier versions of this paper. Any error is of my own responsibility.

² The project was originally titled "Diccionario Yaqui-Inglés-Español" and was approved by CONACYT with the number 489100-5-4101-H.

³ *Ume'e Surem* or *Los Sures* was published in Silva et al. (1998). The narrative will be reproduced in the Yaqui-Español Dictionary with the permission of the authors.

⁴ This material was collected by Guadalupe Yanez and Yudith González during the period they were participating in the activities of the “Diccionario Yaqui-Español” project.

References

- Castillo, María Elena, Zarina Estrada, Adriana Gurrola. s/f. Diccionario yaqui-español-inglés: obra de preservación lingüística para dos naciones. To appear in *Lingüística Mexicana*.
- Crumrine, Lynne S. 1961. The phonology of Arizona Yaqui. *Anthropological Papers* 5. Tucson: University of Arizona Press.
- Demers, Richard, Fernando Escalante and Eloise Jelinek. 1999. Prominence in Yaqui Words. *International Journal of American Linguistics*. Vol. 65. No. 1: 40-55.
- Escalante, Fernando. 1990. Voice and Argument Structure in Yaqui. PhD. Dissertation. Tucson: The University of Arizona.
- Escalante, Fernando and Eloise Jelinek. 1989. Double Accusative in Yaqui. *Proceedings of the Fourth Meeting of the Pacific Coast Conference*. University of Oregon. 120-132.
- Estrada F., Zarina. 2000. Copula constructions in Uto-Aztec languages. *Uto-Aztec: Structural, Temporal and Geographic Perspectives*. Papers in Memory of Wick R. Miller by the friends of Uto-Aztec. Eugene H. Casad and Thomas L. Willett (eds.). Universidad de Sonora. 139-154.
- _____. 1998. Aspectos lingüísticos. *Jiák nokpo etéjoim (pláticas en lengua yaqui)*. Silva Encinas, et al. Hermosillo: Universidad de Sonora. 123-140.
- Fraenkel, Gerd. 1959. Yaqui phonemics. *Anthropological Linguistics*. Vol. 1. Num. 5: 7-17.
- Jelinek, Eloise and Fernando Escalante. 1986. Yaqui minus control. *Proceedings of the Second Annual Meeting of the Pacific Coast Conference*. University of Oregon. 120-132.
- _____. 1988. Verbless possessive sentences in Yaqui. In *Honor of Mary Hass: From the Hass Festival Conference on Native American Linguistics*. University of Oregon. 411-429.
- Johnson, Jean B. 1962. El idioma yaqui. Instituto Nacional de Antropología e Historia. Departamento de Antropología e Historia, Publ. 10, México.
- Kurath, William and Edward H. Spicer. 1947. A Brief Introduction to Yaqui, a Native Language of Sonora. *University of Arizona Bulletin, Social Science Bulletin*, No. 15, Vol. XVIII, No. 1, Tucson, Arizona.
- Lindelfeld, Jacqueline. 1973. *Yaqui Syntax*. Berkeley: University of California Press.
- Lionnet, André. 1977. *Elementos de la lengua cahita*. México: Universidad Nacional Autónoma de México.
- Moctezuma, José Luis and Gerardo López. 1990. *Variación dialectal yaqui-mayo*.

- Noroeste de México 9. Hermosillo: Centro Regional Sonora del Instituto Nacional de Antropología e Historia. 94-106.
- Molina, Felipe y David Leedom Shaul. n/d. A Concise Yoeme and English Dictionary. Tucson Unified School District, Bilingual Education and Hispanic Studies Department, Tucson, Arizona.
- Molina Marilla, Hilario. 1994. Jiak Noki. Lengua Yaqui, Sonora. Primer Ciclo, Parte I. México: Secretaría de Educación Pública.
- _____. 1994. Jiak Noki. Lengua Yaqui, Sonora. Primer Ciclo, Parte II, México: Secretaría de Educación Pública.
- Rude, Noel. 1995ms. Diccionario morfémico de yaqui. Universidad de Sonora.
- Silva Encinas, Manuel Carlos, Pedro Álvarez and Crescencio Buitimea. 1998. Jiák nokpo etéjoim (pláticas en lengua yaqui). Hermosillo: Universidad de Sonora.
- The Hopi Dictionary Project. 1998. Hopi Dictionary. Hopiikwa Lavàytutuveni. A Hopi-English Dictionary of the Third Mesa Dialect with an English-Hopi Finder List and a Sketch of Hopy Grammar. Tucson: The University of Arizona Press.
- Zavala Castro, Palemón. 1989. Apuntes sobre el dialecto yaqui. Hermosillo: Gobierno del Estado de Sonora.



Inderivation in West Greenlandic **- in passive and other valence decreasing structures.**

Karen Langgård
Iisimatusarfik / University of Greenland

1. Definition of inderivation.

Incorporation by means of derivation, i. e. inderivation, is an all permeating feature of the syntax of the Inuit language, Greenlandic. My **definition of inderivation** is as follows (Langgård, 1993):

*In a stem x which is derived morphologically from a stem y, the stem y is syntactically **inderived**, if it still can retain one or more of its possible syntactic relations simultaneously with the derived stem x's syntactic relations.*

The tests rely

- on wordclass and the differences between their subordinated elements
- on coreference
- on scope - and meaning

In the 1980s, disagreements on Greenlandic incorporation sprang from the fact that incorporation in general is seen as a matter of compounding,¹ whereas it is a case of derivation in Greenlandic (see Mithun 1984, 1986; Sadock 1986). Because of this, I have introduced the term inderivation (Langgård 1993). The kind of inderivation based on nouns is the one that has been especially described and discussed, inaugurated by Jørgen Rischel in Rischel 1971, 1972. Later Jerrold Sadock has treated it several times (e. g. Sadock 1980, 1986, 1991, 2000).²

The focus of my presentation at the WAIL/ SSILA conference was on Verb Inderivation structures found in connection with passives and other valence-decreasing derivation. For a longer demonstration of Langgård, 1997, 2000 and forthcom.

Before advancing to the examples of inderivation, a little information about Greenlandic. It has 8 moods, 4 of which are superordinate moods, while the other 4 are moods that are primarily used as

subordinate verbs. The contemporative mood is one of the primarily subordinate moods. Its meaning resembles the English *ing*-form, and the contemporative mood form is subject coreferential with the next higher clause. This subject coreferentiality is an essential feature of the evidence of inderivation.

2. Introduction to the West Greenlandic form of Inderivation.

(EX 1) and (EX 2) are among the possible answers to the following question: *Kaalip Aani qanoq pivaa?* 'What did Kaali say to Aani?' / 'What did Kaali do to Aani'? Both of the examples have the same main verb including the same derivation by means of {qqu} 'to order him to _'

(EX 1)³

Suaarluni aggeqquaa

suaar-luni

cont 3.p c s

shouting

Shouting he ordered her to come

agger-qqu-aa

indic 3.p s-3.p s

he ordered her to come

(EX 1)⁴

Suaarluni aggeqquaa

(EX 2)

Taxarluni aggeqquaa

taxar-luni

cont 3.p c s

going by cap

He ordered her to come in a cap

agger-qqu-aa

indic 3.p s-3.p s

he ordered her to come

(EX 2)

Taxarluni aggeqquaa

The two examples are identical concerning derivation and inflection. Their syntactic structure is ambiguous. However, as is indicated in the glossing, due to the meaning of the lexemes, native speakers of Greenlandic will interpret the examples to differ in structure in the following way:

In (EX 1) the contemporative form *suaarluni* is subject coreferential with the subject of the next higher clause in accordance with the general rules of coreferentiality.

However, in (EX 2) the contemporative mood form *taxarluni* is an adverbial subordinate clause to the verbal action in the stem *agger-*. Even though further derivation has been added, the stem *agger-* still has some syntactic power in constituting the next higher clause in relation to *taxarluni*. This is an instance of inderivation.

The example below (EX 3) (source: a short story written by H. C. Petersen) is a normal, average Greenlandic sentence. It shows noun inderivation as well as verb inderivation:

(EX 3)

Niviarsiaqqanik iliarsunnik ikiorteqartarpoq anaanami arnat suliassaannut sungiusagaanik.
She had some orphan girl helpers, her mother's trainees in women's duties.

Niviarsiaqqanik (3.p)pl instr girls
iliarsunnik (3.p)pl instr orphans

Ikiorte-qar-tar-poq helper(s)-have-use to- indic 3.p s used to have helpers
anaanami POSS 3.p c s- (3.p)s rel her (own) mother's

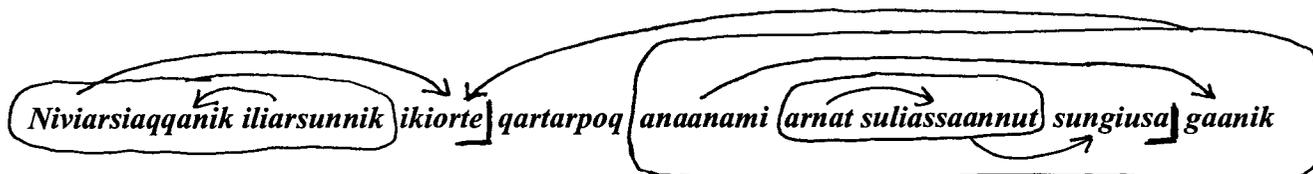
arnat (3.p)pl rel women's
suliassaannut POSS 3.p n-c pl- (3.p)pl term their 'works to be done'

sungiusagaanik
 POSS 3.p n-c s- (3.p)pl instr
 her trainees (i. e. her + passive nominalization of the transitive verb *sungiusarpa* (to train him, -mut/nut for something))

In Greenlandic the nouns are used both as nominal heads and as attributes to a nominal head.

Noun nderivation: The main verb *ikiorteqartarpoq* has an implicit subject (marked in the verb inflection, and known from the context). It is intransitive, but has an nderived (incorporated) object *ikiorte* by means of derivation (*{qar}* to have *_*). To this nderived noun stem a very long attribute in the instrumental case is subordinated: *Niviarsiaqqanik iliarsunnik* + *anaanami arnat suliassaannut sungiusagaanik*.

Verb nderivation: *sungiusagaanik* has a verb stem nderived by means of nominalization. In Greenlandic adverbial elements subordinated to a verb are retained unchanged when the verb is nominalized. The noun phrase (possessor + possessum) *arnat suliassaannut* with the head in the terminal case is still subordinate to the verb stem nderived into *sungiusagaanik*.



As shown in the examples above: An nderived verb stem retains to some degree the verb's power to create a clause and to create relations of coreferentiality.

However, Verb Inderivation is of two different kinds. (EX 4) and (EX 5) will show these two kinds:

(EX 4)

Naammagittarluta inuusariaqarnerput malugaa

<i>Naammagittarluta</i>	<i>inuusariaqar-ner-put</i>	<i>malugaa</i>
Cont 1.p pl	POSS 1.p pl – (3.p)s abs	indic 3.p s-3.p s
We not demanding much	our having to live	he noticed

He noticed that we have to live without demanding much.

Naammagittarluta ^{1.p pl} *inuusariaqar* ^{1.p pl} *nerput* *malugaa*

1.p pl

inuusariaqarnerput is a nominalization of the verb *inuusariaqar-* by means of the derivation {*niq*} (the act to , the fact that). It has a possessor-marking for 1.p plur. The possessor-marking shows what would have been subject in the corresponding clause, and it is the pivot for the coreferentiality in the contemporative mood form *naammagittarluta* in 1.p plur. I use the term 'personal inderivation' for this kind of inderivation, to distinguish it from an 'impersonal inderivation' as is illustrated in the following example:

(EX 5)

Naammagittarluni inuusariaqarneq paasivaa

<i>Naammagittarluni</i>	<i>inuusariaqar-neq</i>	<i>paasivaa</i>
Cont 3.p c s	(3.p)s abs	indic 3.p s-3.p s
one not demanding much	one's having to live	he understood

He understood that one has to live without demanding much.

Naammagittarluni ^{1.p pl} *inuusariaqar* ^{1.p pl} *neq* *paasivaa*

In (EX 5) the nominalization does not have any possessor to show who would have been the subject in the corresponding sentence, and the contemporative mood form can only be 3.p sing used in an impersonal neutral sense.

Nominalizations with their subordinated adverbial elements can be inflected in any of the 8 cases of West Greenlandic and can be used in any syntactic function filled by a noun.

As mentioned (EX 4) is a personal inderivation, while (EX 5) is an impersonal one. The inderivation in (EX 2) was by means of transitivity derivation. This kind of inderivation is always personal, while verb inderivations by means of valence decreasing derivation is always impersonal, even if the agent is explicit in a passive sentence, as will be shown in the next section.

3. Passive

In relation to one of the two passives of Greenlandic, it was noticed already by Samuel Kleinschmidt in the middle of the 19th century, that an odd construction existed (Kleinschmidt (1851) 1968). However, if the theory of inderivation is applied, nothing is odd. The structure we are concerned with is like the one in (EX 6):

(EX 6)

Nanoq qimmit atorlugit piniarneqartarpoq

<i>nanoq</i>	<i>qimmit</i>	<i>atorlugit</i>	<i>piniar-neqar-tar-poq</i>
(3.p)s abs	(3.p)pl abs	cont [3.p c s]-3.p n-c pl	indic 3.p.s
the bear	dogs	using	is customarily hunted

The polar bear is hunted with dogs

Nanoq (*qimmit atorlugit*) *piniar*] *neqartarpoq*

This is the frequent construction of inderivation by passive, i. e. a construction with a transitive contemporative mood form. Since the transitive contemporative mood forms most often do not have any subject marking, we then need a construction with an intransitive contemporative mood form to show that the form is impersonal. Such a construction is grammatically correct, too, and it turns out that such a contemporative mood form can only be in 3.p c sing and it can only have an impersonal meaning, even if the agent is added as it is in (EX 7):

(EX 7)

Qatserisartunit tuaviorani orninneqaramik navianartoorsiortut ilaat toqupput.

<i>qatserisartunit</i>	<i>tuaviorani</i>	<i>ornin-neqar-Gamik</i>
(3.p)pl abl	cont 3.p c s	caus 3.p c pl
by the firemen	not hasting	when they were met

<i>navianartoorsiortut</i>	<i>ilaat</i>	<i>toqupput</i>
(3.p)pl rel	3.p n-c pl-(3.p)s abs	indic 3.p pl
those in danger	some of them	they died

When they were met slowly by the firemen, some of the endangered died

Qatserisartunit (*tuaviorani*) *ornin*] *neqaramik* *navianartoorsiortut* *ilaat* *toqupput.*

Agent

impers 3.ps

4. Other valency decreasing nderivation.

Apart from passive, some other valency decreasing derivations are nderiving and in the impersonal way. First and foremost there is a derivation {nar} that means 'makes one _'. E. g. *NALUAA* means 'he does not know it' and *NALUNARPOQ* 'it makes one not to know' that is 'it is uncertain, not known'; *ALUTORAA* 'he admires it' and with {nar}: *ALUTORNARPOQ* 'it makes one admire' that is 'it is wonderful, admirable':

(EX 8)

Qaqqat isiginnaarlugit alutornarput

<i>qaqqat</i>	<i>isiginnaarlugit</i>
(3.p)pl	cont [3.p c s]-3.p n-c pl
the mountains	when looking at them
<i>The mountains are admirable to look at</i>	

<i>alutore-nar-put</i>
indic 3.p pl
they make one admire them

Qaqqat (*isiginnaarlugit*) *alutor* } *narput*

5. Some non-coreferential verb forms in supposedly coreferential context.

Till now the nderivation theory has been used to explain how some coreferential verb forms can be used with nominalizations, transitivations and passives etc. It has been used to show the pivot of the coreferentiality of some coreferential verbforms.

Now I would like to show some examples where we have a sentence with a main verb and a subordinated verb, and where the subjects of the two verbs are corefering. However, although the coreference is obvious, the subordinated verb is non-coreferential in form. The most frequent structure is a construction with the subordinated verb in causative mood form, the causative mood being used as verb in causal and / or temporal subordinate clauses, as in (EX 9):

(EX 9)

Kaali tillimmat soraarsinneqarpoq.

<i>Kaali</i>	<i>tillimmat</i>	<i>soraarsin-neqar-poq</i>
(3.p)s abs	caus 3.p n-c s	indic 3.p c
Kaali	when he stole	he got fired
<i>When Kaali stole, he (Kaali or another one) got fired.</i>		

Kaali tillimmat (*n.-c.*) *soraarsin* } *neqarpoq.*

The non-coreferential form is used, although it is ambiguous if the meaning of the lexemes or the broader context does not eliminate the ambiguity. However, once more this structure with the non-coreferential form can be accounted for by applying the theory of inderivation. The non-coreferential causative mood form is subordinated to the inderived stem *soraarsiC-*, and not to the whole matrix verb. The structure with the non-coreferential form, which at least today is the unmarked one, is even more astonishing because the coreferential form would be unambiguous, and furthermore because of the appearance of the paradigm of causative:

Its paradigm is as follows for intransitive forms of the lexeme *ANIVOQ* (goes out) - the transitive will show the same pattern -only with object-marking inflection added.

1.p s	ani-gama
2.p s	ani-gavit
3.p c s	ani-gami
3.p n-c s	ani-mmat
1.p pl	ani-gatta
2.p pl	ani-gassi
3.p c s	ani-gamik
3.p n-c pl	ani-mmata

It is obvious that any tendency towards regularisation will mean that the non-coreferential forms are the ones to be changed, and indeed, most native speakers will use the coreferential third person forms whenever the subject of the next higher clause is 1. or 2. person and thereby no ambiguity exists. A coreferential form subordinated to a passive matrix verb in 3. p sing is used too, but as the marked form. Native speakers seem to understand the coreferential form to mean that the point of view is placed with the Patient, the meaning equivalent to a report of his thoughts.

This pattern is not found in connection with valence decreasing inderivation by means of *{nar}*. A subordinated temporal-causal clause will not be subordinated to the inderived stem.

There is in this case a difference between passive by means of *{neqar}* and the valence decreasing derivations by *{nar}*.

However, both the passive *{neqar}* and the verbs derived by means of *{nar}* have the same fixed structure of indirect speech, whenever the inderived stem's lexeme sub-categorizes a +human object and an indirect speech. An example of this is *OQAATIGIVAA* 'he says about him or it, that', and so is *PAASIVAA* 'he understands about him or it, that':

(EX 10)

Oqaatigineqarput Nuummut aallarniartut

oqaatigi-neqar-put

indic 3.p pl

They are said to

They are said to have the intension to go to Nuuk

Nuummut

(3.p)s term

to Nuuk

aallarniartut

part 3.p n-c pl

they are intending to go away

(EX 11)

Paasinarput Nuummut aallarniartut**Paasi-nar-put**

indic 3.p pl

they make one understand that

Nuummut

(3.p)s term

to Nuuk

aallarniartut

part 3.p n-c pl

they are intending to go away

It is understandable that they are intending to go to Nuuk

(EX 10) and (EX 11) can be explained by the theory of inderivation. The verb of the indirect speech is not subordinated to the whole of the inquit verb, but only to the inderived transitive verb stem, *oqatige-* ('say about him that he _') and *paasi-* ('understand about him that he _'). The inderived stem has still retained some of its syntactic power to constitute a clause, that can govern the coreferentiality of its subordinate clause. This being so, both of the sentences in (EX 10) and (EX 11) in fact contain a matrix clause and two layers of imbedded clauses: the inquit + the inderived stem's "clause" + the indirect speech.

Oqaatigi | **neqarput** **Nuummut aallarniartut**
 n.-c.

Paasi | **narpur** **Nuummut aallarniartut**
 n.-c.

(EX 12)

.... taallat ilaat takuinnarlugit nalunarneq ajorput Endaleeqqamit pisuusut.**taallat**(3.p)pl rel
the poems**ilaat**3.p n-c pl- (3.p)pl abs
some of them**takuinnarlugit**cont [3.p c s]-3.p n-c pl
one just glancing at them**nalu-nar-neq ajorput**(compound) indic 3.p pl
they do not, not make it clear**Endaleeqqamit**(3.p)s abl
from Endaleeraq**pisuusut**part 3.p n-c pl
that they come / stem from*... if one just glances at some of the poems one will know that they are written by Endaleeraq.**AJORPAA* in connection with a *-neq*-form: he does not customarily*NALUAA*: he does not know it*NALUNEQ AJORPAA*: he always knows it (he does customarily not, not know it)*NALUNARNEQ AJORPOQ*: it is always obvious (it always makes one not not to know it)

.... taallat ilaat **takuinnarlugit** **nalu** | **narneq ajorput** **Endaleeqqamit pisuusut.**
 c. n.-c.

Without going into details, (EX 12) the inderived verb stem *nalu-* is activated twice as an inderived one: the contemporative subject coreferential mood form *takuinnarlugit* and the non-coreferential participial mood form *pisuusut* are both subordinated to this stem and not to the whole of *nalunarneq ajorput*.

6. Combinations and blockings.

It is not a frequent pattern to have more than one inderivation activated per word. However, they do occur. When they do, they follow Baker's mirror-principle (Baker, 1988), as can be seen in (EX 12), which I noticed in our local newspaper (in June 2001):

(EX 12)

Aatsaalli attartortunik najugarisaminnik pisisoqarpoq

<i>Aatsaalli</i>	<i>attartortunik</i>	<i>najugarisaminnik</i>
adv + clitic	(3.p)pl instr	3.p c pl-(3.p)s instr
But for the first time	tenants	their own dwelling

pi-si-Tuq-qarpoq

indic 3.p s

there are some who buy something

But for the first time there are some tenants who buy their own dwelling.

Aatsaalli *attartortunik najugarisaminnik* *pi* | *siso* } *qarpoq*

In the verb there is a nominal root *{pi}* 'something' inderivated by the means of *{si}* 'to buy something' and this stem is inderivated by the nominalizing *{Tuq}* 'one who _'. This stem *pisisoq-* is then inderived by means of *{qar}* 'to have _' or – as is the case here – impersonal 'there is _'.

The inderived *pi-*, being an inderived object, takes an attribute *najugarisaminnik* 'their own dwelling' in instrumental case, while the inderived stem *pisisoq-*, being an inderived object too, takes *attartortunik* 'tenants' as attribute in instrumental case.

However, as stated above, the structure with two activated inderivations in one word does not occur frequently. The frequent pattern used to express complex meanings is a chain of inderivations in the sense that the subordinated element of an inderived stem itself contains an inderived stem which takes another subordinated element and so on. The constrains as to how long such chains caused by a recursive application of inderivation can be, without becoming grammatically unacceptable, are comparable to those of relative clauses or coordinations in Indo-European languages.

For an example of a short chain, please see (EX 3) above.

The patterns of inderivation varies with the inderiving morpheme, but are there any blockings? I have already shown an example with two activated inderivations in one word. Obviously such a structure is not blocked, at least not always, since we have this example. In fact, it is not blocked, even when the lack of blocking can lead to ambiguity:

(EX 14)

.... . *Arnamik toqutsisoqarmat*

<i>arnamik</i>	<i>toqutsi-Tuq-qar-mat</i>
(3.p)s instr	caus 3. p n-c s
a woman	since there one who murders somebody

... *Since there was a murderer of a woman / a murderess*

The ambiguity springs from the fact, that *arnamik* 'women' can either be oblique object to the inderived halftransitive stem *toqutsi-* 'to murder someone' (and the meaning: one who murders someone who is a woman) or attribute to the inderived object *toqutsi-Tuq-* 'one who murders somebody' (and the meaning: one who murders someone and who is a woman).

.... . *Arnamik toqutsi* | *soqarmat*

.... . *Arnamik toqutsiso* | *qarmat*

On the other hand, blockings occur in some cases:

(EX 15)

Nukappiaraq ajunngitsunik soqutisaqarpoq

<i>Nukappiaraq</i>	<i>ajunngitsunik</i>	<i>soqutisaqarpoq</i>
(3.p)s abs	(3.p)pl instr	indic 3.p s
The boy	good	has interests

The boy has good interests

Nukappiaraq *ajunngitsunik* | *soqutisa* | *qarpoq*

If we put a transitivity derivation *{fige}* ('to have as an opportunity to _') on the verb stem, we could get:

(EX 16)

Nukappiaqqap timersorneq computerilerinerlu *ajunngitsunik soqutigisaqarfigai

Nukappiaqqap

(3.p)s rela

The boy

timersorneq

(3.p)s abs

sport

computerilerinerlu

(3.p)s abs + clitic

and use of computers

****ajunngitsunik***

(3.p)pl instr

good

The boy has good interests in sport and computers

soqutigisaqarfigai

indic 3.p s-3.p pl

has them as interests

**ajunngitsunik* can no longer be attribute to the inderived object *soqutigisaq-* after the further derivation {fige} has been added. We could only get an adverbial element as *ajunngitsunik* 'in a good way'.

Blockings may also be due to lexicalisations. On the other hand, when inventing new expressions, one often uses nominalisations and uses the patterns that are possible by means of the inderived verb stems.

To clarify all the aspects involved here, much more research is needed in the years to come.

7. Inderivation and general theories of linguistics.

What are the consequences of the inderivational structures in Greenlandic for general theories of linguistics? How does one describe these Greenlandic structures? Does the Lexicalist Hypothesis work? Should the border between morphology and syntax be removed?

Are the patterns of inderivation to be regarded as a feature of the lexemes of Greenlandic? Since the derivation is so dominant and so productive as it is in Greenlandic, it is not possible to make a synchronic list of lexicalised words. However, even if the lexicon then can be described as consisting of both lexicalised lexemes and non-lexicalised ones, this won't solve the problem either, since some of these lexemes then will have to sub-categorize for more than one set of syntactic relations.

Furthermore, the out-put of the word-formation cannot be described just as the sum of the derivations, because there might be some blockings - and lexicalisations too. The language users seem not only to have access to use lexicalised items, but seem to actively produce lexemes and seem to have some unconscious knowledge of the lexemes as constructed out of various elements. The language user seems to have some unconscious knowledge of possible combinations concerning inderivation.

Research about acquisition of Inuit languages seems to show that children learn certain elements earlier than children learning such languages as the European ones. For instance, they become aware of derivational morphemes much earlier and learn how to use them in a productive way (Fortescue and Olsen 1992, and especially: Allen, 1997).

Still, by retaining the border between morphology and syntax, one will get the simplest and most economical description of Greenlandic. However, at the same time it is necessary to accept that in a language with a typology as Greenlandic, the syntactic process does not only have 'access' to the

features of inflection, but also 'access' to the features of the nderiving derivation, which is to say that one has to accept that some words in Greenlandic have more than one set of grammatical relations simultaneously, and that this is part of an active production - and can be found in connection with lexicalised lexemes but most often in connection with non-lexicalised lexemes produced on the spot.

References

Allen, Shanley:

- 1997: *Aspects of Argument Structure Acquisition in Inuktitut*. Amsterdam/Philadelphia.

Baker, Marc:

- 1988: *Incorporation*. Chicago Press.

Fortesque, Michael and Lise Lennert Olsen:

- 1992: The acquisition of West Greenlandic. IN: Dan Isaac Slobin (ed): *The Crosslinguistic Study of Language Acquisition*. Vol 3, 111-220. Hillsdale NJ.

Katamba, Francis:

- 1993: *Morphology*. Modern Linguistics Series.

Kleinschmidt, Samuel:

- 1851(1968): *Grammatik der Grönländischen Sprache*. Hildesheim.

Langgård, Karen:

- 1993: Inderivation af verbalstammer. IN: *Grønlandsk kultur- og samfundsforskning 93*.
Ilisimatusarfik / Atuakkiorfik.

- 1997: Coreferentialitet, oratio obliqua og nderivation – i grønlandsk. IN: *Grønlandsk kultur- og samfundsforskning 93*. Ilisimatusarfik / Atuakkiorfik.

- 2000: 'Indérivation en kalaallisut, Groenland occidental'. IN: Nicole Tersis og Michèle Therrien (eds.): *Les langues eskaléoutes. Sibérie, Alaska, Canada, Groënland*. Sciences du langage. CNRS editions. Paris.

- Forthc.: 'Inderivation in Greenlandic' in an anthology on *Incorporation*, by Ole Nedergård Thomsen et al. IN: *Travaux du cercle linguistique de Copenhague*.

Mithun, Marianne:

- 1984: The evolution of noun incorporation. IN: *Language* 60.847-894.

- 1986: On the nature of incorporation. IN: *Language* 62.32-37.

Rischel, Jørgen:

- 1971: Some characteristics of noun phrases in West Greenlandic. IN: *Acta Linguistica Hafniensia* 12.213-45.

- 1972: Derivation as a syntactic process in Greenlandic. IN: Ferenc Kiefer (ed.): *Derivational processes*, 60-73. Stockholm.

Sadock, Jerrold:

- 1980: Noun incorporation in Greenlandic: A case of syntactic word formation. IN: *Language* 56.300-319.
- 1986: Some notes on noun incorporation. IN: *Language* 62.19-31.
- 1991: *Autolexical syntax*. Chicago. University of Chicago press.
- 2000: *Morphologie dérivationnelle en quatre dimensions. Kalaallisut, groenland occidental*. IN: Nicole Tersis og Michèle Therrien (eds.): *Les langues eskaléoutes. Sibérie, Alaska, Canada, Groënland*. Sciences du langage. CNRS editions. Paris.

Notes

¹ This view is still current in introductions, cf e. g. Katamba (1993).

² For more details on the reseach done on Incorporation / Inderivation, cf Langgård, forthcom.

³ Abrivations: indic = indicative; caus = causative (a mood form used to form temporal /causal clauses); part = participial; cont = comtemporative; abs = absolutive; rel = relative; term = terminalis / allative; instr = instrumental / modalis; abl = ablative; c = coreferential; n-c = non-coreferential.

⁴ The notation: the arrows point from the subordinated element toward the head. The head is a word or an inderived stem. The arrows are above the elements whenever the head is a nominal one, below whenever it is a verbal one. The circles inscribe phrases. The notation-system can be described as immediate constituents with dependences added.

Positional Verbs and Relational Nouns in Zaniza Zapotec

Natalie Operstein

University of California, Los Angeles

1. The Zapotec family of languages is spoken mainly in the State of Oaxaca, Mexico. Zapotec languages possess a number of classificatory elements that can enter the systems of nominal, pronominal, numerical, and verbal morphology. Within the realm of nominal morphology, the classificatory elements can take the form of prefixes on nouns that denote living things; the sole reconstructible function of these prefixes is that of animacy markers¹ (cf. Lachixío Zapotec *ni=hyo* ‘man’, *ni=ko* ‘dog’ and Zaniza Zapotec *bi=nya* ‘woman’, *bi=ziny* ‘mouse’, *bi=dxidx* ‘pineapple’). The animacy markers present just one instance of the general Zapotec (and Otomanguan) tendency to use superordinate terms in compounds with nouns to indicate lexical classes: the same process can be observed, for example, in Amuzgo and Mixtec (Smith Stark and Tapia García 1984:212; Macri 1983). A great number of Zapotec languages currently distinguish between alienably and inalienably possessed nouns, while others preserve at least traces of this distinction. Third person pronouns form the basis of a semantic division of Zapotec nouns into three to six gender classes. Of these, only the division into animates and inanimates seems to be common to all Zapotec languages, while the further subdivision of the animate class into human and animal, and of the human subclass into smaller categories based on such parameters as age, sex, sacredness, and relative social status of the referent, are mostly language-specific (Munro 2000 and 2001; Operstein 2001).² The earliest recorded form of Zapotec makes an extensive use of numeral classifiers that classify nouns mostly according to shape and animacy (Córdova 1578b:197-98; Operstein 2001). This usage is attested, in varying degrees, in modern Zapotec languages as well (cf. Munro 2001). The striking peculiarity of both the lexical and numeral classification of Zapotec nouns, however, consists in the fact that, at least synchronically in any given language, they do not cover all of the relevant vocabulary. For example, many of the nouns that denote living things in Zaniza Zapotec do not carry animacy markers, while a number of objects in Córdova’s (1578b:197) description of numeral classification are being counted without the use of classifiers (among them such obviously shape-endowed objects as pans and pumpkins).

The positional verbs, which classify nouns according to their spatial orientation, make up a closed class of forms. They form such an essential part of the Zapotec verbal system, that they received a fairly extensive treatment already in the first Zapotec grammar ever produced, published in the second half of the sixteenth century by Juan de Córdova. The following fragment is typical of Córdova’s Latin-based but eminently adequate description of this phenomenon: ‘El segundo compuesto es. Adsum, estar presente. Este se dize por muchas maneras. . . . Si estoy en pie, naçoa, si assentado, tipeea. Si echado, naaya . . . y assi los diferencias conforme ala postura o asiento de la cosa de que hablan’ (Córdova 1578b:42). Since Córdova’s early treatment, however, no study of positional verbs in any Zapotec language appeared in print, not even as part of the descriptive grammars of the individual languages.

The present paper is intended as a preliminary description of the semantics of positional verbs in just one language, Zaniza Zapotec, a member of the Western branch of the family. The data for this description consist of elicited responses to three sets of pictures (around two hundred sentences), and two spontaneous narratives. The Zaniza forms presented here have been supplemented, wherever possible, by data from several other Zapotec languages.³ The focus of the present study being purposely limited to the semantic characterization of the positional verbs, most details of their morphology have been omitted from consideration, and the verbs are identified by their stems.

2. The positional verbs in Zaniza can be divided into two unequal groups based on the dimensions along which they classify the spatial orientation of the object. The first group of verbs refers to the position and to some extent shape of the objects, and is composed of the following three verbs:

(a) Zaniza *zub* ‘to be seated’ (of animates), ‘to be placed, to stand’ (of inanimates). In case of both animates and inanimates, the verb is used of objects that can be conceived of as having a visible base. In Zaniza, this verb can be used of a man sitting at the table, of a dog sitting on the floor, of a tree growing on a hill top, and of cups, lamps, or bottles standing on the table. The causative of *zub* means ‘to put, place’. Related to *zub* are such forms in other Zapotec languages as Yatzachi El Bajo *zob* ‘estar presente’, Chichicapan *zobã* ‘está sentado’, SLQZ *zùub* ‘is sitting; is located, exists (sitting of projecting)’, Córdoba’s Zapotec *nazoba (quete)a* ‘estoy (en lo baxo)’, *naçoba(peeza)ya* ‘estoy (esperando)’, *naçobaya (xicheni / laoni)* ‘estoy (defuera)’, Reyes 1700 Valley Zapotec *Rizoo baya* ‘estar’, Anonymous 1823 *Rizoobaya*, *ricoobaya* ‘estar derecho o en pié’, ‘estar absolutamente’. Although not genetically related, the following forms are semantically parallel: Córdoba’s Zapotec *tipeea* ‘estoy assentado’, Reyes 1700 Valley Zapotec *narija* ‘estar sentado’, Zoogocho *dxi* ‘sit’ (of persons, animals, and things) and possibly *zhia* ‘sit’ (of inanimates and persons on a horse). The following is an example of the use of this verb in Zaniza:

txun	almet=e	zub=ny	lo	mezh . . .	(PV:46)
three	bottle=deictic	is=3s inanimate	face	table	
Three bottles are standing on the table . . .					

(b) Zaniza *zu* ‘to stand’, ‘to be’ is used of vertically oriented objects that have no visible base. This verb can be used of a cat standing on the floor, ladder up against a wall, or a table standing on its legs. The causative of *zu* means ‘to put, place’. This verb is matched by such cognates as Chichicapan *zũ* ‘está parado’, SLQZ *zuu* ‘is standing, is located (standing)’, Zoogocho *zoa* ‘to be standing (of inanimates)’, ‘to be’, ‘to live’, Córdoba’s Zapotec *naçoa* ‘estoy en pie’, *tizoo(lahui)a*, *naço(lahui)a* ‘interuenir o entreuenir’, *naço(nñij)a*, *naço(nñij)ni* ‘estoy, esta en pie hablando’, *naçoo(pani)a* ‘vivir’, *nazoo(yago)a*, *naço(zago)ni* ‘estar en pie comiendo’, Reyes 1700 Northern Zapotec *rizooa*, *baazooa* ‘estar’, Reyes 1700 Valley Zapotec *riizooya* ‘estar en pie’, Anonymous 1823 *rizooya* ‘estar en pié’. It is unclear whether Zoogocho *zee* ‘to stand (of animates)’ and Yatzachi El Bajo *ze* ‘estar parado’ are related. At least in Zaniza and Zoogocho, this

verb is also used as a general existential predicate meaning something like ‘to exist’ or ‘there is’, by extension ‘to live’. Both uses of *zu* can be exemplified as follows:

igey guj-bit=i, zu=y lat nis-gey (PP 5)
 man be sad=3s m stand=3s m in the middle of rain
 The man is sad, he is standing in the rain.

zu tib lyez (CT:1)
 be one opossum
 There was an opossum.

bany-bala, a zu=r? -- zu=ã (CT:2)
 old lady Q be=2s be=1s
 - Old lady, are you there? – I am.

By incorporating adverbials, *zu* can qualify the manner in which the object is standing, cf. *zu-dxib* ‘to stand on edge’, *zu-tyily* ‘to stand upside down’, *zu-gwas* ‘to stand inclined’. In Zoogocho, the first two notions are expressed by two lexically distinct positional verbs, cf. *shtulhe* ‘to be lying on side’, *ndosa* ‘to be upside down’.

(c) Zaniza *mix*, morphologically the stative of *bix* ‘to fall’, means ‘to be lying’, and is used with both animates and inanimates. It can refer to a cat lying on the floor, spoon lying on the table, a ball or a rolled-up rope lying on the ground. The causative of *bix* means ‘to fell, drop’. An exact morphological and semantic parallel is provided by Córdoba’s Zapotec *nábixia* ‘estar echado’, and Whitecotton and Whitecotton 1993 *nabixi* ‘revolcado estar’. Zoogocho *zxoá* ‘to lie in an extended fashion’, Yatzachi El Bajo *xoa* ‘estar extendido’, Chichicapan *nâ’a* ‘está acostado’, SLQZ *nàa’tga’ah* ‘is lying down, is (located) in a lying position’ (*nàa* ‘is; is (located) in, is a part of; exists in’), Córdoba’s Zapotec *naaya* ‘estoy echado’, *naa(yaci)a* ‘estar echado dormiendo’ and *tixóbaya*, *naxóbaya* ‘estar recostado’, although not cognate with the Zaniza form, are semantically related. An example of the use of this form would be:

kwit bay-ij=e mix tib kutyar (MB:24)
 side handkerchief=deictic lie one spoon
 Next to the handkerchief there lies a spoon.

The second, larger group of Zaniza positional verbs consists of the verbs that refer to the manner in which the objects are placed on the supporting surface. The spatial orientation of the objects becomes irrelevant, but in the case of one (or possibly two) verbs their number assumes importance:

(d) Zaniza *(y)ub*, *(b)ib* ‘to be (placed) on a surface’ (of a single object). This pair of verbs appear to be synonymous as they can replace each other without the apparent change of meaning. They can be used of all kinds of objects on all kinds of surfaces, for example a box or a tablecloth on the table, a boat on the water, a sweet potato on the stump, a ball on a flat stone. The causative morphologically corresponding to the first of

the two verbs means 'to put, place'. It is unclear whether any of the following are related to either of the Zaniza forms: SLQZ *bèe'b* 'is located on', Chichicapan *dzi'iba* 'está puesto encima', Yatzachi El Bajo *chi'ib* 'estar sentado', Whitecotton and Whitecotton 1993 *chiba* 'estar encima de algo', Córdova's Zapotec *tichijbaya, nachijbaya* 'estar encima como sobre algo', Zoogocho *ndobe* 'to be folded or wrapped around'. An example of the use of these verbs is:

gu=e rib=ny/ub=ny ru7 yag-bity (PV:23)
 sweet=deictic lie=3s inanimate edge stump
 potato

The sweet potato is (lying) on the stump.

(e) Zaniza *ngwa7* 'to be placed on a surface' (of multiple objects). This can be said of beans scattered on the table and balls lying on the ground. The causative of this verb means 'to place multiple objects on a surface' (examples given by the speaker included placing pieces of fruit on the table and pieces of meat on the grill). Zoogocho *nkwa* 'to be heaped' is cognate with the Zaniza form, while Chichicapan *rĩ* 'está a granel' provides a semantic match:

del pelote ngwa7=ny lo mezh (PV:8)
 all balls are=3s inanimate face table

All the balls are on the table.

(f) Zaniza *ka* 'to be stuck on a surface', 'to have a close contact with the surface' (of one object, or of multiple objects conceived as separate entities). This verb can be used of a fly on the wall, letters written on a T-shirt, medallion on a chain, orange or oranges on a tree, picture on a stamp, band wrapped around a candle. The causative of this verb means 'to stick on'. This verb is cognate with Chichicapan *kă* 'está pegado' and SLQZ *càa* 'is hanging, is hung, is located; is clinging to, is spread on':

tyug-lyeg=e ka=m kwit yu7 (MB:52)
 snail=deictic be=3s animal wall house

The snail is on the wall.

(g) Partly overlapping with the preceding is *ta7* 'to be stuck on a surface', which is used of mass objects (such as butter or paint) or multiple objects conceived as an undifferentiated mass (for example, fruits on a tree). The causative, consequently, means 'to spread' and 'to paint'. This verb is cognate with Zoogocho *daa* 'to be stuck on' and Yatzachi El Bajo *da'a* 'estar pegado':

ur=wi ix ta7=ny to7 yag (MB:45)
 orange sweet be stuck=3s inanimatehead tree

The oranges are on the tree.

(h) Zaniza *yu7* ‘to be inside an enclosed space’, ‘to be’ (in general). The causative of *yu7* is an intransitive verb that means ‘to enter’. Cognates to this verb include Zoogocho *yoo* ‘to be inserted in, wrapped around’, Yatzachi El Bajo *yo’o* ‘estar adentro’, Chichicapan *nñu* ‘está metido’, SLQZ *nu’uh* ‘exists (in a location)’ [often has a habitual sense], Córdoba’s Zapotec *nayooa*, *nooya*, *tooya*, *noo*, *nooya(lani)ni* ‘estar dentro en algo’, *noo yagoa*, *nooyagoni* ‘estoy comiendo, estar comiendo’, Reyes 1700 Valley Zapotec and Anonymous 1823 *nooya* ‘estar dentro’, *huayooya* ‘estar dentro de otra cosa’. The locative meaning of this verb is exemplified below:

almet=e	yu7=ny	xi	gitx-kiw	(PV:60)
bottle=deictic	is=3s inanimate	belly	basket	

The bottles are in the basket.

(i) The basic meaning of Zaniza (*b*)*iny* is ‘to be in the middle of’, whether the supporting surface is as an enclosed or open space. The causative of this verb means ‘to place in the middle’. The various uses of (*b*)*iny* can be gleaned from the following examples:

txi-giny	riny=m	xi	yag	(MB:67)
bird	be=3s animal	belly	tree	

The bird is in the middle of the tree (=in the hollow of the tree).

seya	riny=ny	xi	sobra	(MB:3)
stamp	be=3s inanimate	belly	envelope	

The stamp is in the middle of (=on) the envelope.

bej=e	riny=ny	yeny=m	(MB:51)
necklace=deictic	is=3s inanimate	neck=3s respect	

The necklace is in the middle of (=around) her neck.

(j) Zaniza *zab* ‘to be freely hanging, or floating (in air or water)’, *zab-giny* ‘to be placed in a hanging position’ has cognates and/or semantic matches in Chichicapan *zá’abi* ‘está colgado’, SLQZ *zèèi’by* ‘is hanging’, ‘is floating (in air or water)’, Córdoba’s Zapotec *nazaabi* ‘colgado’, Zoogocho *nala* ‘to hang’, *zehe* ‘to hang’ (especially at a higher altitude, like an electrical wire), Yatzachi El Bajo *ze* ‘estar colgado’. Examples of the use of both verbs can be seen below:

za	zab=ny	to7	lo-gi	(MB:36)
cloud	hang=3s inanimate	head	hill	

The cloud is hanging above the hill.

nyedy	zab-giny=ny	lo	du-gedx	(MB:37)
clothes	hang=3s inanimate	face	rope	

The clothes are hanging on a line.

3. The following remarks pertain to the use of positional verbs in existential, locative, and possessive sentences.

(a) Zaniza Zapotec has a basic VSO word order. A sentence such as the one below, which has the unmarked word order, can have both existential and locative readings:

zu igey gi7 yag (PP 12)
stand man foot tree

There is a man under the tree. / The man is standing under the tree.

The intended meaning can be clarified by utilizing the focussed position before the verb. The locational reading can be shown by placing the nominal in this position (notice also the relative order of the nominal and the locative phrase, cf. Clark 1978):

tor kas gur ze7t zu=m (SS:7)
bull black very far stand=3s animal

The black bull is standing very far.

The existential reading can be emphasized by placing the locative phrase before the verb:

xi bej=e ka do7 (MB:57)
belly necklace=deictic be stuck saint
On the necklace there is a saint (=an image).

i zu tib mezh, lo mezhe zu tib dya7n . . . (SS:11)
here stand one table, on table stand one plate
Here there is a table, on the table there is a plate . . .

(b) The Zaniza verbs most frequently used as general existential predicates are *yu7* and *zu* (*zoa* is used similarly in Zoogocho). The concrete locative meanings of *yu7* and *zu* seem to be completely suppressed when they are used in this function.

(c) As far as could be determined, only two positional verbs are used in possessive constructions in Zaniza: *zu* seems to be used when the possessed object is animate, *mix* when it is inanimate, as in the following examples:

zu binya=y
be wife=3s m
He has a wife.

mix gid-ij=ã
lie cover-head=1s
I have a hat.

4. The data on spatial relationships in any Zapotec language would not be complete without some reference to spatial prepositions and relational nouns. The number of prepositions in Zapotec languages is in general very small; in some, the only true prepositions have been borrowed from Spanish. Relational nouns, on the other hand,

form an essential part of the grammar not only of Zapotec, but also of other Otomanguean languages (cf. Brugman 1983; Hollenbach 1990). For Zapotec, no general comparative study of relational nouns is available, but this subject has been addressed in Lillehaugen (2001a and 2001b), Smith Stark (2000), and, much more comprehensively, in MacLaury (1989). The relational nouns of Zaniza Zapotec include the following forms:

(a) Zaniza *lo* ‘face’, ‘eye’. As a relational noun: ‘on’, ‘above’, ‘in front of’, ‘next to’, and goal in general, including indirect object marking. This form is common Zapotec, cf. Chichicapan *lo* ‘cara’, Ayoquesco *lo-* ‘face’, ‘in front of’, SLQZ *lohoh* ‘face’, ‘on, in front of’, Córdova’s Zapotec *láo(ni)* ‘cara o rostro de animal’, *lào* ‘sobre’, ‘en, vt en la pared en la cruz’, ‘encima de algo’ (1578:158, 382, 161), colonial Zapotec texts *lao*, *loo* ‘in, to’. The prepositional use of *lo* can be seen in the sentence below:

pelot=e	mix=ny	lo	yu	(PV:7)
ball=deictic	lie=3s inanimate	on	ground	

The ball is (lying) on the ground

(b) Zaniza *ru7* ‘mouth’: ‘at the edge of’. This likewise appears to be common Zapotec in both its literal and extended uses, cf. Chichicapan *ruaa* ‘boca’, Ayoquesco *ro’o-* ‘lips, mouth’, SLQZ *ru’uh* ‘mouth’, ‘at the edge of’, Córdova’s Zapotec *tòhua*, *tòua*, *tòa* ‘boca generalmente’, *tóa* ‘en, vt en la pared en la cruz’ (1578:158), colonial Zapotec texts *rua*:

xal=e	ub=ny	ru7	gitx-kiw	(PV:2)
band=deictic	be placed=3s inanimate	edge	basket	

The band is (lying) on the side of the basket.

(c) Zaniza *gi7* ‘foot’: ‘at the bottom of’, ‘under’ is matched by Ayoquesco *ye’e-* ‘foot, foreleg’, ‘at the foot of’ and SLQZ *ni’ih* ‘foot’, ‘under, beneath’:

zu	igey	gi7	yag	(PP 12)
stand	man	foot	tree	

There is a man under a tree. / A man is standing under a tree.

(d) Zaniza *xi* ‘belly’: ‘inside’, ‘in the middle of’ is semantically matched by Chichicapan *la’an* ‘estómago’, Ayoquesco *la’ayn-* ‘stomach’, ‘belly’, ‘inside’, SLQZ *là’iny* ‘stomach’, ‘in’, Córdova’s Zapotec *làni* ‘barriga’, ‘delantera de olla, cantaro, arbol pared cruz o todo assi’, *làni* ‘en, vt en la pared en la cruz’, ‘en por dentro de’ (1578:158), colonial Zapotec *lani(j)* ‘in’:

tyup	pelot=e	riny=ny	xi	gitx-kiw	(PV:56)
two	ball=deictic	be placed	belly	basket	
		in the middle			

Two balls are in the basket.

instance, the verb *(b)iny* ‘to be in the middle’ does not take relational nouns whenever the surface is a body-part, while it does require relational nouns in other cases, cf. the following sentences:

seya riny=ny xi sopra (MB:3)
 stamp be=3s inanimate belly envelope
 The stamp is on (=in the middle of) the envelope.

bay-ij=e riny=ny to7=m (MB:46)
 kerchief=deictic be=3s inanimate head=3s respect
 The kerchief is on (=in the middle of) her head.

(k) Spatial prepositions are few in Zaniza and include the following: *neny* ‘inside’, ‘under’ (its scope overlaps with those of *xi* ‘belly; inside’ and *gi7* ‘foot; under’); *lat* ‘in’, ‘among’, ‘while’, ‘during’, and *zew* ‘in the middle of’, ‘while’.

Conclusion. In Zaniza Zapotec two different groups of positional verbs can be distinguished depending on their main semantic stress. The first group, which in Zaniza is composed of three verbs, classifies objects in terms of their spatial orientation (*zub/zu/mix*: sitting/standing/lying). Departures from the canonical orientation of the object are taken care of by incorporating adverbials. The second group, which in Zaniza contains nine verbs, places stress on the manner of placement of the object on the supporting surface (or what is perceived as such) (*(y)ub/(b)ib/ngwa7* ‘to be on the surface’, *ka/ta7* ‘to be stuck on the surface’, *(b)iny* ‘to be in the middle of the surface’, *yu7* ‘to be inside the surface’, *zab/zab-giny* ‘to be hanging on or floating on or over the surface’). In one case, the objects are differentiated with respect to their number, single versus multiple (*(y)ub*, *(b)ib/ngwa7* ‘to be on the surface’), and in another what might have originally been a similar distinction is partially obscured (*ka/ta7* ‘to be stuck on the surface’).

The semantic difference between the two groups of verbs is especially noticeable in the pattern of their cooccurrence with relational nouns and spatial prepositions. The verbs that focus on the spatial orientation of the object show no preference for any particular group of relational nouns, while the verbs that have to do more with the characteristics of the surface do seem to show such a preference. For example, the verb *(b)iny* ‘to be in the middle’ occurs in my sample only with *xi* ‘in the middle of’, *neny* ‘inside/under’ and *lat* ‘in the middle of’, while *(y)ub* ‘to be on’ occurs only with the relational nouns that characterize the surface, namely *titx* ‘on’, *ru7* ‘at the edge of’, *lo* ‘on’, and *to7* ‘on’.

Acknowledgments

I thank Dr. Terrence Kaufman and Dr. John Justeson for the opportunity to work on Zaniza Zapotec as part of the Project for the Documentation of the Languages of Mesoamerica during the field seasons of 1999 and 2000. I am also thankful to Dr. Pamela Munro for her helpful comments, to Mark Sicoli for data on Lachixío Zapotec, and to

Aaron Sonnenschein for making available to me Melissa Bowerman's Topological Relations Picture Series and Picture Series for Positional Verbs.

Abbreviations

MB: Melissa Bowerman's Topological Relations Picture Series
PV: Picture Series for Positional Verbs
PP: project pictures
SS: Snack Surprises
CT: El Cuento del Tlacuache

Notes

¹Terrence Kaufman, personal communication; also Kaufman 1994.

²MacLaury (2001) is cited in Munro (2001:note 26) as the most comprehensive survey of Zapotec pronominal systems.

³The data for Zaniza Zapotec is based on my field notes, that for Chichicapan Zapotec on Smith Stark 2000, for SLQZ on Munro and Lopez 1999 and Lillehaugen 2001b, for Zoogocho on Sonnenschein 2000, for Yatzachi El Bajo on Butler 1980, and for earlier dialects on Córdova 1578a and 1578b, Reyes 1700, Anonymous 1823, Whitecotton and Whitecotton 1993, and Lillehaugen 2001a.

References

- Anonymous. 1887 [1823]. Gramática de la lengua zapoteca, ed. Antonio Peñafiel. México.
- Brugman, Claudia. 1983. The use of body-part terms as locatives in Chalcatongo Mixtec. *Survey of California and Other Indian Languages* 4:235-90.
- Butler, Inez M. 1980. Gramática zapoteca: zapoteco de Yatzachi El Bajo. México, D.F.: Instituto Lingüístico de Verano.
- Clark, Eve V. 1978. Locational: existential, locative, and possessive constructions. *Universals of human language*, ed. by Joseph H. Greenberg, vol. 4, pp. 85-126. Stanford: Stanford University Press.
- Córdova, Juan de. 1987 [1578a]. Vocabulario en lengua zapoteca. México: Ediciones Toledo.
- _____. 1886 [1578b]. Arte del idioma zapoteco. Morelia: Imprenta del Gobierno.

- Hollenbach, Barbara E. 1990. Semantic and syntactic extensions of Copala Trique body-part nouns. *Homenaje a Jorge A. Suárez*, ed. by Beatriz Garza Cuarón and Paulette Levy, pp.275-96. México: El Colegio de México.
- Kaufman, Terrence. 1994. Proto-Zapotec Reconstructions. Ms.
- Lillehaugen, Brook Danielle. 2001a. A preliminary look at prepositions in colonial Zapotec documents. UCLA American Indian Seminar handout.
- _____. 2001b. Acquisition of body part prepositions in San Lucas Quiaviní Zapotec. UCLA American Indian Seminar handout.
- MacLaury, Robert E. 2001. Zapotecan personal pronouns: typology and reconstruction. Ms.
- _____. 1989. Zapotec body-part locatives: prototypes and metaphoric extensions. *IJAL* 55:119-54.
- Macri, Martha J. 1983. Two noun class systems in Mixtec. *Survey of California and Other Indian Languages* 4:291-306.
- Munro, Pamela. 2001. Hierarchical pronouns in discourse: third person pronouns in San Lucas Quiaviní Zapotec narratives. *Southwest Journal of Linguistics* (to appear).
- _____. 2000. Elaboration in Valley Zapotec pronouns. Paper presented at the conference *La Voz Indígena de Oaxaca*, Los Angeles.
- _____, and Felipe H. Lopez. 1999. *Di'csyonaary X:tè'n Diizh Sah Sann Lu'uc*. 2 vols. Los Angeles: UCLA Chicano Studies Center Publications.
- Operstein, Natalie. 2001. Personal pronouns and Zapotec dialectology. Ms.
- Reyes, Gaspar de los. 1891 [1700]. *Artes de las lenguas cerrana y del valle*. Oaxaca.
- Smith Stark, Thomas C. 2000. Relaciones espaciales: zapoteco de Chichicapan. PDLMA handout.
- _____, and Fermin Tapia García. 1984. Los tonos del amuzgo de San Pedro Amuzgos. *Anales de Antropología* 21:199-220.
- Sonnenschein, Aaron. 2000. Positional verbs in Zoogocho Zapotec. UCLA American Indian Seminar handout.
- Whitecotton, Joseph W., and Judith Bradley Whitecotton. 1993. *Vocabulario zapoteco-castellano*. Nashville, Tennessee: Vanderbilt University Publications in Anthropology.

**Cross-referencing in Emérillon (Tupi-Guarani):
a hierarchical agreement system**

Françoise ROSE¹

Université Lyon II (France), Laboratoire Dynamique du Langage – IRD
francoise.rose@etu.univ-lyon2.fr

This paper will deal with the cross-referencing system of Emérillon and consider the distribution of the two person markers sets which determine the kind of agreement system Emérillon displays. Cross-referencing is one of the major themes of Tupi-Guarani morpho-syntax, which has been studied both within particular languages (Harrison, 1994 for Guajajara; Seki, 1990 for Kamaiurá; Rodrigues, 1990 for Tupinamba; Leite, 1990 for Tapirapé) and for the whole family (Jensen, 1990); Payne, 1994). Our aim is to point the specificities of the Emérillon system, which has so far not been included in the discussions due to insufficient description.

Emérillon is a language spoken only in French Guyana by about 400 speakers who call it Teko. It belongs to the 8th subgroup of the Tupi-Guarani family (Rodrigues, 1984-85). The Tupi-Guarani family consists of about 40 languages spread out through Brazil and neighbouring countries and "is noted for a high degree of lexical and morphological similarity among its member languages in spite of their extensive geographical separation." (Jensen, 1999). We will therefore always refer to the analyses proposed for other Tupi-Guarani languages or for Proto-Tupi-Guarani. As far as the Emérillon language is concerned, very little work has been done to date: a short grammatical sketch of the language (Maurel, 1998), two articles on its morpho-syntax (Queixalós, 2001; Couchili, Maurel & Queixalós, n.d.) and a qualifying paper on the phonology of the language by the author of this paper (Rose, 2000).

The paper will first present the Emérillon cross-referencing system. It will then focus on two important challenges discussed in the literature on Tupi-Guarani cross-referencing systems, using new data on Emérillon: first, the hierarchy of person, then the hypothesis that such a system, which is based on a hierarchy of person is an inverse system. After considering these two points, we will readdress the issue of what type of cross-referencing system characterizes the Emérillon language: whether it is active/stative, nominative/accusative, absolutive or inverse.

1. The cross-referencing system of Emérillon, in relation with other Tupi-Guarani languages as they are usually presented²

Grammatical relation marking in Emérillon is realized mainly by cross-referencing person markers, which are divided in two main sets, traditionally called set I and II in Tupi-Guarani literature (Table 1).

	set I	set II
1 st sg	a-	e-
2 nd sg	ele-	de-
1 st incl	si-	nõde-
1 st excl	olo-	ole- / olone-
2 nd pl	pe-	pe- / pene-
3 rd	o-	i-
indeterminate	za-	zo-

Table 1: Emérillon person markers of set I and II.

We now observe the distribution of these two sets in order to determine what type of organization it marks. We will first look at their use in independent clauses (intransitive then transitive) then in dependent clauses.

1.1 Person marking in independent intransitive clauses

In these clauses, set I marks S, as shown in (1).

- (1) a-ʔita ʔi-pope
 1sg/I³-swim river-in
 I swim in the river.

As far as set II is concerned, it occurs on words that have often been described by tupi-guaranists as stative (intransitive) verbs called "descriptive" verbs (Jensen, 1998; Seki, 1990). These stative verbs are opposed to active verbs on the basis of the person markers they take and the agreement system is said to display a split-intransitivity pattern, where set I marks the S of active verbs (2a) and set II the S of stative verbs (2b).

Proto-Tupi-Guarani, Jensen 1998:

- (2) a- *a-ʔár
 1sg/I-fall
 I fall.
- b- *če-katu
 1sg/II-be good
 I am good.

But in fact, an alternative analysis seems to be spreading among linguists working on Tupi-Guarani languages, analysis to which we subscribe for Emérillon: that some words which refer to qualities or states and are often considered as adjectives or descriptive verbs, are actually not different from possessible nouns (Rodrigues, 1996; Couchili, Maurel & Queixalós, no date). Preceded with a set II marker, these words can be analyzed as nouns preceded by the possessive marker and (2b) should thus be glossed "my goodness". They can also predicate (3a), like any possessible nouns (3b). The same structure "set II marker plus noun" can constitute therefore either an argument or a complete predication.

(3a) e-kaneʔõ
1sg/II-fatigue
my fatigue / I am tired. (My fatigue exists.)

(3b) e-kija
1sg/II-hammock
my hammock / I have a hammock. (My hammock exists.)

It follows from this analysis, that there is in Emérillon no "split-intransitivity" between active and stative predicates, but a set of independent intransitive verbs all marked with set I prefix as in (1), and a set of "descriptive" words considered as possessible nouns and not as stative intransitive verbs.

1.2 Person marking in Emérillon independent transitive clauses

We must first note that only one of the two participants is marked on the verb in independent transitive clauses. The marking of either A (with set I) or P (with set II) seems to depend on their relative position on a person hierarchy⁴.

Three cases must be distinguished:

a) one participant is a speech act participant, the other a 3rd person.

This first case shows a clear hierarchy: 1>3 and 2>3, the participant higher on the hierarchy being marked on the verb, whether it is A or P. Thus when A1→P3 or A2→P3, only A is marked, with set I, as illustrated in (4):

(4) ele-nupã
2sg/I-hit
You hit him.

And when A3→P1 or A3→P2, only P is marked, with set II:

(5) zawal e-suʔu
dog 1sg/II-bite
A dog bit me.

b) both participants are third persons:

When a 3rd person acts on another 3rd person (A3 → P3), prefix *o-* of set I is systematically used, whatever the arguments may refer to.

- (6) patu-pope o-ijnuŋ
 pot-in 3sg/I-put

She puts them in the pot. (The third persons are interpreted as feminine and plural in context.)

This second case shows that the person hierarchy is irrelevant, another hierarchy takes over: A>P.⁵ The *o-* prefix of set I (marking either Ss or As) can logically be described here as an A marker. Its presence marks indirectly that P is also a third person: if P was a speech act participant, it would have to be marked instead of A due to its higher position on the person hierarchy. A consequence of these two hierarchies is that the *i-* prefix of set II never occurs on verbs. It is used only in nominal phrases (as a possessive marker) or in postpositional phrases.

c) both participants are speech act participants ("local" forms):

The hierarchy between first and second person is more difficult to establish and will be specifically dealt with in section 2.

Consider for instance how, when a second person acts on a first person (A2 → P1), A is marked with set I on the verb, and another person marker follows the verb (the meaning of which will be discussed in section 2.a).

- (7) ele-nupã eleŋ
 2sg/I-hit ?
 You hit me.

And when a first person acts on a second person (A1 → P2), some special forms are used: *olo-* (I/we → you singular), *apolo-* (I/we → you plural), forms to which we will return in section 2.b .

- (8) olo-ʔu-tal
 A1/P2-eat-FUT
 I'm going to eat you.

In conclusion, in the case of independent transitive verbs of Emérillon, a clear 1/2>3 hierarchy operates when only one speech act participant is involved, and a semantic hierarchy A>P when both A and P are third persons. Before focusing on the specific hierarchy problem encountered in the "local" forms, let us finish the presentation of the Emérillon agreement system with the dependent clauses.

1.3 Person marking in dependent clauses

In Emérillon, dependent verbs follow the same pattern as independent verbs, although in Proto-Tupi-Guarani the agreement system in "dependent constructions" (i.e. subordinate clauses, nominalizations, serial verbs and oblique-topicalized constructions) was reconstructed differently from the one in independent verbs.

On dependent Proto-Tupi-Guarani verbs, the only mark is always of set II, marking P on transitive verbs and S on all intransitive verbs. This system is said to be absolutive in the sense that what is marked on the verbs is P and S, with the same marks for P or S, just like in an ergative-absolutive system where only the absolutive is marked. Proto-Tupi-Guarani, Jensen 1998:

(9a) *o-có i-moʔé-βo
3/I-go 3/II-teach-SER
He went to teach him.

(9b) *kwecé i-ʔár-i
yesterday 3/II-fall-OBTOP
Yesterday he fell.

A 3rd set of person markers has been reconstructed for Proto-Tupi-Guarani. It marks an intransitive serial verb subject (S) when this subject is coreferential with the subject of the main clause, whether S or A (10).

Proto-Tupi-Guarani, Jensen 1998:

(10) *a-có wi-poracéj-ta
1sg/I-go 1sg/III-dance-SER
I went to dance.

This set has no reflex in Emérillon (except the third person *o-* coreferential with the subject, used in nominal phrases, and whose form is similar to the set I prefix on the verb and can be confused with it in this language). The coreferentiality of S with the subject of the main clause is not relevant: in these cases, the Emérillon agreement system is the same as in independent clauses.

The absolutive system does not function in Emérillon. Instead, the system which is characteristics of the Tupi-Guarani main verbs of independent clauses has been extended to all syntactic structures. Therefore in example (11a) from Emérillon, A is marked with set I on the relativized verb (where P would be marked with set II in Proto-Tupi-Guarani). In (11b), set I (not set II) is used to mark S on the verb of the temporal subordinate clause.

(11a) a-ekal kija ele-baʔe māʔě.
1sg/I-search hammock 2sg/I-make REL
I'm looking for the hammock you made.

- (11b) a-wig-a-nam, o-ho-pa.
 1sg/I-arrive-a-when 3/I-go-COMPL
 When I arrived, he had gone.

Example (12) shows that the coreferentiality of the three subjects is not indexed by a special set of marker: set I is used also on dependent verbs.

- (12) a-wut-tal a-ho t-a-po?o inga
 1sg/I-go up-FUT 1sg/I-go purpose-1sg/I-pick sp.fruit
 I'm going to come up to pick the fruit.

To summarize, the main points of the agreement system of Emérillon are:

- On transitive verbs, only A or P are marked (with set I and II respectively) according to their relative position on a person hierarchy or on a semantic hierarchy, according to the persons involved.
- On intransitive verbs, S is marked by set I. On nouns, set II is used as possessive markers.
- This is valid both in independent and dependent constructions.

We can now focus our discussion on the different hierarchies operating in transitive constructions and provide an overview of the Emérillon hierarchical system.

2. Person hierarchy and semantic hierarchy in Emérillon

While the person hierarchy which is attested in most Tupi-Guarani languages is usually presented as 1>2>3, the Emérillon data points only at a 1/2>3 hierarchy in a clear fashion. Note that both in Emérillon and in Proto-Tupi-Guarani the marking is very clear and systematic when only one speech act participant is involved, but works differently in the "local" forms, when both first and second person are involved. This difference of treatment is frequent in languages involving a person hierarchy. A similar example is given by Gildea (1994) in the inverse system of Caribe. DeLancey (no date) proposes the deictic nature of inverse and hierarchical systems as an explanation to the fact that these systems give a special status to speech act participants.

We will now look at the Emérillon "local" forms in more detail (first when A2 → P1, then when A1 → P2), with constant reference to the corresponding forms in other languages of the Tupi-Guarani family.

2.1 Local form: A2 → P1

Table 2 presents the four instances where a second person (singular or plural) acts on a first person (singular or plural).

A	P	
2 nd sg	1 st sg	<i>ele-nupã ele-ŋ</i> 2sg/I-hit 2sg/I-ŋ You (sg) hit me.
2 nd sg	1 st excl	<i>ele-nupã olone-kom</i> 2sg/I-hit 1excl/II-pl You (sg) hit us.
2 nd pl	1 st sg	<i>pe-nupã pe-ŋ</i> 2pl/I-hit 2pl/I-ŋ You (pl) hit me.
2 nd pl	1 st excl	<i>pe-nupã olone-kom</i> 2pl/I-hit 1excl/II-pl You (pl) hit us.

Table 2: A2 → P1 examples in Emérillon.

We already said that when A2 → P1, A is marked by a prefix of set II, and another person marker follows the verb. Since A is marked by the prefix, we would expect this last marker to refer to the patient, but this is not always the case.

- *Olonekom* refers explicitly to the patient first person exclusive in that *olone* is the set II first person exclusive marker and *kom* is the plural marker of P so this case is clear.

- Neither *eleŋ* nor *peŋ*, on the other hand, refer directly to the patient. However, their presence is necessary to distinguish A2 → P1 from A2 → P3, as shown in the following example (13a,b)

(13a) *ele-nupã eleŋ*
2sg/I-hit ?
You hit me.

(13b) *ele-nupã*
2sg/I-hit
You hit him.

In fact, it could be that *ele* and *pe* (without the final *ŋ*) are the set I person markers for second singular and plural shown in Table 1 earlier, but that would mean that they mark only the A. It would remain to explain how a double marking of a second person A could mean A2 → P1. The only explanation we can give for this oddity is that it could be a residue of the 1>2>3 hierarchy of Proto-Tupi-Guarani, more precisely of the independent pronoun marking A when A2 → P1. Indeed, in Proto-Tupi-Guarani, when A2 → P1, P

which is higher on the hierarchy by virtue of being 1 is marked with set II. A is then expressed by an independent pronoun after the verb, reconstructed by Jensen (1998) as **jepe* in the singular and **pejepe* in the plural. The examples in (14) are from Tupinamba, a conservative Tupi-Guarani language.

(14a) **sjé r-epják jepé* (Tupinamba, Jensen 1998)
 1sg/II r-see 2sg/PRO
 You (sg) see me.

(14b) **sjé r-epják pejepé* (Tupinamba, Jensen 1998)
 1sg/II r-see 2pl/PRO
 You all see me.

In these examples, *jepe* and *pejepe* refer to A2, since P is marked with a set II prefix.

Our diachronic hypothesis would be confirmed if I could explain precisely the origin of these markers, in particular the *ɲ*. It is possible that this nasal is bound to these second person markers in a more ancient way. Data from Sateré-Mawé⁶ (Franceschini, to appear) show the existence of *en* for a second person singular A, when P is first person. The form *ejpe* (second person plural A) does not contain any nasal.

2.2. Local form: A1 → P2

When A1 → P2, some special prefixed forms are used in Emérillon: *olo-*, *apolo-*:

(15a) *olo-nupã* (Emérillon)
 ?-hit
 I/we hit you.

(15b) *apolo-nupã*
 ?-hit
 I/we hit you all.

The first remark to be made is that *olo-* is homonymous with the set I marker for first person exclusive. (15a) may thus have another meaning:

(16) *olo-nupã* (Emérillon)
 A1-hit
 We hit (it/she/he/them).

The second remark is that Jensen 1998 proposes the following reconstructions for Proto-Tupi-Guarani: **oro-*: A1/P2sg, **opo-*: A1/P2pl. These forms are traditionally referred to as portmanteau, on behalf of the fact they refer to both A and P in a unique morpheme, analysis that I will reconsider. An interesting fact about these forms is that they display a great variety within the Tupi-Guarani family. And to the point here is the fact that, when Montserrat and Soares (1983) give five groups of Tupi-Guarani languages

for which the person hierarchy functions only partially, it is precisely in the A1 → P2 case that all five groups break the hierarchy. We will now focus on the analysis of these two forms in Emérillon.

The *olo* Emérillon morpheme is quite clearly a reflex of Proto-Tupi-Guarani **oro*. Its analysis as a "porte-manteau" is based on the fact that it reveals at the same time the persons of A and P, a fact that verified also for every marker we have seen until now. An analysis of *olo-* as an A or P marker seems therefore more relevant. Its homonymy with the set I marker for first person exclusive makes us lean toward its analysis as an A marker. Only context can solve the ambiguity between (15a) and (16). If we follow this hypothesis, we have to explain why in (15a), *olo-* can refer to a plural A as well as to a singular A. A possible solution is to be found in facts of politeness. It is a general fact about communication that the situation A1 → P2 creates a confrontation between the speaker and his/her addressee who is in a lower position by virtue of being the addressee and a patient. Languages often use devices like pluralization or substitution of a person for another to soften this confrontation (Kerbrat-Orecchioni, 1990, following Brown & Levinson, 1987). For example, in French, a unique addressee may be referred to with a second person plural pronoun "vous". A possible analysis of the Emérillon *olo-* is therefore that its plural form is a means of weakening the first person agent, making the relation less "threatening" for the addressee-P.

As far as the other form is concerned, that of *apolo*, which is not so clearly a reflex of **opo*, it could be segmented as follows: *a-* 1st person singular set I, referring to an A, and *polo* which is not part of the traditional person markers paradigms. However, *polo* is described by Jensen (1998) as an incorporated generic human object marker in Proto-Tupi-Guarani. In fact, here is an example from Emérillon, in which the object *polo* means *people*:

- (19) zo-tuna?ijɲ wane polo-ilu-o
 ind/II-heart good people-bring?-PROG
 It is the heart that makes people last/ survive.

The use of *polo* in the A1 → P2pl case could also be explained in terms of the confrontation between the speech act participants as was the case for *olo*. In this case, the use of an indeterminate marker for a second person would create a distance that softens the confrontation. In fact, the same hypothesis has been recently developed by Cabral (2001) for a few Tupi-Guarani languages.⁷

2.3. Conclusion on the hierarchisation

Describing the "local" forms (involving both first and second person) for Emérillon is problematic. The Emérillon A2 → P1 forms are different from those of the other languages of the family and we have followed here a new analysis for A1 → P2 forms.

The facts of Emérillon lead us to reconsider the 1>2>3 Tupi-Guarani hierarchy, at least as far as Emérillon is concerned. Once again, it is typologically common that a person hierarchy is clearly distinct with respect to 1>3 and 2>3 but less so when both speech act participants are involved. This may help account for the fact that the Emérillon language succeeded in reorganizing the Proto-Tupi-Guarani hierarchy concerning specifically the "local" forms.

In summary, our analysis of the Emérillon "local forms" are as follows:

- When A2 → P1, A has the priority to be marked. The second marker refers to P either directly (*olonekom*), or indirectly, when it is a residue of a Proto-Tupi-Guarani A marker (*elep*, *pep*).
- For A1 → P2sg, *olo-* is analyzed as an A marker.
- For A1 → P2pl, A is marked by set I (*a-*), whereas P is marked with a generic object (*polo*).

Our conclusion is that there is no obvious person hierarchy between the two speech act participants. However, in every case, A has priority over P: A fills the pre-verbal position specific to person marker, whereas P is put after the verb, incorporated between the set I person prefix and the verb, or simply implied from the context. Therefore, we propose that the relevant hierarchy for all these local forms is a semantic hierarchy A>P, hierarchy we already used to explain A3 → P3 (Cf.1.2.b).

There are now two different hierarchies: a 1/2>3 person hierarchy, and an A>P semantic hierarchy. Which hierarchy should be used first? We will follow here Couchili, Maurel & Queixalós's proposal: the "hierarchy of hierarchies" is person > semantic roles. For these authors, the person hierarchy is 2>1>3, and therefore the semantic hierarchy is used only when both participants are third persons. But since we described the person hierarchy as 1/2>3, we extend this analysis to the local configuration: it is also outside the scope of the person hierarchy, and must then obey the semantic hierarchy.

Since the existence of such person hierarchy in Tupi-Guarani languages has led to the interpretation of these agreement systems as "inverse" systems, we now turn to this interesting question.

3. Hierarchy of person in Tupi-Guarani and Emérillon: an inverse system?

It was Doris Payne (1994) who first formulated the hypothesis that Tupi-Guarani languages could be described as having an inverse system. On the basis of Givón's definition of the inverse, she explains that in these languages, there is an "inherent topicality hierarchy" 1>2>3. If A1 → P2/3 or A2 → P3, the action flows in the natural direction (A is more topical than P), and set I is used. This is considered to be direct. If the action flows the other way around (when P is more topical than A), then set II is used instead, and that is considered to be inverse. Further more, a canonical inverse language is a language that expresses the inverse direction with morphosyntactic devices in a transitive construction.

Then, the author applies this definition to the Tupi-Guarani system. Direct situations are marked with set I for A, inverse situations with set II for P. She proposes that the *r-* prefix be analysed as an inverse marker in Tupi-Guarani languages. Since this prefix is not present any more in independent Emérillon verbs, this discussion is irrelevant in a synchronic analysis of this language. If we still wanted to describe the Emérillon system as an "inverse" system, it would not be a canonical inverse in relation to the prototype of inverse (the Algonquian languages) since no morpheme can be interpreted as an inverse marker and moreover, the system is limited to situations in which one but only one speech act participant is involved.

To make the "inverse" analysis possible for Emérillon, one should be willing to consider only the sheer distribution of person markers into two sets as an inverse system (keeping in mind the fact that some markers do not belong clearly to one or the other set). This assumption would be possible with a purely functional definition of inverse like those of Givon (1994) or Klaiman (1991) which find inverse whenever P is more topical than A, but A is still topical. In Thomas Payne (1997), one subtype of inverse is called "special verb agreement for inverse situations", and is illustrated by examples of Wayampi, a language very close to Emérillon.

Therefore, although one can easily identify an inverse function in the Emérillon language, it is for us an insufficient argument to speak in terms of an inverse system. First, it is not a system, since two cases out of three are not concerned (when both participants are third persons, and when both participants are speech act participants), and second direction is not expressed by a specific morpheme. The analysis of the person marking system as an inverse system thus does not seem necessary to understand the data: positing person and semantic hierarchies appears to suffice to explain the cross-referencing system. We will now briefly conclude on the type of alignment system the Emérillon exemplifies.

Conclusion: a hierarchical agreement system

We will now conclude on the issue of the type of alignment system the Emerillon exemplifies.

Two types of agreement system were quickly dismissed: the absolutive, since it is not used any more in dependent clauses in Emérillon (1.3), and the active/stative system since the analysis of "descriptive" words as nouns rather than verbs make it impossible to have a split-intransitivity (all Emérillon intransitive verbs are marked with set I, Cf.1.1). However, Couchili, Maurel & Queixalós (no date) suggest that the Emérillon system can be called active/stative if we accept to speak in terms of predicates rather than verbs, since some single-place predicates are marked with set I -intransitive verbal predicates- and others with set II -nominal predicates (1.1).

The Emérillon agreement system also shares some characteristics with the nominative/accusative system, since S and A are both marked with set I that could be interpreted as a nominative paradigm, and P only is marked with set II, the accusative paradigm. But an important point is that only one argument is actually marked on transitive verbs, and this according to the person and semantic hierarchies.

Furthermore, we discussed in 3 the possibility to analyse the Emérillon system as an inverse system, but preferred to speak in terms of hierarchies only.

Thus, hierarchy is the primary organizing pattern of the Emérillon agreement system, which we will therefore call a hierarchical agreement system. In this, we follow Nichols (1992), who recognizes the hierarchical system as an "alignment" system, along with neutral, accusative, ergative, three-way and stative-active, where "inverse" is just a sub-category of hierarchical systems.

Notes

1- I would like to thank Francisco Queixalós, Andrej Kibrik and Scott DeLancey for their helpful remarks.

2- A: subject of the transitive clause; S: subject of the intransitive clause; P: object of the transitive clause. A1 → P2 stands for "the first person agent acts on a second person patient".

3- The term "person hierarchy" that I use corresponds roughly to other designations like "animacy, agency, referential or inherent topicality hierarchy". No semantic or pragmatic definition is necessary, given that in Emérillon, the hierarchy clearly orders grammatical persons.

4- 1sg/I means *a-* is a prefix of set I referencing 1st person singular.

5- This analysis was originally proposed in Couchili, Maurel & Queixalós, n.d.

6- Sateré-Mawé belongs to the Tupi family, but not to the Tupi-Guarani branch.

7- Cabral proposes that the proto-form for A1sg → P2pl be **a-poro* (and not **apo*), **a-poro* being reduced phonologically in some languages (which explains Jensen's reconstruction **apo-*).

Abbreviations

I, II, III	Set I, II or III
pl	plural
sg	singular
excl	exclusive
incl	inclusive
ind	indeterminate
PRO	pronoun
FUT	future (immediate future)
SER	serialization marker
OBTOP	oblique-topicalization marker

COMPL	completive suffix
REL	relativization marker
PROG	progressive

References

- Brown P. & Levinson S. 1987. *Politeness*. Cambridge: Cambridge University Press.
- Cabral A.S. 2001. "O desenvolvimento da marca de objeto de segunda pessoa plural em Tupí-Guaraní". *Estudos sobre Línguas Indígenas*, ed. by A.S. Cabral & A.D. Rodrigues. Belem: Universidade Federal do Pará.
- Couchili T., Maurel D. & Queixalós F. no date. *Classes de lexèmes en émérillon*, to appear in *Amerindia*.
- DeLancey S. no date. *Deixis, Topicality, and the Inverse*. ms. <http://www.uoregon.edu/~delancey/papers/inverse.html>
- Franceschini D. to appear. *A voz inversa em Sateré-Mawé (Tupi)*. *Atas do I Encontro Internacional do GTLI*.
- Gildea S. 1994. *Semantic and pragmatic inverse : 'inverse alignment' and 'inverse voice' in Carib of Surinam*. *Voice and inversion*, ed. by T. Givón. Amsterdam/Philadelphia: John Benjamin's.
- Givón T. 1994. *The pragmatics of de-transitive voice: functional and typological aspects of inversion*. *Voice and inversion*, ed. by T. Givón. Amsterdam/Philadelphia: John Benjamin's.
- Harrison C.H. 1984. *Verb proeminence, verb initialness, ergativity and typological disharmony in Guajajara*. *Amazonian languages*, ed. by D.C. Derbyshire & G.K. Pullum, vol. I. Berlin: Mouton de Gruyter.
- Jensen C. 1990. *Cross-referencing changes in some Tupí-Guaraní languages*. *Amazonian linguistics, Studies in Lowland South American Languages*, ed. by D. Payne. Austin: University of Texas Press.
- Jensen C. 1998. *Comparative Tupí-Guaraní Morpho-syntax*. *Handbook of Amazonian languages*, ed. by D.C. Derbyshire & G.K. Pullum, vol. IV. Berlin: Mouton de Gruyter.
- Kerbrat-Orecchioni C. 1992. *Les interactions verbales, tome II*. Paris: Colin.
- Klaiman M. 1991. *Grammatical voice*. (Cambridge Studies in Linguistics 59) Cambridge: Cambridge University Press.
- Leite Y. 1990. *Para uma tipologia ativa do Tapirapé, os clíticos referenciais de pessoa*. *Cadernos de estudos linguísticos* 18: 37-56.
- Maurel D. 1998. *Eléments de grammaire émérillon*. (Chantiers Amerindia 1.23) Paris.
- Montserrat R. & Soares M.F. 1983. *Hierarquia referencial em línguas Tupi*. *Ensaio de lingüística* 9: 164-187.
- Nichols J. 1992. *Linguistic Diversity in Space and Time*. Chicago/London: University of Chicago Press.
- Payne D. 1994. *The Tupí-Guaraní inverse*. *Voice: form and function*, ed. by B. Fox & P.J. Hopper, 313-340 (*Typological Studies in Language* 27). Amsterdam/Philadelphia: John Benjamin's.

- Payne T.E. 1997. Describing morpho-syntax: a guide for field linguists. Cambridge: Cambridge University Press.
- Queixalós F. 2001. Le suffixe référentiel en émerillon. *Des noms et des verbes en tupi-guarani: état de la question*, ed. by F. Queixalós. (LINCOM Studies in Native American Languages 37) München: Lincom Europa.
- Rodrigues A.D. 1984-1985. Relações internas na família lingüística Tupí-Guaraní. *Revista de Antropologia* 27/28: 33-53.
- Rodrigues A.D. 1990. You and I = neither you nor I: the personal system of Tupinambá. *Amazonian linguistics, Studies in Lowland South American Languages*, ed. by D. Payne. Austin: University of Texas Press.
- Rodrigues A.D. 1996. Argumento e predicado em tupinamba. ABRALIN 19.
- Rose F. 2000. *Eléments de phonétique, phonologie et morphophonologie de l'émerillon (teko)*. Mémoire de DEA, Université Lumière Lyon II.
- Seki L. 1990. Kamaiurá (Tupí-Guaraní) as an active-stative language. *Amazonian linguistics, Studies in Lowland South American Languages*, ed. by D. Payne. Austin: University of Texas Press.

Nominal Classification in Miraña

Frank Seifart

Max Planck Institute for Psycholinguistics

0. Introduction

This paper presents some data and analysis from an ongoing research on the system of nominal classification of Miraña.¹ I want to draw attention on two closely related issues of theoretical import raised by the Miraña data: the status of this system in a general typology of systems of nominal classification and the notion of agreement in noun class.

Miraña is a dialect of Bora (cf. Thiesen 1996) and belongs to the Witotoan language group (Aschmann 1993). It is spoken along the Caquetá river in the Colombian Amazon not far from the Peruvian and the Brazilian borders. It has now less than 100 speakers left, none of them children, for a population of about 400. It is being replaced by Spanish. Miraña has a complicated tonal system. It is a fairly polysynthetic language with a strong preference for suffixing. The basic word order is SOV, but word order is relatively free.

The system of nominal classification in Miraña is made up of at least 62 class markers which are phonologically bound forms used on a variety of targets to mark agreement within noun phrases, for cross-referencing and anaphora. Their function in combination with nouns is mostly derivational. Example (1) shows four instances of one class marker in various functions.

- (1) *ó-dì* *í^hkà:-bà* *tsà:-bà* *mu^hwù:-bà* *ʔúβì:-bà*
1S-POS be-SCM.cont one-SCM.cont big-SCM.cont basket-SCM.cont
"I have one big basket" (lit. What is to me, is one big basket)

1. Inventory of class markers: general and specific class markers and repeaters

There are two clearly identifiable subtypes of class markers, which are in paradigmatic relation within the same morphosyntactic environments, but differ in their concordial function and in their semantics: "general" class markers and "specific" class markers. A third set of morphemes may occur in some, but not all class marker specific contexts. These are the so called "repeaters".

1.1. General class markers

There are six general class markers, five for animates which combine semantics of number (singular, dual and plural) and sex (masculine and feminine), and one for inanimates which does not distinguish number (table 1). All animate class markers have allomorphs according to morphosyntactic context, including some suppletive forms, and some are morphologically complex.

forms of general class markers	meaning
-:bɛ, -di, - ^h pi ~ -:pi	masculine singular
-dʒɛ, -pi-dʒɛ	feminine singular
-muu-tʃi, -tɛ-tʃi	masculine dual
-muu-pi, -tɛ-pi	feminine dual
-mɛ, -tɛ, -muu	animate plural
-nɛ	inanimate

Table 1: The forms of general class markers

1.2. Specific class markers

The bulk of class markers are the specific class markers (at least 56). They have the same form in all constructions in which they are found (with only one exception) and do not mark number, which is marked by a separate suffix. The set of specific class markers is heterogeneous in terms of forms and semantics, ranging from monosyllabic forms with relatively general semantics to polysyllabic ones with very specific semantics. The semantics of specific class markers is based mainly on shape distinctions.

The set of specific class markers is given in two sections: Table 2 contains the forms that do not have a recognizable nominal origin while table 3 contains the ones with a nominal origin within the language.²

N°	specific class marker	approximate meaning	examples
1	-ʔaʲ	oval	u ^h kàʔàʲ "beard"
2	-ʔɛ	tree, bush, plant	kó:huʔɛ "avocado tree"
3	-ʔi	bunch, river	bɛ:ʔi "bunch of a coconuts"
4	-ʔo	oblong	à:ʔò "maraca fruit"
5	on noun: -ba ~ -ʔba in all other contexts: -ʔba	a variety of meanings, e.g. thick liquids; fruits; seasons; oblong objects; musical instruments; etc.	nɛ:g ^w à:ba "stone" màʔníbà "natural tar"
6	-dʒi	a bunch of fibers	ádʒúúwùdʒi "eyelashes"
7	-g ^w a	plank-like shaped	bòʔdóg ^w à "paddle"
8	-hi	flat, round	màri:múhì "taricaya turtle"
9	-i	stick	kò:i "wooden stick"
10	-ko	shaft	kà:núwò "mortar"
11	- ^h kɛ	climbing vines	g ^w áʔdà ^h kɛ "guaya vine"
12	- ^h paʲ	liquid, broth	tɛ ^h paʲ "broth"
13	- ^h tò	pointed	níhíʔɛ ^h tò "cumare palm thorn"
14	-rò	bottle, pipe	àdòrò "drinking bottle"

15	-u	round, three dimensional	í:ʔúúú "egg"
16	-ʔa ^h ku	trunk	úúúʔa ^h ku "tree trunk"
17	-ʔa ^h tsi	clearing	dʒáʔa ^h tsi "patio"
18	-ʔa:mi	leaf	g ^w áhákúʔá:mì "book"
19	-ʔbaba ^j	bag	g ^w ará ^h kòʔbábá ^j "bag to sieve coca"
20	-ʔehu	hole	tú:ʔéhu "nostril"
21	-ʔi:g ^w a	open space on a riverbank	pá ^h tèʔí:g ^w a "port on a river"
22	-ʔi:ba	small palm tree	tó:kèʔí:bà "small palm tree, sp."
23	-ʔo:ba	short gable	háʔo:ba "short gable of a house"
24	-βi:u	chunk	kóʔbàβi:ù "chunk of wood"
25	-dʒi:ʔo	bud	mé:méʔédʒi:ʔò "bud of chontaduro palm"
26	-dʒi:hu	powder	bádʒi:hú "ash, powder"
27	-i:ʔo	little stick	ká:túúúí:ʔò "pencil"
28	-i:hi	platform	g ^w á:hì "raft, floor"
29	-g ^w a ^j :ʔe	crumbs	í:búʔg ^w a ^j :ʔè "leftover of coca"
30	-g ^w a ^j :ʔò	bundle of sticks	í:búʔg ^w a ^j :ʔò "bundle of coca sticks"
31	-hu:ʔo	palm leaf	níhìʔéhu:ʔò "cumare palm leaf"
32	-kaha	swamp, creek	ínèkáhá "flooded canangucho plantation"
33	-miha	wooden stick for illumination	kóg ^w áʔámíhá "copaiba stick for illumination"
34	-mi:ʔe	skin, soft shell	báhímí:ʔè "ray fish skin"
35	-mi:ʔo	hard shell	kú:múhímí:ʔò "turtle shell"
36	-ni:ʔo	female which has offspring	ókáhini:ʔò "female tapir which has offspring"
37	- ^h pa ^j ko	liquid	a ^j βè ^h pa ^j kò "liquor"
38	- ^h tsu:ʔo	pack	úúú ^h tsú:ʔò "pack of salt"
39	-pa ^j ʒi	planted field	úúú ^h épa ^j ʒì "a planted field"
40	-pa ^h tsi	roll, ring	móʔòpá ^h tsì "a roll of liana"
41	-pa:hi	cave, hole	úúú ^h épa:hì "a hole in a tree"
42	-to ^h ko	corner	háto ^h kò "a corner of a house"
43	-tʂo:ʔo	medium-sized palm tree	ínéʔétʂó:ʔò "medium-sized canangucho palm tree"
44	-roʔdʒo	completely twisted	dú:ruúróʔdʒò "a completely twisted candle"
45	-uʔo	basin	ká:núúúʔò "basin for pounding fruit"
46	-u:ʔa ^j	grains	nég ^w a ^j ú:ʔá ^j "sand"
47	-me ^h kei	thin part	úúú ^h éme ^h kéì "thin part of a tree"
48	-tʂa:rag ^w a	fibers sticking out	ní:g ^w áútsá:rag ^w a "head with hair sticking out"

Table 2: Specific class markers that do not have a recognizable origin

N°	specific class marker	meaning	example	nominal origin
49	-?bau	round protuberance	té?bàù "hill"	í-?bàù (POS.3S-belly) "his/her belly"
50	-ha	cover, shelter	ká:méèhà "shirt"	ha "house"
51	-huu	tube, story, song	bá'nèhù "cigarette"	í-?hù (POS.3S-mouth) "his/her mouth, story, language"
52	-mi	means of transport	kú:huúg"ámì "motorboat"	mì:nè "canoe"
53	-mo	big river	ókáhimò "Caquetá river" (river of the tapir)	mò:àì "big river"
54	-ta	metal	mé?ó ^h tsig"ámì:?óné imít'ótà" nail clipper"	rá:tà "tin can"
55	-tu ^h ke	stem of a fruit or a leaf	mu ^h tsi:tsíbatu ^h ke "stem of a pear apple fruit"	tu ^h kénù "beginning"
56	-.ba	deep pool, container	àdó:bà "salty well" g ^w á':bà "hammock"	bánè:bà "pool"

Table 3: Specific class markers with a recognizable origin

1.3. Repeaters

Repeaters are either reduced forms of a noun ("partial" repeaters; table 4) or homophonous with a noun ("full" repeaters; table 5). They occur in class marker specific contexts, but they are associated only with the one noun from which they derive. Repeaters apparently have the same meaning as the noun from which they derive, as opposed to specific class markers with a nominal origin, which have a more general meaning than the corresponding noun. Example (2) illustrates the use of a "full" repeater.

- (2) ó-dì í^hkà-báhù tsá-báhù báhù
 1S-POS be-RP.forest one-RP.forest forest
 "I have one (stretch of) forest."

N°	repeater	corresponding noun	meaning
1	-bo	bó:à	anaconda
2	-i:	í:bû:	coca
3	-i?d3o	d3íri?d3ò	pot
4	-ma	námè	feces
5	-pe	d3ì?pè	rat, sp.
6	-to	tò?hù	earthworm
7	-tʃi	tʃi?tʃi	tucupí sauce
8	-ra	d3ò:rà	parrot, sp.

Table 4: "Partial" repeaters

N°	repeater	corresponding noun	meaning
9	-ba ^h ke	bá ^h kè	root
10	-g ^w á ^h kà	g ^w á ^h kà	branch
11	-ba ^h ku	bá ^h ku	bone
12	- ^h ku ^{ba}	^h ku ^{bà}	leg
13	- ^h tu ^{ʔa} i	tu ^{ʔa} i	foot
14	-ʔe:ba	í-ʔé:bà "his/her waist"	waist
15	-ʔo ^h tsi	ʔó ^h tsì	hand
16	-ʔumi	ʔúmì	face
17	-bahu	báhu	forest, plantation
18	-dehu ^{ʔri} u	dèhu ^{ʔri} u	tree without roots
19	-g ^w airo:hi	g ^w áiro:hì	very thin
20	-g ^w a ^h ko	g ^w a ^h kò	hook
21	-i:hu	ì:hù	anteater
22	-ko:hi	kò:hì	day
23	-ko:mi	kó:mì	community
24	-mí ^h ko	mí ^h kò	corral
25	-mu ^h ko	mu ^h kò	place for bathing
26	-o:ʔi	ò:ʔí:-bè (jaguar-GCM.masc.sg) "jaguar, dog"	jaguar, dog
27	-pe ^h ko	pè ^h kò	night
28	- ^h ku:βe	^h ku:βè	afternoon
29	-ʔí ^h ko	ʔí ^h kò	nest
30	-ro ^ʔ pe ^g a	rò ^ʔ pég ^w à	twisted

Table 5: "Full" repeaters

2. Morphosyntactic contexts of class markers

Class markers can be found on multiple sites in a clause: on nouns whether root or derived, on modifiers of the noun, including relative clauses, and on predicates. Both subtypes of class markers may occur in any of these contexts.

2.1. On nouns

Not all nouns carry a class marker at all times. Some seem to function as concept nouns when bare and to acquire individuality with the affixation of a class marker. The use of the class marker may have definite derivational value, as shown with the various nouns naming different parts of the banana plant, or rather, aspects of the banana concept (example 3).

- (3) a. *úhì*
banana
"banana"
- b. *úhì-ʔò*
banana-SCM.oblong
"a banana (fruit)"

- | | | | |
|----|---|----|--|
| c. | <i>úhì-kò</i>
banana-SCM.shaft
"a banana plant" | d. | <i>úhì-g"à</i>
banana-SCM.plank
"one half of a banana" |
| e. | <i>úhì-ʔì</i>
banana-SCM.bunch
"a bunch of bananas" | f. | <i>úhì-hì</i>
banana-SCM.flat.round
"a banana seed " |
| g. | <i>úhì-ʔbábàj</i>
banana-SCM.bag
"a bag of banana" | | |

Successive multiple affixation of different class markers on the same noun is also possible (example 4).

- (4)
- | | |
|----|--|
| a. | <i>úhì-kó-ʔá:mì</i>
banana-SCM.shaft-SCM.leaf
"a leaf of a banana plant" |
| b. | <i>úhì-ʔó-βí:ù</i>
banana-SCM.oblong-SCM.chunk
"a chunk of a banana" |
| c. | <i>úhì-dʒí:huì-rò</i>
banana-SCM.powder-SCM.bottle
"a bottle of pulverized dried banana" |

The majority of nouns do not exhibit the alternation of root with and without class markers and must always be suffixed with a class marker (example 5).

- (5)
- | | | | |
|----|--|---|--------------------|
| a. | <i>úmè-ʔè</i>
tree-SCM.tree
"tree" | * | <i>úmè</i>
tree |
| b. | <i>í:ʔú-ù</i>
egg-SCM.round.3D
"egg" | * | <i>í:ʔú</i>
egg |

In the case of specific class markers, the number marking presupposes the presence of a class marker (example 6). The general animate class markers encode number as well as noun class (example 7).

- (6)
- | | |
|----|---|
| a. | <i>kò:mì</i>
milpeso.palm
"milpeso(s)" |
| b. | <i>kó:mì-ù</i>
milpeso.palm-SCM.round.3D
"(a) milpeso palm fruit" |

- c. *kó:mì-ú-kù*
milpeso.palm-SCM.round.3D-DL
"two milpeso palm fruits"
- d. *kó:mì-ú:nè*
milpeso.palm-SCM.round.3D-PL
"milpeso palm fruits (>2)"
- (7) a. *á'tónù:-bè*
brother/sister.in.law-GCM.masc.sg
"a brother-in-law"
- b. *á'tónù-dzè*
brother/sister.in.law-GCM.fem.sg
"a sister-in-law"
- c. *á'tónù-mútsì*
brother/sister.in.law-GCM.masc.dual
"two brothers-in-law"
- d. *á'tónù-mù*
brother/sister.in.law-GCM.masc.pl
"brothers-in-law (>2)"

Some of the combinations of nouns (including derived nouns) and class markers semantically resemble compounding processes (example 8).

- (8) a. *á'βè-^hpá'kò*
pain-SCM.liquid
"liquor"
- b. *kuú:huúg^wà-^hpá'kò*
fire-SCM.liquid
"gasoline"
- c. *g^wàhákuè-ʔá:mì*
know.NMZ-SCM.leaf
"book"
- d. *kà:nú-kò*
to.pound.NMZ-SCM.shaft
"mortar"
- e. *kà:nú-ì*
to.pound.NMZ-SCM.stick
"pestle"

2.2. Class markers on other nominal expressions

Class markers combine with the modifiers of the noun, such as numerals, demonstratives, possessives, and various determiners (example 9). Any of these modifiers may also be used headlessly as anaphoric device. It is noteworthy that, as a rule, any nominal expression in Miraña that has the potential to show agreement in noun class, must include a class marker.

(9) NUMERALS

- a. *tsà-ʔì* *úhì-ʔì*
 one-SCM.bunch banana-SCM.bunch
 "one bunch of banana"

DEMONSTRATIVES

- b. *í-gʷà*
 DEM.PROX-SCM.plank
 "this (plank, bench, etc.)"
- c. *è-gʷà*
 DEM.DIST-SCM.plank
 "that (plank, bench, etc.)"

OTHER NOMINAL EXPRESSIONS

- d. *táʔné-gʷà*
 my-SCM.plank
 "my (plank, bench, etc.)"
- e. *pà-ʔbà*
 complete-SCM.3D
 "this whole (fruit, drink, etc.)"
- f. *tsì-ʰpì*
 other-GCM.masc.sg
 "another (male)"
- g. *tsì-mì*
 other-SCM.transport
 "another (canoe, motorboat, etc.)"

The modifiers taking class markers include stative verbs used as adjectives (example 10) and relative clauses (example 11).

(10) STATIVE VERBS USED AS ADJECTIVES

- a. *múhù-hì* *ku:mu-hì*
 big-SCM.flat.round turtle-SCM.flat.round
 "big turtle"
- b. *ku:βè-gʷà* *bòʔdó-gʷà*
 dark-SCM.plank paddle-SCM.plank
 "dark paddle"

(11) RELATIVE CLAUSES

- a. *hà:* [ò *í:té*]-*hà* *ái:βé-ŕì*
house [1S.SUB see]-SCM.cover burn-PRED
"The house I saw burned down"
- b. *ó ì:té-ŕí í^hhà-g^wà* [ò:-kè *ù* *á^hkuù*]-*g^wà*
1S see-PRED bank-SCM.plank [1S-ACC 2S.SUB give]-SCM.plank
"I saw the bank that you gave to me"

2.3. Class markers on verbs

Finally, class markers are used for the indexing of the subject on predicates, in the case of intransitive verbs as well as transitive verbs (example 12).

- (12) a. *í:té-dzè*
look-GCM.fem.sg
"She looks"
- b. *ká:túβè-g^wà*
fall.down-SCM.plank
"The (plank, etc.) fell down"
- c. *g^wá?dáínù:-bé* *àhì*
cut-GCM.masc.sg palm,sp.
"He cuts the palm tree"
- d. *á^hkuù:-bé* *báínè-hú-βù* *g^wá^hkódzì-kè*
give-GCM.masc.sg tobacco-SCM.tube-ADL proper name-ACC
"He gives a cigarette to *g^wá^hkódzì* (proper name)"

3. Variable, anaphoric and 'absolute' use of class markers

In order to give a realistic picture of the functioning of nominal classification in Miraña, some aspects of the actual use of class markers in natural discourse will be mentioned in this section.

3.1. Variable use of class markers

In Miraña, it is possible to see a variation between types of class markers in many context except the noun. Syntactically, a general class marker as well as a specific one can be used to mark agreement, cross-reference, and anaphoric relation with a noun that belongs to a specific class (example 13).

- (13) a. *ó-dì* *í^hkà-kò* *tsà-kò* *pì^hkuú-kò*
1S-POS be-SCM.shaft one-SCM.shaft to.fish.NMZ-SCM.shaft
"I have one fishing rod"

- b. *ó-dì* *í^hkà-nè* *tsà-nè* *pì^hkuú-kò*
 1S-POS be-GCM.inanim one-GCM.inanim to.fish.NMZ-SCM.shaft
 "I have one fishing rod"
- c. *kuú:muù-hì* *á:kítè-hì*
 turtle-SCM.flat.round fall.down-SCM.flat.round
 "The turtle fell down"
- d. *kuú:muù-hì* *á:kítè:-bè*
 turtle-SCM.flat.round fall.down-GCM.masc.sg
 "The turtle fell down"

This variation is at least partially dependent on register and age of speaker. The forms showing agreement that include a specific class marker (examples 13a and c) are more likely to be used by older speakers and in formal registers that often involve magical practices. In everyday conversation, the use of specific class markers in positions other than the noun (such as in examples 13a and c) is actually extremely rare and general class markers are used instead (such as in examples 13b and d). However, constructions such as 13a and c are not only accepted by native speakers, but considered to be the "proper ones" or "the way the elders spoke".

3.2. Anaphoric use

Most of the modifiers in Miraña can also be used independently of a noun phrase as anaphoric devices. Also, predicates that include a class marker can occur without an over subject noun phrase. In these expressions the use of one type of class marker as opposed to the other may depend on the degree to which the speaker wishes to specify the participant at a given point in discourse: he might use the specific class marker, which is usually also an integral part of the noun which encodes the participant, in combination with a pronominal stem, on a relative clause or on verb, or a general class marker, which specifies the participant only for animacy and possibly also for number and sex.

This can be observed in the following example (14), which is taken from a text, where a Miraña speaker explains the making of a blowgun. The name of the blowgun is mentioned once at the beginning of the text (example 14a). It includes the specific class marker *-huu* (SCM.tube). All through the text, there is no need to specify the reference to the blowgun beyond "inanimate", because it can easily be established by the hearer, so the general inanimate class marker *-nè* is used in different kinds of constructions to establish reference to the blowgun (example 14b). Only at the very end of the text, the specific class marker *-huu* (SCM.tube) is used again, in combination with the pronoun *tè-* (example 14c).

- (14) a. *tódzì:-huù*
 blowgun-SCM.tube
 blowgun
- b. *tè-nè* *ími-nè*
 PN-GCM.inanim good-GCM.inanim
 "It (the blowgun) was good"

- c. *ábáhá-hpì-kè* *ó á^hkwù-:ʔì* *té-hwù-βwù*
 owner-GCM.masc.sg-ACC 1S give-PRED.FUT PN-SCM.tube-ADL
 "I will give it (the blowgun) to the owner."

3.3. 'Absolute' use

Miraña class markers can have an 'absolute' use in which they are neither syntactically nor anaphorically linked to any noun and where the class marker itself clearly conveys a meaning. Since Miraña class markers are bound forms, they have to combine with one of the many expressions that have a class marker slot, e.g. a pronominal stem. In the story of the blowgun, when it comes to making a hole in it, this notion is introduced by a construction using the semantically weak pronominal stem *té-* in combination with the class marker *-pa:hi* (SCM.hole) (example 15).

- (15) *té-pà:hì* *mé-kádzòìʔkwù-ʔì*
 PN-SCM.hole SAP-scrape-PRED
 "One scrapes the hole"

This 'absolute' use of class markers shows the importance of the class markers in the reference system of the language, not only to track down a participant but also to introduce a new one.

4. Issues of theoretical interest raised by the Miraña data

4.1. The status of the Miraña system within a general typology of systems of nominal classification

Within recent approaches to a general typology of systems of nominal classification, noun classes and classifiers (with their various subtypes according to morphosyntactic context, e.g. noun classifiers, numeral classifiers, etc.) have been recognized as the major types (Grinevald 2000; Aikhenvald 2000). These major types differ mainly in their degree of grammaticalization: noun classes are grammaticalized agreement systems, while classifiers are of a lexico-grammatical nature. Miraña class markers share some important properties of both of these types of systems.

The specific class markers are much like classifiers in that they constitute a large, somewhat open system with transparent semantics, which is sensitive to discourse conditions. Since no "primary" morphosyntactic context can be determined for Miraña class markers, they constitute a "multiple classifier system" in the terminology of Aikhenvald (2000). It is however unusual for classifiers to be bound forms and –more importantly– to take part in agreement.

It is the major defining criterion of noun class systems to be realized in agreement systems. So the Miraña system can also be characterized as a noun class system. It is, however, a system at a much earlier stage of grammaticalization than prototypical noun class systems, which are found for instance in Bantu languages (Grinevald/Creissels/Seifart 2001). The less grammaticalized nature of the system is apparent in the size of the system, the lexical nature of many of the class markers and the variability of their use according to speakers and discourse conditions.

The presence of large systems of nominal classification with transparent semantics which fulfill a variety of functions is actually an areal phenomenon of the North-West Amazon. Such systems have been described for other Witotoan languages (Petersen de Piñeros 1994), Eastern Tucanoan languages (e.g. Barnes 1990; Gómez 1982; for an overview cf. Gómez-Imbert 1996), Arawakan languages (Aikhenvald 1994) and Yagua (Payne 1986). (Cf. Payne 1987 and Derbyshire and Payne 1990 for an overview of systems of nominal classification in Amazonian languages.) So the issue of their typological status does not only concern Miraña, but the North-West Amazon area more generally.

4.2. Agreement in noun class

Virtually all coreferential expressions in Miraña must include either a specific or a general class marker. If a head noun is present, the choice of a class marker of either subtype is clearly governed by this noun's properties: either the specific class it belongs to or its animacy and possibly its number and sex. The use of either subtype of class markers clearly constitute agreement in that sense.

A complicating factor for the analysis of class markers as agreement markers is the 'absolute use' of specific class markers and the use of specific class markers in natural discourse in general, where usually only general class markers are being used to mark agreement in a more narrow sense, while specific class markers are used predominantly for derivational purposes on nouns and sometimes to establish anaphoric reference. The use of specific class markers in contexts other than the noun does not seem to be governed by the morphosyntactic constraint of an agreement relation. In fact, the relation between expressions the include specific class markers with antecedents that belong to that class may be better analyzed as a coreferentiality relation that is established in a discourse situation by virtue of the lexical meaning of that class marker and not as a syntactic agreement phenomenon (cf. Barlow 1999). In the great majority of discourse situations, however, there is no need to invoke the lexical meaning of a specific class marker to indicate coreferentiality. Since we do not usually talk about various differently shaped objects at the same time, general class markers are most often used in these contexts.³

The phenomenon of split agreement in a broad variety of morphosyntactic contexts where noun class is marked is typologically the most outstanding feature of nominal classification in Miraña. A number of languages have been reported to have more than one noun class system and thus more than one type of agreement. However, the different types of agreement are usually realized through different sets of morphemes and on different targets. Thus, these systems are independent from each other (cf. Aikhenvald 2000: 70ff. for examples). On the other hand, Payne (1986: 122) reports that in Yagua, a neutral class marker may replace a specific class marker in some contexts. In Miraña, however, general class markers constitute a whole system which may be used alternatively to another system, that of specific class markers and repeaters.

Given the large number of specific classes and the fact that class markers are obligatory in all coreferential nominal expressions and in one type of predicates, the existence of at least one alternative class marker, a general one, such as in Yagua, or an alternative, general system, such as in Miraña, seems to be necessary. Possibly, the case of Miraña lends support to the establishment of a general constraint on agreement systems, namely the impossibility of being large and obligatory in a variety of morphosyntactic contexts at the same time.

Notes

¹ The data comes from my own fieldwork in 1999 and 2000 in Colombia which was supported by a grant from the Deutsche Akademische Austauschdienst (DAAD). For a more detailed account of this system cf. Seifart (Forthc.). I wish to thank the Miraña communities for their collaboration, Colette Grinevald, Jon Landaburu and the members of the Language and Cognition Group of the Max Planck Institute for Psycholinguistics for discussion and comments.

The following abbreviations are used: 1 = first person; 2 = second person; 3 = third person; 3D = three dimensional; ADL = adlative; ACC = accusative; ADL = adlative; cont = container; COP = copula; DEM = demonstrative; DIST = distal; DL = dual; FEM = feminine; FUT = future tense; GCM = general class marker, inanim = inanimate; MASC = masculine; NMZ = nominalization; PAS = past tense; PL = plural; POS = possessor; PRED = predicator; PN = pronoun; PROX = proximal; RP = repeater; S = singular; SAP = speech act participant; SCM = specific class marker; SUB = subordinate clause; TEMP = temporal; VBZ = verbalization.

² The inventory of specific class markers given here corresponds to the set of forms that can occur in all of the class marker specific contexts according to the grammaticality judgments of a number of native speakers. There is, however, a big variation between speakers in the acceptability of some rare constructions including specific class markers, e.g. in combination with verbs (cf. §2.3), which makes it difficult to define the inventory precisely.

³ In a recently run experiment, where speakers were asked to assemble a variety of differently shaped objects, specific class markers were used in a variety of different constructions to distinguish these objects. This data is still being processed.

References

- Aikhenvald, Alexandra Y. 1994. Classifiers in Tariana. *Anthropological Linguistics* 36 (4): 407-465.
- 2000. *Classifiers. A typology of noun categorization devices*. Oxford: Oxford University Press.
- Aschmann, Richard P. 1993. *Proto Witotoan*. (Publications in Linguistics 114.) Dallas: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Barlow, Micheal. 1999. Agreement as a Discourse Phenomenon. *Folia Linguistica* 15/1-2 XXXIII (2):187-210.
- Barnes, Janet 1990. Classifiers in Tuyuca. *Amazonian linguistics. Studies in Lowland South American Languages*, ed. by Doris L. Payne, 273-92. Austin: Texas University Press.
- Derbyshire, Desmond C., and Doris L. Payne. 1990. Noun classification systems of Amazonian languages. *Amazonian linguistics. Studies in Lowland South American Languages*, ed. by Doris L. Payne, 243-72. Austin: Texas University Press.
- Gómez, Elsa. 1982. De la forme et du sens dans la classification nominale en Tatuyo (langue Tucano orientale d'Amazonie colombienne). (TDM 19.) Paris: Université Sorbonne / Ed. ORSTOM.
- Gómez-Imbert, Elsa. 1996. When animals become "rounded" and "feminine": Conceptual categories and linguistic classification in a multilingual setting. *Rethinking linguistic relativity*, ed. by John J. Gumperz and Stephen C. Levinson, 438-69. Cambridge: Cambridge University Press.
- Grinevald, Colette. 2000. A morphosyntactic typology of classifiers. *Systems of nominal classification*, ed. by Gunter Senft, 50-92. Cambridge: Cambridge University Press.

- Grinevald, Colette, Denis Creissels and Frank Seifart. 2001. Noun classes in African and Amazonian Languages. Toward a Comparison. Paper presented at the International Symposium "Typology of African languages", Sankt Augustin, May 2001.
- Payne, Doris L. 1986. Noun Classification in Yagua. Noun classes and categorization, ed. by Colette Craig, 113-32. Amsterdam: John Benjamins.
- 1987. Noun classification in the Western Amazon. *Language Science* 9 (1): 22-44.
- Petersen de Piñeros, Gabriele 1994. La lengua Uitota en la obra de Konrad Theodor Preuss. Bogotá: editorial universidad nacional.
- Seifart, Frank. Forthcoming. El sistema de clasificación nominal del miraña. (Lenguas Aborígenes de Colombia Descripciones 13.) Bogotá: Universidad de los Andes / CCELA.
- Thiesen, Wesley. 1996. Gramática del idioma Bora. (Serie Lingüística Peruana 38.) Yarinocchoa, Pucallpa (Perú): Instituto Lingüístico de Verano.

Toward a typology of position class: comparing Navajo and Ket verb morphology

Edward Vajda

Western Washington University

0. Introduction. This paper argues that the Navajo verb – despite its overall linear complexity, its interdigitation of lexico-derivational and inflectional elements, and the degree of fusion among certain of its components – does not differ fundamentally from processes of lexical stem derivation followed by inflection typical of morphologically simpler 'stem-and-affix' languages. In this view I support a growing number of Athabaskanists – notably Rice & Saxon (1994), McDonough (1997, 2000), and Rice (2000) – in arguing that Athabaskan verb morphology is no more 'templatic' than that of most other synthetic languages, at least insofar as how that term is defined by Stump (1997). To demonstrate that position class plays no special role in Navajo, I go beyond Athabaskan to examine a language where position class does represent a typologically significant feature of the verbal morphology, since it is used as a component of derivation rather than inflection alone. This language is Ket, an isolate spoken in Central Siberia. With reference to data introduced in Vajda (2000, 2001, in press), I demonstrate that the polysynthetic Ket verb displays a previously undocumented type of interface between the derivation and inflection – one in which the morpheme positions used to express subject and object agreement are chosen idiosyncratically from among several lexically competing position class models, even though the agreement morphs that fill these positions are determined syntactically. Based upon its use of inflection position as an integral and largely unpredictable component of verb stem derivation, I conclude that Ket, unlike Navajo, does employ position class as a non-trivial morphological category. Ket, in fact, appears to exhibit what Stump (1997) referred to as 'templatic derivation', whereas Navajo (and, more broadly, Athabaskan as a whole) does not.

1. Navajo verb morphology. The template-like scheme of Navajo verb structure usually presented – Young & Morgan (1987:37-38) and, most recently, Young (2000: 18-26) – focuses no special attention on which positions are derivational as opposed to inflectional. When that emphasis is added, an important detail emerges. As a general rule, each position is either strictly derivational or strictly inflectional. In Table 1, the Roman numerals and lower-case letters follow Young & Morgan's traditional morpheme

position designations. The letter L marks positions filled during stem creation, while G marks grammatically functional morpheme positions:

Table 1. Navajo verb morpheme positions

0	Ia	Ib	Ic	Id/e	II	III	IV	V	VI	VII	VIII	IX	X
G	G	L	G	L	L	G	G	G	L	G	G	L	L

The grammatical positions function as follows: 0 – postpositional object or possessor of a verb-prefixed noun; Ia – subject of an intransitive inchoative verb, direct object of a transitive inchoative verb, or indirect object of certain non-inchoative verbs; Ic – reciprocal; III – distributive plural; IV – direct object of non-inchoative verbs; V – subject agentivity or non speech-act participant number; and VII – tense-mood-aspect; VIII – subject agreement (except in the case of intransitive inchoative verbs). As far as concerns position X (the verb stem), Leer (1979) demonstrated that modern tense- or mood-linked alternations in the shape of position X morphemes derive from the historical addition of various aspectual suffixes to the verb root. The resultant partial suppletion in the verb base functions today to convey tense/mood/aspect forms. Rice (2000: 296-298) has argued that the original, underlying suffixes are best regarded as derivational elements rather than inflections, in which case position X could be considered entirely lexical. It is also possible to view the verb base as a portmanteau of two positions, one lexical (the root), the other either grammatical or lexical, depending on how one interprets the function of the original suffixes. In either case, my division of the Navajo verb form into discrete lexical vs. grammatical components would remain valid.

Even though most inflections are interdigitated between various portions of the verb stem, which is thus often discontinuous, there is no need to posit a morphological template as a component of Navajo verb stem derivation. The position of every inflection is selected according to a universal morphosyntactic pattern. Different grammatical features of the subject are predictably cross-referenced in positions V and VIII (Rice & Saxon 1994; Rice 2000: 244); the direct object of a transitive verb is normally cross-referenced in position IV (or by a non-agreement unspecified object marker); a postpositional object in position 0; and an indirect object without a postposition in position Ia. Transitive inchoative stems cross-reference their objects in position Ia, while intransitive inchoative verbs cross-reference their subject in this position. Otherwise, Navajo verb inflection displays nothing resembling the functional plasticity that we find associated with agreement morpheme positions in Ket (section 3 below). For example,

no Navajo verb uses position VIII (the subject position) to cross-reference a direct object. All inflection positions function according to a clearly predictable pattern across the entire Navajo verbal lexicon. Once processes of fusion are accounted for, Navajo verb forms therefore easily succumb to an attempt to divide their constituent morpheme positions into a lexical stem plus one or more inflectional affixes. One must agree with views expressed recently by both McDonough (2000) and Rice (2000) – despite their quite different estimations regarding the motivating principle behind Navajo verb form creation – that the template is nothing more than a descriptive convenience and not an independently relevant morphological category of the language. Or, it could be added, if there is a template then it is an epiphenomenon of the inflectional morphosyntax, just as is demonstrably the case in linearly more conventional 'stem-and-affix' languages (cf. Stump 1997). There is no evidence that position class plays any idiosyncratic role in Navajo verb stem derivation at all. At best it provides a convenient vehicle for illustrating the relative positions of lexical elements in discontinuous stems. All verbs in Navajo subscribe to the same set of overall grammatical rules for actant agreement marker placement.

2. Ket verb stems. Evidence provided in Vajda (2000, 2001, and in press) shows that the situation in Ket is radically different with respect to subject/object agreement position selection. Although Ket and Navajo appear at first glance to share the same general type of interdigitation of lexical and inflectional morphemes, the subject-object positions in Ket are selected idiosyncratically during stem derivation, even though the morphs that fill them are selected in the syntax during verb phrase formation. For this reason, it is impossible to provide a fully specified morpheme position model for all Ket verb stems equivalent to that often used for describing Navajo verb morphology. Table 2 provides a generic position class model for Ket finite verb form creation:

Table 2. Generic position class specifications in Ket finite verb form creation

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
actancy	incor- porate	actancy +spatial viewpoint	determiner viewpoint	durative marker or actancy	actancy	preterite or imperative	actancy	base	actancy

Most of the six slots marked 'actancy' have the potential of containing either a subject or direct object agreement marker, depending on the verbal lexeme in question. Which slot will be used for what grammatical function in any given stem is a purely lexical choice.

In terms of their agreement strategy, most Ket verb stems belongs to one of five actant conjugations distinguished by their choice of positional combinations. In the Ket verbal lexicon as a whole, the functions of the various possible actant positions overlap to such a high degree that it is only possible to specify them for each actant conjugation separately. Of all the morpheme positions involved in expressing the Ket verb's three morphosyntactic categories – tense, mood, and subject/object agreement – only the tense-mood slot (P2) is universal across the verbal lexicon. It is the only slot that truly resembles the sort of inflection position selection typical of Navajo and most other synthetic languages. With regard to subject/object agreement, on the other hand, Ket uses inflectional morpheme position selection as a component of stem derivation. This means that morpheme position selection – and therefore actant agreement position itself – is an idiosyncratic lexical category in Ket in a way quite unlike Navajo or other supposedly 'templatic' languages.

3. Ket actant conjugations and allo-templates. Instead of the type of fully specified inflectional model found in Navajo, each individual Ket verb stem conforms to one of five distinct 'allo-templates'. The choice between these lexically competing inflection position models is not grammatically meaningful. Table 3 shows the syntactic agreement morphs that appear in each derivationally selected slot. Abbreviations follow Dixon (1999): A = transitive subject, O = direct object, Sa = active intransitive subject, So = inactive intransitive subject:

Table 3. Subject/object agreement affixes

position	P8	P6	P4	P3	P1	P-1
agreement	(person/class)	(person/class/number)	3 an.	3 n.	1/2 or redundant	(plural)
functions	all A	some O	some 3an O	some neuter O	some 1/2 O	most A
	some Sa and So	some Sa and So	some 3an So	some neuter So	some 1/2 So	
redundant marker		redundant A, Sa, or So marker in some verbs			redundant A or S marker in some verbs	
1S	<i>di</i>	<i>ba ~ bo</i>	-	-	<i>di</i> (rarely Ø)	-
2S	<i>ku</i>	<i>ku</i>	-	-	<i>ku</i> (rarely Ø)	-
3M	<i>du</i>	<i>a ~ o ~ bu</i>	<i>a</i> (o in PT)	-	<i>a</i> (rarely Ø)	-
3F	<i>da</i>	<i>i ~ u ~ bu</i>	<i>i</i>	-	<i>a</i> (rarely Ø)	-
3N (S or PL)	<i>da</i>	Ø ~ <i>u ~ bu</i>	-	<i>b</i> (rarely Ø)	<i>a</i> (rarely Ø)	-
1PL	<i>di</i>	<i>dəŋ</i>	-	-	<i>daŋ</i>	<i>n</i>
2PL	<i>ku</i>	<i>kəŋ</i>	-	-	<i>kaŋ</i>	<i>n</i>
3AP	<i>du</i>	<i>aŋ ~ oŋ ~ bu</i>	<i>aŋ</i> (oŋ in PT)	-	<i>aŋ</i>	<i>n</i>

The P6 labialized variants (*o/u*) occur in verbs conveying action that involves movement away from of the term they cross-reference, while the non-labialized P6 variants (*a/i/Ø*) convey action directed toward that term or a stative event. P6 *bu* is a redundant 3rd person subject marker sometimes used derivationally to build involuntary causatives, semelfactives, or auto-instrumentals. Also, the P6 agreement markers is normally preceded by a postpositional consonant expressing a type of lexical situation aspect; this is usually *t* (internal state or action on a surface) or *k* (dynamic action toward or away from the actant cross-referenced in P6). These elements are not glossed separately in the examples below.

Overall, Ket cannot be characterized as conforming to any single agreement typology, since certain lexical actant sub-patterns are ergative, others active/inactive (split-S), and still others nominative/accusative. The five position configurations used productively in the language to express subject/object agreement semantically overlap to such a degree that Ket subject/object agreement must be regarded as a derivational category that expresses grammatical function.

Despite the proportionately large number of agreement slots (six out of a total of 10), Ket shows almost no one-to-one correspondences between individual semantic roles or syntactic functions and particular morpheme positions. What is unique about Ket typologically is that lexical stem derivation rather than grammatical typology determines the morphosyntactic agreement pattern characteristic of each particular verb (in other words, which lexical allo-template will be used by a particular stem). Tables 4-9 specify the functions each actant agreement position fulfills in each actant conjugation. Note that all five conjugations contain monovalent as well as bivalent verbs, and all can express either active or stative events. The examples follow Vajda (in press), which provides more complete paradigms illustrating each conjugation.

One common lexical actant agreement strategy, called Active Conjugation, resembles the binary split-S pattern found in many languages (no Ket verbs are fluid-S):

Table 4. Active Conjugation allo-template

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
active agent (person/class)	incorporate	- determiner		durative marker or 3animate patient	3neuter patient	preterite or imperative	1, 2 patient	base	active animate (plural)

In this conjugation, most intransitives require their animate subjects to be marked in P8 regardless of finer situational nuances of control or volition: *dàngìsta* 'he is hanging' [*du*⁸-*añ*⁷-*k*⁵-(*s*)-*ta*⁰ 3M.SJ⁸-hang⁷-away⁵-be.extended⁰]. And most inanimate subjects are marked in P3: *àngìpta* 'it is hanging' [*añ*⁷-*k*⁵-*b*³-*ta*⁰ hang⁷-away⁵-3N.SJ³-be.extended⁰]. For Active Conjugation transitives, S is obligatorily cross-referenced in P8 and O in P3 (inanimate), P4 (animate 3rd person) or P1 (1st or 2nd person). Active subject agreement involves a sort of circumfix composed of P8 (person/class) and P-1 (number): *dílkànsin* 'we dressed you.PL' [*di*⁸-*il*²-*kan*¹-*s*⁰-*n*¹ 'we dress you.PL' [1SJ⁸-PT²-2PL.O¹-dress⁰-AP¹*dílkànsin*].

In Possessive Conjugation, the subject is cross-referenced by possessive pronominal proclitics in P7 rather than regular actant agreement morphs:

Table 5. Possessive Conjugation allo-template

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
-	incorporate	-	-	durative marker	non-	preterite	non-	base	-
+ genitive prefix (person/number/class)					agreement	or imperative	agreement		

Most Possessive Conjugation verbs are intransitive and denote the production of sound: *ablákèjbátà* 'I am clapping' [*ab/lakej*⁷-*b*³-*a*¹-*ta*⁰ my/clapping⁷-IC³-SR¹-be.extended⁰].

Absolutive Conjugation uses P6 to mark any intransitive subject (whether active or inactive, animate or inanimate), as well as any direct object. A transitive subject is marked in P8. Semantic roles have no independent effect on the expression of syntactic agreement functions.

Table 6. Absolutive Conjugation allo-template

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
transitive subject (person/class)	incorporate	intransitive subject or direct object	determiner	durative marker	non-agreement	preterite or imperative	-	base	transitive subject (plural)
		+ spatial viewpoint							

Absolutive Conjugation verbs are virtually unrestricted in the meanings they can convey, and the difference with Active Conjugation is purely derivational, with no systematic correlation in semantics or syntax. Compare the Active intransitives *dólàq* 'he walked out' [*du*⁸-*o*⁴-*il*²-*aq*⁰ 3M.SJ⁸-D⁴-PT²-go⁰] and *íyàjbet* 'he spends the day' [*iγ*⁷-*a*⁴-*bet*⁰ day⁷-

3M.SJ⁴-make⁰] with the functionally parallel Absolutive intransitives *óγòvinden* 'he ran out' [*o/k⁶-o⁴-b³-in²-den⁰* 3M.SJ⁶-D⁴-intensity³-PT²-go⁰] and *áyìssal* 'he spends the night' [*a/k⁶-(s)-sal⁰* 3M.SJ⁶-spend.night⁰], which mark the subject in P6 rather than P8.

The choice between which transitive verbs belong to Active Conjugation and cross-reference their objects with inactive markers (P4-3-1), and which belong to Absolutive Conjugation and marks their object in P6, also lacks any grammatical significance whatsoever. This leads to many synonymous or nearly synonymous pairs. One is the Active transitive *dábbàk* 'she drags it (once)' [*da⁸-b³-bak⁰* 3F.SJ⁸-3N.O³-drag⁰] and its iterative counterpart *dabágdèñúyàvet* 'she drags it (often)' [*da⁸-bag/den⁷-u/g⁶-a⁴-bet⁰* 3F.SJ⁸-drag/ITER⁷-3N.O⁶-D⁴-ITER⁰], which belongs to Absolutive Conjugation. Intransitive synonyms include *déqsàq* 'he hears' [*du⁸-eq⁷-(s)-aq⁰* 3M.SJ⁸-L⁷-L⁰], an Active verb with P8 subject marking, and *áyàbda* 'he hears' [*a/k⁶-a⁴-b³-da⁰* 3M.SJ⁶-D⁴-IC³-hear⁰], an Absolutive verb with the subject marked in P6.

The fourth derivational subtype is called Coreferential Absolutive Conjugation because it obligatorily requires a subject marker in P8 as well as P6. Once again, agreement with a transitive subject requires a combination of P8 (person/class) and P-1 (number). P6 redundantly cross-references subject person (as well as subject number in the case of 1st or 2nd person subjects). The direct object is marked in P1 (for 1st and 2nd person; in P3 for 3rd person neuter, and in P4 for 3rd person animate objects): *dabúgdìt* 'she carries me' [*da⁸-bu/k⁶-di¹-t⁰* 3F.SJ⁸-3RS⁶-1S.O¹-L⁰], *dabúgbìt* 'she carries it' [*da⁸-bu/k⁶-b³-t⁰* 3F.SJ⁸-3RS⁶-3N.O³-L⁰], *dabúgàjit* 'she carries him' [*da⁸-bu/k⁶-a⁴-(j)-t⁰* 3F.SJ⁸-3RS⁶-3M.O⁴-L⁰]. Even though it expresses grammatical agreement, the coreferential subject marker in P6 functions as a derivational element expressing that the action is performed without the aid of a sled or other external means of conveyance.

Table 7. Coreferential Absolutive Conjugation allo-template

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
subject (person/class)	incorporate	redundant subject marker (person/number) +spatial viewpoint	determiner	durative marker or 3 anim.obj	3 neuter obj	preterite or imperative	1, 2 object	base	subject (plural)

In the fifth and final derivational actant marking type, the subject is obligatorily cross-referenced in P1 and P8 for animates: *dájàtij* 'he grows' [*du⁸-a⁴-(j)-a¹-tij⁰* 3M.SJ⁸-D⁴-3S.RS¹-grow⁰]; or P1 and P3 for inanimate: [*a⁴-b³-a¹-tij⁰* D⁴-3N.SJ³-3S.RS¹-grow⁰]. The

direct object is marked in P6. Transitive verbs are unproductive in this lexical actant agreement type, but they do occur: *dóksàva* 'he sells him off' [*du*⁸-*o/k*⁶-(*s*)-*a*¹-*qa*⁰ 3M.SJ⁸-3M.O⁶-3AP.RS¹-sell⁰].

Table 9. Coreferential Inactive Conjugation allo-template

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
animate or transitive subject (person/class)	incorporate	direct object +spatial viewpoint	determiner	durative marker	inanimate inactive subject	preterite or imperative	redundant subject marker (person/number)	base	-

Unlike other conjugations that use P8 to mark an active or transitive subject, Coreferential Inactive verbs do not normally cross-reference a plural animate subject with *-n* in P-1. Subject number is only expressed by the otherwise redundant P1 subject marker.

Although most Coreferential Inactive verbs are intransitive, monovalent verbs with similar meanings may belong to any conjugation. Compare the Coreferential Inactive verb *daájàtij* 'she grows' [*da*⁸-*a*⁴-(*j*)-*a*¹-*tij*⁰ 3F.SJ⁸-D⁴-3S.RS¹-grow⁰] with the Active Conjugation verb *daqáyàkan* 'she grows big' [*da*⁸-*qa*⁷-*a*⁴-*qan*⁰ 3F.SJ⁸-big⁷-D⁴-flow⁰], or the Coreferential Inactive verb *datájàraq* 'she falls' [*da*⁸-*t*⁵-*a*⁴-(*j*)-*a*¹-*daq*⁰ 3F.SJ⁸-L⁵-D⁴-3S.RS¹-fall⁰] with the Absolutive Conjugation verb *úgbùn* 'she slips' [*u/k*⁶-*b*³-*un*⁰ 3F.SJ⁶-IC³-slip⁰]. Also compare the Active Conjugation semelfactive *dílsìbet* 'he takes a breath' [*du*⁸-*il*⁷-(*s*)-*bet*⁰ 3M.SJ⁸-breath⁷-make⁰] with its iterative counterpart *déliŋ ísàbet* 'he breathes' [*du*⁸-*eliŋ*⁷-*i*-(*s*)-*a*¹-*bet*⁰ 3M.SJ⁸-breaths⁷-3S.RS¹-make⁰], which belongs to Coreferential Inactive Conjugation.

Extensive synonymy among verb stems belonging to different actant conjugations underscores the derivational function of subject-object agreement position selection. In general, a verb belonging to any conjugation has at least the potential of being roughly synonymous with a verb of any other:

(1) Derivational patterns that build involuntary causatives

P8: *dītáàl* 'I get cold' [*di*⁸-*təə*⁰ 1SJ⁸-freeze⁰]

P4-3-1: *úsàren* 'I fall asleep' [*us*⁷-*a*⁴-*di*¹-*en*⁰ sleep⁷-D⁴-1S.SJ¹-intransitive⁰]

sítòna 'he woke up' [*si*⁷-*o*⁴-*in*²-*a*⁰ wake⁷-3M.SJ⁴-PT²-state⁰]

P6: *ábàtavan* 'I sweat' *a*⁷-*balt*⁶-*a*⁴-*qan*⁰ heat⁷-1S.SJ⁶-D⁴-flow⁰]

P8 + P6: *dabútì lok* 'she shuddered' [*da⁸-bul⁶-il²-ok⁰* 3F.SJ⁸-3RS⁶-PT²-up⁰]

P6 + P3 *b*: *ógbùn* 'he slips' [*ulk⁶-b³-un⁰* 3F.SJ⁶-IC³-slip⁰]

éjbà yaptal 'I freeze to death' [*ej⁷-balk⁶-a⁴-b³-tal⁰* kill⁷-1S.SJ⁶-D⁴-IC³-freeze⁰]

(2) Subject marking in auto-causatives (actions involving no object or external stimulus)

P6: *úyàtn* 'she goes' [*ulk⁶-a⁴-tn⁰* 3F.SJ⁶-D⁴-go⁰]

P8: *ī daésàkut* 'the sun rises' [*da⁸-es⁷-a⁴-qut⁰* 3F.SJ⁸-up⁷-D⁴-move⁰]

daqá yàkan 'she grows big' [*da⁸-qa⁷-a⁴-qan⁰* 3F.SJ⁸-big⁷-D⁴-flow⁰]

P8 + P1: *da.ájàtij* 'she grows' [*da⁸-a⁴-(j)-a¹-tij⁰* 3F.SJ⁸-D⁴-3S.RS¹-grow⁰]

díndì ruk 'I returned' [*di⁸-in²-di¹-duk⁰* 1SJ⁸-PT²-1S.RS¹-back⁰]

P8 + P6: *súgbà yonden* 'I returned' [*di⁸-sug⁷-balk⁶-o⁴-in²-den⁰* 1SJ⁸-back⁷-1S.RS⁶-D⁴-PT²-go⁰]

P7 'self' + P0 incorporated SJ: *bíṣìvej* 'wind blows' [*bin⁷-k⁵-(s)-bej⁰* self⁷-away⁵-wind⁰]

(3) Derivational patterns that build stative resultatives

Active marking: *daséstà* 'she is sitting' [*da⁸-ses⁷-a⁴-ta⁰* 3F.SJ⁸-place⁷-D⁴-extend⁰]

Inactive marking: *dóndì bet* 'I have a knife' [*don⁷-di¹-bet⁰* knife⁷-1S.SJ¹-have⁰]

P6 + P1: *bótàjòksàjabet* 'he's been made rich' [*botaj⁷-olk⁶-(s)-a¹-(j)-abet⁰* rich⁷-3M.SJ⁶-SR¹-make⁰]

(4) Derivational patterns specifying the use of an external tool

P7: *dadónbàtet* 'she stabs me' [*da⁸-don⁷-balk⁶-a⁴-tet⁰* 3F.SJ⁸-knife⁷-1S.O⁶-D⁴-hit⁰].

P3 *b*: *dúgdàptan* 'I drag it (by sled)' [*di⁸-ulk⁶-d⁵-a⁴-b³-tan⁰* 1SJ⁸-3N.O⁶-across⁵-D⁴-AL³-drag⁰].

Redundant P1 subject: *datísàro* 'she loads' [*da⁸-ti⁷-(s)-a¹-ro⁰* 3F.SJ⁸-bowstring⁷-3S.RS¹-stretch⁰]

(5) Derivational patterns that build auto-instrumentals (blocking external tool)

P0: *díptàn* 'I drag it (by hand)' [*di⁸-b³-tan⁰* 1SJ⁸-3N.O³-drag⁰]

Redundant P6 subject marker: *dabúgdìt* 'she carries me' [*da⁸-bul⁶-di¹-t⁰* 3F.SJ⁸-3RS⁶-1S.O¹-L⁰]

(6) Derivational patterns that build iteratives or distributives

Special P0 morpheme: *dítàramin* 'we are lying prone' [*di⁸-t⁵-a⁴-damin⁰* 1SJ⁸-static⁵-D⁴-PL.SJ⁰]

Special P7 morpheme: *datóqṣìbèt* 'she keeps stepping' [*da⁸-toqṣ⁷-bet⁰* 3F.SJ⁸-footsteps⁷-make⁰]

Non-agreement P3-1: *ablákèjbátà* 'I am clapping' [*abllakej⁷-b³-a¹-ta⁰* my/clap⁷-IC³-SR¹-extend⁰]

P5-4-3: *dasáqtàptet* 'she stabs him' [*da⁸-saq⁷-t⁵-a⁴-do⁰* 3F.SJ⁸-step⁷-ITER⁵⁻⁴⁻³-hit⁰]

Nor do example sets (1-6) exhaust the derivational patterns available to convey each semantic group.

5. Conclusion. Modern Ket appears to be typologically unique among documented living languages for its multiple, highly idiosyncratic derivational split in subject/object agreement marking. In terms of its overall grammatical typology, Ket is best described as a language that cross references the syntactic subject and object noun phrases (both of which are zero marked in the absolutive case) verb internally, with the positional marking of S sometimes reflecting of its semantic role (in the conjugations that use P3 as an agreement marker). The actant conjugations themselves are merely competing lexical expressions of the same grammatical function: number/person/class agreement with the syntactic terms coordinated with the verb during verb phrase formation. Even when a particular agreement pattern appears to correlate with a recognizable semantic feature, it is a derivational rather than grammatical correlation, and one finds similar meanings conveyed by verbs belonging to other conjugations. For this reason, any overall position-class model for the Ket verb must remain partly under-specified, as given previously in Table 2. Otherwise, the functional specifications of the slots involved in actant marking become contradictory or redundant, as in Table 10:

Table 10. Fully functionally-specified position-class model for the Ket verb

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
SJ agr. + SJ agr. (person, class)	incorporate + SJ agr. (person/class/numb.) or redundant SJ (person/numb.) + situation aspect	SJ or O agr. (person/class/numb.) or redundant SJ (person/numb.)	determiner (person/class/numb.) or redundant SJ (person/numb.)	durative or 3 animate SJ or O agr. (person/numb.)	3 neuter O agr. or instrumental applicative, or involuntary causative	preterite or imperative or instrumental applicative, or involuntary causative	1 or 2 SJ or O agr. (person/number) or redundant SJ, or stative resultative	base	animate plural SJ agr.

The grammatically overlapping labels given under slots P8, P6, P4, P3, P1, and P-1 reflect the fact that Ket verb structure involves five lexically competing strategies for expressing actant agreement. In terms of morpheme position class configuration, these strategies represent lexico-derivational 'allo-templates' of a more abstract inflectional master template. Non-templatic derivational morphology, such as that found in Navajo and most other languages, is associated with a single-level, fully specified model of inflectional morpheme position configuration; and there is no real derivational morpheme position template to speak of.

The unparalleled degree to which the modern Ket verb lexically determines the linear configuration of its agreement inflections certainly ranks as the language's most

striking typological feature. In most synthetic or even polysynthetic languages, verb stems resemble nominal stems in that both created in the lexicon and modified in the syntax by elements prefixed, infix, or suffixed into a specific inflectional zone or series of slots, which may be discontinuous, as in the Navajo verb. All stems belonging to each particular form class normally inflect according to a single formula. The formula differs lexically only insofar as some verbs are monovalent, others polyvalent; or in the case of split-S or fluid-S languages, according to some generalizable semantic dichotomy (Mithun 1991). This overall, predictable uniformity in inflection position selection is typical even of such morphologically complex languages as Navajo. Despite the interdigitation of Athabaskan inflections within the verb stem, inflection position selection remains operationally distinct from the process of stem derivation. In the case of a flectional language such as Russian or most other Indo-European languages, selecting the proper inflectional allomorphs often requires a look-back from the grammar to the lexicon to establish the proper declension or conjugation membership. This is certainly true in the selection of tense, mood and aspect affix forms in the Navajo verb (and in selecting the form of the P2 tense-mood inflectional affix in the Ket verb). But there is normally no lexical look-ahead to the inflection-bearing template during stem derivation, since inflection position configuration is globally specified in the syntax for each form class. All Latin nouns or Navajo verbs inflect according to a single position-class model, so one could say that Latin or Navajo has a single 'allo-template' for each of its form classes. Despite the non-discrete nature of many Navajo verb stems, there is no need for positing templatic structures in the lexicon. And in the syntax, such structures are trivial epiphenomena that tend predictably to reflect scopal relations in Universal Grammar (Rice 2000). Ket uses this strategy of derivation followed by inflection for all form classes except its finite verb. Among the verbal morpheme positions that express grammatical categories, only the P2 tense and mood slot is added by a global grammatical rule (thus one could say there is a single template for Ket tense and mood inflection). The configuration of subject-object positions, by contrast, is lexical and idiosyncratic, even though the morphs that fill these positions are chosen syntactically. This rare technique of stem formation could be called 'paradigmatic derivation' since it involves a lexical choice between several competing inflection position configurations. Ket verb morphology demonstrates by comparison that Athabaskan verb structure does indeed show a greater typological affinity with 'non-templatic', stem-and-affix languages as regards its morphology-syntax interface.

Abbreviations

agr. - agreement

AN - animate class

AL - applicative infix (action performed using a tool or means of conveyance)

AP - animate plural (denotes the animate plural subject agreement suffix)

APS - animate plural subject affix

D - durative marker (appears in many verbs that denote states or activities)

F - feminine class

IC - involuntary causative (denotes an event occurring by accident)

L - lexico-derivational element with opaque semantic content

M - masculine class

N - neuter (inanimate) class

O - verb-internal direct object agreement affix

PL – plural

PT - preterite (=past tense)

RS - redundant subject agreement marker (obligatory derivational affix)

S – singular

SJ - verb-internal subject agreement affix

SR - stative resultative infix (denoting a state caused by a previous event);

References

- Dixon, R. M. W. 1999. Semantic roles and syntactic functions: the semantic basis for a typology. *Chicago Linguistics Society* 35/2: 323-41.
- Leer, Jeff. 1979. *Proto-Athabaskan verb stem variation, part 1: phonology*. Fairbanks, AK: ANLC.
- McDonough, Joyce. 1997. Athabaskan redux: against the position class as a morphological category. *Morphologica* 1996.
- McDonough, Joyce. 2000. On a bipartite model of the Athabaskan verb. *The Athabaskan languages: perspectives on a Native American language family*, eds. Theodore B. Fernald & Paul R. Platero, 139-66. Oxford: Oxford University Press,
- Mithun, Marianne. 1991. Active/agentive case marking and its motivation. *Language* 67: 510-46.

- Rice, Keren, & Leslie Saxon. 1994. The subject positions in Athapaskan languages. *The morphology-syntax connection*, eds. Heidi Harley & Colin Phillips, 173-95. *MIT Working Papers in Linguistics* 22. Cambridge, MA: MIT WPL.
- Rice, Keren. 2000. *Morpheme order and semantic scope: word formation in the Athapaskan verb*. Cambridge: Cambridge University Press.
- Stump, Gregory T. 1997. Template morphology and inflectional morphology. *Yearbook of morphology 1996*, 217-41, eds. Geert Booij & Jaap van Marle. Amsterdam: Kluwer.
- Vajda, Edward J. 2000. Aktantnye sprjazhenija v ketskom jazyke [Actant conjugations in Ket]. *Voprosy jazykoznanija* 67/3:21-41.
- _____. 2001. The role of position class in Ket verb morphology. *Word* 52/3:369-436.
- _____. [in press]. Ket morphology. *Morphologies of Asia and Africa*, ed. Alan Kaye. Winona Lake, IN: Eisenbrauns.
- Young, Robert W., & William Morgan, Sr. 1987. *The Navajo language: a grammar and colloquial dictionary*, 2 ed. Albuquerque: University of New Mexico Press.
- Young, Robert W. 2000. *The Navajo verb system: an overview*. Albuquerque: University of New Mexico Press.

The semantics of Yurok Intensive infixation*

Esther Wood and Andrew Garrett

University of California, Berkeley

1. Introduction. In his 1958 grammar of the Yurok language of northwestern California, R. H. Robins introduced the term “intensive” for a central grammatical category of that language. We will retain his term in our discussion of the semantics of this category, though it remains to be seen whether actual “intensive” meanings exist and (if so) how they arise. Our analysis is based on published and unpublished texts and preliminary fieldwork, and our results at this point are provisional.

The Yurok Intensive has two main surface forms, both resulting from an underlying infix *-eg-*. Typically, as illustrated in (1), the Intensive surfaces with precisely this form: an infix positioned after the onset consonant or cluster of its base.

(1)	BASE		INTENSIVE
	ko'moy-	“hear”	k-eg-o'moy-
	laay-	“pass”	l-eg-aay-
	hlkyorkw-	“watch”	hlky-eg-orkw-
	trahk-	“fetch water”	tr-eg-ahk-

Alternatively, if the base vowel is *-e-* and is followed by a velar, the Intensive has an apparent ablaut variant. Before a nonlabial velar the Intensive takes the form of *-ii-* ablaut (*CeK- → *C-eg-eK- > *C-iy-iK- > CiiK-); before a labialized velar it takes the form of *-uu-* ablaut (*CeKw- → *C-eg-eKw- > CuuKw-). These variants are illustrated in (2).

(2)	BASE		INTENSIVE
	a. kweget	“visit”	kwiiget
	lekoo(t-)	“stab”	liikoo(t-)
	nek-	“put”	niik-
	b. lekwohl-	“fall, fill the air”	luukwohl-
	new-	“see”	nuuw-
	pew(om)-	“cook”	puuwom-

Finally, note that some forms show an apparent double Intensive formation **-eg-eg-* > *-iig-* (e.g. *cwin-* “speak” → *cwiigin-*), which we will discuss later in the paper.

Though the meaning of the Yurok Intensive has not previously been studied in detail, brief statements exist. For example, Robins (1958: 82) wrote that its “basic meaning ... is intensity, plurality, or iteration ... of the action, state, or process denoted by the verb.” In a pair of 1985 articles he called the formation “frequentative”, which is reminiscent of the term “iterative” used by A. L. Kroeber as well as the brief description given by Kroeber’s colleague T. T. Waterman, who wrote that the formation “means *often* or *habitually*, adding the idea of frequent or customary occurrence.”¹

Clearly, insofar as its meanings involve frequency, habituality, iteration, and plurality of action, the Intensive belongs in the typological arena of pluractionality or verbal plurality. Such phenomena were examined in Sapir’s classic treatments of Takelma (1922: 127-34) and Southern Paiute (1930: 148-59, 236-41, 256-62) and have been investigated more recently by Cusic (1981), Mithun (1988a, b), Lasersohn (1995), and others (Newman 1990, Xrakovskij

1997, Yu 2001). Cusic and Lasersohn in particular highlight the important typology in (3).

- (3) THREE KINDS OF ITERATION OR REPETITION (Cusic 1981, Lasersohn 1995)
- Event-internal (event-phase) repetition, e.g. *bite* → *nibble*, *fly* → *flutter*
 - Event-external repetition, single occasion, e.g. *keep biting*, *fly back and forth*
 - Event-external repetition, multiple occasions, e.g. *bite regularly*, *fly frequently*

One kind of verbal plurality, called “repetitive” aspect by Cusic, involves repetition of the internal phases of an event. For example, as suggested in (3a), an event of nibbling or fluttering consists of iterations of the basic action associated with biting or flying respectively. Crucially, the iterated internal phases – each bite in (3a), or each flap of the wing – do not instantiate the same type of event as the overall event.

By contrast, the two kinds of pluractionality in (3b) and (3c) can be seen as event-external repetition. In each case a single type of event is repeated, either on one single occasion as in (3b) or on multiple occasions as in (3c). Cusic (1981) calls event-external repetition “repeated” aspect, while Bybee, Perkins, and Pagliuca (1994) use the terms “iterative” and “frequentative” for event-internal and event-external repetition respectively.²

In Yurok, there is an aspectual opposition between the Intensive and a reduplicative formation we will call the Repetitive (Garrett 2001). Their basic functions are as follows. Repetitive reduplication expresses event-internal repetition, mainly with semelfactive base verbs as in (4).

- | | | | | |
|-----|--------------|-----------------------|--------------------|-------------------------------|
| (4) | BASE | | REPETITIVE | |
| | ckem | “to count” | ckem-ckem | “to make small tattoo marks” |
| | pegon- | “to split” | peg-pegon- | “to split in several places” |
| | prkwrh(s-) | “to peck or knock” | prkw-prkwrh(s-) | “to peck or knock repeatedly” |
| | tek(toy-) | “to grow” (of plants) | tek-tekon- | “to grow in tufts” |
| | t’k’eroh(s-) | “to thump ...” | tk’er-t’k’eroh(s-) | “to thump repeatedly” |

The morphological details of reduplication are not the object of this paper, but we should note that the reduplicant is not always a single heavy syllable as in (4).

In contrast to the Repetitive, the Yurok Intensive typically expresses event-external repetition. To use the terminology of Bybee, Perkins, and Pagliuca, the Repetitive expresses iterative aspect while the Intensive expresses frequentative aspect. Typical Intensives appear in (5).

- (5) a. niko’hl ho ’ne-**kgrrtk**-rk’
 always PAST 1-go.fishing.INT-SG
 “I always used to fish.”
- b. kwilek wit kit mehl sonow-o’m k’i numi cu sook k’e-**negep**-ek’
 well this NR.FUT CIRC be.like-2SG the very all *sook* 2-eat.INT-SG
 “You are getting like this because you eat everything *sook* [= ‘reptiles ... as well as human foetuses’ (Buckley 1980: 156)].”
- c. k’i toomeni sonow-oni c’uc’iš numi cu **kegohce**’w
 ART all.sorts happen-ATTR.3PL bird very all **catch**.INT.3SG
 “He used to catch all sorts of birds.”

- d. hewon we-new-oy-hl woogey **negi'ihl**
 at.first 3-see-PASS-3PL white.man **be.two(human).INT**
 "When they were first seen, the white men went in pairs."

We can also cite minimal pairs to illustrate the Intensive. For example, in (6a), the first sentence refers to one of a series of deer swimming by, while the second sentence refers to the whole series.

- (6) a. kem pulekw niki **rur** ... nii'nowur k'is kwen **regur-in**
 again downriver CONS **swim** look where **swim.INT-ATTR.3PL**
 "Then it swam by downstream ... [The dogs said:] 'Look where they swim!'"
- b. nek kwelek kic 'i **sonow-ok'** 'oolekwohl 'o **segonow-oni**
 I well PERF LOC **be.like-1SG** people LOC **be.like.INT-ATTR.3PL**
 "I am now the way people get to be."

Likewise, in (6b), the verb meaning "be" or "happen" lacks the Intensive infix in the first clause, referring only to the present state of the subject. In the second clause, referring to how things will be for everybody in the future, the Intensive form appears.

The contrast between the Repetitive and Intensive can be seen in sentences like (7).

- (7) moco ku **megokw** kimi wey 'u-**mokwo-mokwoc-ek'** ko srmrt-rk'
 if that **dog** NEG.FUT finish 3-**REP-bark-SG** TEMP kill-1SG
 "If that dog does not stop barking I shall kill it."

The verb "bark" is Repetitive, since actual barking consists of a rapidly repeated series of barks, while the word for "dog" in (7) is itself derived as the Intensive of the same verb root; a dog is a habitual barker.

Words with verbal morphological structure can thus come to function as nouns. Such nouns may show Repetitive reduplication or Intensive infixation, and the expected semantic contrast between these categories is also seen in the nominal domain. In (8a) we give some examples of nouns derived as Intensives – like "dog" in (7) – and in (8b) we give a pair of Repetitive nouns.

- | (8) | BASE | | DERIVATIVE | |
|-----|------------|-------------------------|-------------------------|---|
| a. | muhl(koc-) | "sell" | meguhl | "store" (Berman 1982: 206) |
| | mesik- | "be thin" (worms, etc.) | megesik | "weasel" |
| | na'aw- | "catch surf fish" | nega' | "surf fish net" |
| | teykelum- | "bite" | tege'y | "flea" (3 sg. -') |
| b. | hlkrmrkikr | "single knot in a rope" | hlkrm-hlkrmrkikr | "knots in a rope" (Spott & Kroeber 1942: 216) |
| | mrkwrhl | "peak" | mrkwr-mrkwrhl | "series of peaks, mountain chain" |

The Intensive derivatives refer to characteristics or habitual activities, while the Repetitive nouns denote collections of point-like objects.

Before proceeding, we should note that the aspectual opposition between Repetitive and Intensive formations in Yurok appears to be of Algic antiquity. Underlying these formations historically is a pair of reduplicative constructions, preserved in the Repetitive with some reduction and transformed in the Intensive into the present-day infix (Garrett 2001). The re-

constructed ancestors of both Yurok formations are documented in Algonquian languages with comparable functions. The manifestly parallel character of the two reconstructed reduplicative formations supports our treatment of their Yurok descendants as aspectual complements.

2. Event-external single-occasion repetition. Two sorts of event-external repetition can in principle be distinguished, as indicated in (6), depending on whether an event is repeated on the same occasion or multiple occasions. Habitual and characterizing meanings fall into the multiple-occasion class, of course, but the Yurok Intensive can also express repetition of a single type of event on the same occasion. In such cases translations like “do repeatedly” or “keep doing” may be appropriate.

In an Intensive expressing repetition on a single occasion, the precise interpretation depends partly on the aspectual properties of the base verb. We have so far found examples of two varieties. First, if the base verb denotes an activity, then a single-occasion Intensive simply denotes the repetition of that activity. A few examples of this general type are cited in (9).

(9)	BASE		INTENSIVE (SINGLE-OCCASION REPETITION)
	a. crwrhs-	“point”	cuwrh(s-) “point repeatedly”
	b. nii'n(ow-)	“look (for)”	negii'n(ow-) “look around (for)”
	c. paahcew(-)	“move”	pegaahcew(-) “move around”
	d. swrrk'ws	“the wind gusts”	swrgrrk'ws “the wind gusts repeatedly”

We will give two textual examples of the Intensive in (9c). The first appears in (10).

(10)	'o	new	kaap'olihl	yo'	'o	pegaahcew
	LOC	see	plant	there	LOC	move.INT
	“(Then he almost heard someone speaking to him and looked back ...) He saw a plant moving ...”					

Kroeber translates “moving” in (10), but a plant is rooted and must therefore be making a series of plant-sized movements; in his field notes, Kroeber in fact wrote both “moving” and “shaking”. The sense of “moving around” is also present in our second example, in (11). Here the situation is that Coyote is lurking outside the sweathouse. In the first sentence he hears plural subjects eating together, while in the second sentence he hears a single subject repeatedly moving – as it were, first washing, then dressing, then having breakfast.

(11)	'ap	'o	ko'mo'y	'oc	ko'l	ka	nep-e'm	cmeyonen
	PAST	LOC	hear.3SG		something	PRES	eat-PL	evening
	kwushi'	'owoohl	koy numi	koy	'o	ko'moy-o'm	pegaahcew-o'm	kolin
	then	next.morning	very early		LOC	hear.3SG	move.INT-3SG	one
	“In the evening he heard people eating ... Next morning very early he heard one of them moving.”							

Second, if the base verb denotes an accomplishment “do X with result Y”, then a single-occasion Intensive means “do X with result Y, repeatedly”. Intensives of this type are cited in (12) from Proulx (1985); their morphology shows that these verbs are accomplishments. All are formed with a medial element expressing manner of motion – “drift”, “fly”, “run”, and “swim” in the examples cited – and an initial element delimiting that motion. The de-

limiting initials in (12) mean “back” and “in a circle”. The Intensives in (12) all express the repetition of activity and result phases of their accomplishment base verbs – for example, flying repeatedly back to a bird feeder.

(12)	ACCOMPLISHMENT BASE		INTENSIVE (RESULT REPETITION)	
	yohp-en-	“drift in a circle”	yegohp-en-	“drift around in circles”
	kwomhl-en-	“drift back”	kwegomhl-en-	“move back and forth on the water”
	kwomhl-ohl-	“return (of a bird), fly back”	kwegomhl-ohl-	“fly back and forth (of a bird), e.g. to a feeder”
	kwomhl-o’rep-	“run back”	kwegomhl-o’rep-	“run back and forth”
	kwomhl-ur-	“swim back”	kwegomhl-ur-	“swim back and forth”

Interestingly, we have not seen clear cases where the Intensive of an accomplishment verb expresses the repetition of only the activity phase but not the result phase. Such a hypothetical Intensive could mean something like “fly around and then back”. If this gap is not accidental, it is of potential interest for at least two reasons. First, as we will discuss later, there *are* some contexts where the Intensive is construed semantically with a morphological subpart of its base. Second, not all pluractional categories show the same restriction as Yurok. For example, Latin has an Iterative suffix expressing event-internal repetition (Garrett & Slatin in preparation). With an accomplishment verb, as in (13), the resulting interpretation is “to repeatedly engage in the activity phase of the base verb”.

(13)	LATIN ACCOMPLISHMENT VERB		ITERATIVE	
	advenīre	“arrive”	adventāre	“approach”
	cadere	“fall”	cassāre	“totter”
	dēlicere	“entice away (from)”	dēlectāre	“exert fascination on”
	labefacere	“loosen”	labefactāre	“work at loosening”
	prehendere	“seize, take hold of”	prensāre	“grasp at, keep grasping”

We take the difference between pluractional categories that do and do not license this activity-phase repetition interpretation to be a consequence of the fundamental difference between the event-internal repetition expressed by the Latin category and the event-external repetition expressed in Yurok.

3. Intensification. Robins chose the term “intensive” because the Yurok infix sometimes has what he considered an “intensive” function. This is an interesting area of research for at least two reasons. First, it is not obvious why a category that expresses event-external repetition should also come to express “intensification”, whatever exactly that is. Second, the development of apparently “intensive” meanings in categories of repetition is documented elsewhere, for example in Latin and Alabama (Hardy & Montler 1988, Kimball 1991).

We have found apparent “intensive” meanings in two main contexts, negative and stative contexts. The first of these was already noted by Robins, who wrote in his grammar that when a negative particle is construed with the Yurok Intensive “the negation is intensified” (1958: 82). Ordinary semantic composition would predict, if Intensives express event-external repetition, that the negative of an Intensive would mean that it is not the case that the base action is repeated. The actual meaning, as illustrated in (14), is that the base action never takes place.

- (14) a. **nima** 'wegookwc'
 NEG 3.gamble.INT
 "He never gambled."
- b. **nimi** 'u mep **kego'moy-ok'** wi'iit we-sook
 NEG PAST PAST hear.INT-1SG it 3-sort
 "I have never heard anything like it."
- c. **hikon** kwelekw **nimi** wi' mehl ho **regoowo's** 'oohl
 formerly well NEG it CIRC PAST smoke.INT.3SG people
 "In former times no one used pipes like this for smoking", literally "people did not smoke in this way."

For example, in (14a) the interpretation is "he never gambled" rather than "he didn't gamble regularly". The Intensive still expresses regularity, but regularity of the negative. We attribute this to what Horn calls "NEG-raising", defined as "the availability of a lower-clause reading or understanding for a higher-clause negation" (Horn 1989: 308). In the case at hand, a negative Intensive signifies not that an event is not regular but that it is, as it were, regularly not; i.e. $\text{NEG}(\text{REPET}(X)) \rightarrow \text{REPET}(\text{NEG}(X))$. Note that the relation between this meaning and simple negation involves semantic "intensification" only in a loose sense.

Stative predicates provide the second context where apparent "intensive" meanings arise. Of course, often enough the Intensive form of a stative verb is used when there are multiple instantiations of the state. This may occur if a state holds of many subjects, for example the salmon in (15a), or if a stage-level state holds of one subject at many different times, as in (15b-c).

- (15) a. **ki** **kwegomhl-e'mmi'** **kegesomewt-ehl** so mr'wrmry
 FUT return.INT-PL because be.homesick.INT-3PL to head.of.river
 "(It shall come to pass that the salmon) shall return, because they are homesick, to the head of the river."
- b. **nek** cpi yo' **kegesomew-ok'**
 I always PTCL be.lonely.INT-1SG
 "I always get lonesome."
- c. **ciweyk'** **cegiweyk'**
 be.hungry.1SG be.hungry.INT.1SG
 "I am hungry ... I am often hungry."

Note that the verbs in (15a-b) are morphologically very closely related.

Even an individual-level stative verb may show an Intensive form in contexts with a coerced nonstative reading, as in (16).

- (16) **cu** **ki** ni'iin-o' to' **kic** ni **negi'iin-o'**
 HORT FUT be.two-PL and PERF LOC be.two.INT-PL
 "Let us cohabit. We have been going together all this time."

This example shows a nice minimal pair. The first sentence is literally "Let us be a couple." The second should likewise mean "We have been a couple", but an activity reading instead emerges: "We have been acting as a couple." The Intensive form expresses the repetition of this activity.

Yet even putting aside examples of these types, there are some Intensives based on stative verbs where the interpretation does seem to involve “intensification” and a translation with “very” seems appropriate. Actual textual examples are relatively uncommon, and what follows is work in progress.

We begin in (17a) with a Yurok expression reported by Robert Spott, who elsewhere defined the same word *segeyoyhl* as “the time of year when the grass turns brown” (Spott & Kroeber 1942: 168).

- (17) a. kic **segey-oy-hl**
 PERF **burn.INT**-PASS-3PL
 “Growth is burned dry.”
- b. nimi ki kooyc ko'mi **tegenonihl**
 NEG FUT buy excessively **be.expensive.INT**
 “I shall not buy it, it is too expensive.”

Ignoring the force of the Intensive infix, the form plainly means “they have been burned”, referring to the grass that has all turned brown by this time of year. An Intensive is justified by the mass or plural status of the burned grasses. Any burning of grass, including natural summer browning, is of course a situation in which the degree of overall burning increases as the quantity of individual burnt grasses increases. In this context an intensive meaning “thoroughly burn” applied to a mass subject emerges more or less automatically from a multiple-event interpretation applied to individuated subjects. A burning is truly thorough if everything has burned.

This example suggests one general pathway to “intensification”. A predicate in which a state results from distributed predication may potentially yield an “intensive” reading if event multiplication is applied to the distributed predication. This mechanism also underlies the example in (17b). The verb *tenonihl* “be expensive” is derived from an initial element *ten-* and a medial element *-onihl*. This medial element means “cost”, “pay”, or the like, and it is clear from the data in (18) that *ten-* means “many” or “a lot of” rather than “to a great degree”.³

(18)	DERIVED VERB		MEDIAL ELEMENT
	ten-	“be much, be many”	
	ten-onihl	“be expensive”	-onihl “cost”, “pay”, etc.
	ten-owen-	“take much of, take more of”	
	ten-owohl-	“talk too much”	-owohl “talk”
	ten-pey-	“eat much”	-pey “eat”
	ten-sew-	“catch a lot of”	-sew “catch”
	ten-unow-	“grow thickly, grow in clumps”	

In short, the word “expensive” conceals a complex predicate meaning something like “costs a lot” – whether modern coins and bills or the traditional currency of dentalium shells. It is this mass to which the quantificational force of the Intensive applies.

We illustrate a final class of “intensification” examples with the two quotations in (19).

- (19) a. muscen kwilekw neka' ne-**tegenumonoksim-ek'**
 really well I 1-**wish.bad.luck.on.INT**-SG
 “I sure want (him) to die.”

- b. nek soo kic kegeycec mi' cpaanik' wi k'e-me'wom-e'm-o'w
 I think PERF **be.tired.INT** because be.long here 2-come.from-PL-2PL
 "I am sure you are tired, for your voyage here has been long."

The verbs in these quotations denote experienced states, and in (20) we give a full dossier of examples of this type known to us. The examples in (20a-b) are distinguished by suffix.

(20)	BASE VERB (EXPERIENCED STATE)		INTENSIVE	
a.	keyc-ek	"be tired"	kegeyc-ek	"be very tired" (19b)
b.	kaam-oksim-	"dislike"	kegaam-oksim	"really dislike" (K 74.25)
	tenumon-oksim-	"wish bad luck on"	tegenumon-oksim-	"wish lots of bad luck on" (19a)
	t-oksim-	"admire"	teg-oksim-	"praise" (R 257)
	wa's-ok(sim-)	"pity, be sorry for"	wega's-ok(sim-)	"be very sorry for"

(Waterman & Kroeber 1938: 76)

The medial suffix *-ok(sim)* in (20b) refers to thinking, feeling, or experience, and is further illustrated in (21).

(21)	INITIAL ELEMENT		DERIVED VERBS IN <i>-oks(im)</i> (Proulx 1985)
	cpaa (ni)	"far, late, long"	cpaan-oks-
	hoor-	"weave"	hoor-oks-
	kaam-	"bad"	kaam-oksim-
	kimol-	"be bad"	kimol-oksim-
	hlmey-	"fear"	hlmey-oksim-
	skew-	"good"	skew-oksim-
	kim soon-	"be badly"	kim soon-oksim-
	skuy soon-	"be well"	skuy soon-oksim-
	*t- (> ')-	"be"	t-oksim-
	wa's-	"poor"	wa's-ok(sim-)

The data in (21) show that a verb formed with the suffix *-ok(sim)* means basically to have feelings about someone or something that have the quality of the verb's initial element: good feelings, bad ones, long ones, and so on.

The relationship between the verb type in (20a) and the verb type in (20b) and (21) is shown by the verbs in (22).

- (22) a. keyc-ek "be tired" (20a)
 b. keyc-oksim- "get tired of" (Proulx 1985)

Both verbs mean "feel tired"; the difference is that *keyc-oksim-* means "have tired feelings directed at something or someone else", while *keyc-ek* just means "have tired feelings". In (23) we give a few additional uninflected verbs of experienced state in *-ek*.

(23)	kelom-	"turn"	kelomek	"be worried"
	tm-	"shoot"	tmek-tmek	"have an aching pain" (Repetitive)
	tool-	"sideways"	toolek	"get caught (rope, etc.)"

It should be clear at this point that the verbs in (20) not only share general semantic characteristics but belong to a derivational subsystem governed by a single argument structure scheme. The subjects of such verbs have feelings or thoughts of which their initial element is predicated, and within this subsystem the Intensive yields the semantic effect of “intensification”. We suggest that this is because the Intensive quantifies over the implicit thoughts or feelings expressed by the medial suffix. In effect, the Intensive verbs in (20) mean “have a *lot* of tired thoughts”, “a *lot* of bad thoughts”, and so on. Thus, as shown very schematically in (24), the verb *kaam-oksim-* “dislike” in (20b) involves an Experiencer x who dislikes, a Theme or Cause argument y , and, we suggest, also an Experience argument z . It is this element, the Experience argument, which is pluralized by the Intensive.

- (24) *kaamoksim-*: $\exists s$ & Experiencer(s,x) & Theme(s,y) & Experience(s,z) & *kaam*(z)
kegaamoksim- “really dislike”: $\exists s$ & ... & Experience(s,z) & *kaam*(z) & **PLUR**(z)

To summarize, the Yurok Intensive seems to express actual semantic “intensification” only in two kinds of context. First, with negatives the process of NEG-raising produces an interpretation “never”, which Robins viewed as an intensification of the meaning “not”. Second, complex stative predicates may have internal elements that are subject to Intensive quantification. In some cases, by saturation of an internal element, this may produce the effect of “intensification”. The Intensive by itself does not express an “intensive” meaning *per se*, however, and examples that seem to be “intensive” do not seem to be found with the full range of stative predicates. For example, we have not yet found apparent “intensification” with the common and morphologically underived class of verbs like “big” or “long”, or with color predicates.

4. The double Intensive. A final question we will consider is the semantic effect of double Intensive formation of the type shown in (2) above. In his grammar Robins wrote only that the double formation expresses “further intensification” and is “not ... as freely productive” as the ordinary Intensive (1958: 84), but he later suggested that double Intensives may convey the meaning “always” as opposed to “often”, comparing *begi* “it is often said” and *hiigi* “it is always said” (Robins 1985a: 640). While the examples in (5) show that the single Intensive can mean “habitually” or “always”, the double Intensive always has this meaning. In other words, the single Intensive, like the plain unfixed verb, is compatible with a translation “always” but does not by itself express that sense.

Isolated glosses are the least probative class of evidence, but for what it is worth we note that at least the two double Intensive verbs in (25) are glossed with habitual adverbial expressions.

- (25) a. kemoloc- “be jealous” *kiigemoloc-* “be jealous by nature”
 b. na’aw- “catch surf fish” *niiga’aw-* “catch surf fish as an occupation”

In (25b) a gloss like “catch a *lot* of surf fish” would in principle have been possible, but it was not chosen. To be sure, the more common pattern in isolated glosses is that this effect is absent, and that single and double Intensives are listed with no distinction noted. Here we pursue the question using textual examples of the five additional verbs in (26a). The verbs mentioned in (26b) also appear to pattern like those in (26a), but for reasons of time or analytic complexity we will to discuss them here.

(26)	BASE VERB	INTENSIVE	DOUBLE INTENSIVE	BASE GLOSS
a.	ke'yonem- (29a)	[no example]	kiige'yonem- (29b)	"release"
	raayo'r-	regaayo'r- (28a)	riigaayo'r- (28b)	"run past"
	syo'oogec-	[no example]	syiigo'oogec- (30b)	"make noise, sing"
	terum- (31a)	tegerum- (31b)	tiigrwm- (31c-d)	"speak"
	tmool-	tmegool- (27a)	tmiigool- (27b)	"shoot"
b.	cwin(kep-)	cwegin(kep-) (R 166)	cwiigin(kep-) (R 282)	"speak"
	nr'rmrh	nrgr'rmrh- (K 74.4)	niigr'rmrh- (K 74.5)	"snare"
	rohsim-	[no example]	riigohsim- (R 142)	"throw"

We begin in (27) with Intensives based on the common verb *tmool-* "shoot".

- (27) a. si **tmegool-i-'**
 IRR **shoot.INT-PASS-3SG**
 "He came near being shot more than once."
- b. si mehl **tmiigool-i-'** mehl hookckehl
 IRR CIRC **shoot.INT.INT-PASS-3SG** with arrow-flaker
 "It would have been shot at with arrow-flakers [had Wohpekumeu not spoiled it]."

The examples are comparable because both are passive and both occur with the irrealis particle *si*. If we translate the passives as actives with generic subjects, then the example in (27a) means "it nearly happened that they shot him more than once" and the example in (27b) means "it would have happened that they could shoot at it with arrow-flakers". The shooting in (27a) is episodic and repeated, while in (27b) the shooting would have been perpetual and habitual. In (27b) "often" would not add the intended meaning, and likewise in (27a) "always" could not be used.

In (28) we cite Intensives based on another common verb, *raayo'r-* "run past" (*raay-* "pass" + *-o'r-* "run")

- (28) a. co hl'o'ronep-e'm k'i kwen co ko **regaayo'r-ep-e'm**
 IMPV stop.moving-2SG wherever PTCL **run.past.INT-REFL-2SG**
 "Stop at each place you pass (and wherever there is a fishing rock leave some of your scales there.)"
- b. tu' wi'iit **riigaayo'r** ku wrgrs
 and it **run.past.INT.INT.3SG** the fox
 "A fox used to cross over it [a log]."

The simple Intensive in (28a) is used to express repeated running past on a single occasion, while the double Intensive in (28b) has habitual force.⁴

A third double Intensive example, based on *ke'yonem-* "release", is cited in (29b).

- (29) a. nes-ek' ti'n co ki mehl **ke'yonem-e'm** k'e-kap
 1.think-SG FUT CIRC **release-2SG** 2-medicine
 "I thought you would let me have your medicine."
- b. kelew kwelekwit cpi numi mehl **kiige'yonem-e'm** k'e-kap
 you them always very CIRC **release.INT.INT-2SG** 2-medicine
 "You always let them have your medicine."

A contrast is intended between (29a) and (29b), which occur in close proximity in a single text. In (29b), the sense is universal, habitual, or characterizing. Note that we cannot contrast the double Intensive in (29b) with a single Intensive, which is not attested in the text.

In (30) we cite another verb (*yo'oogec-* “make noise, sing”) whose simple Intensive we have not yet found attested textually.

(30) a. “After a time he [the small salmon] knew that they [the other fishes] all slept, so he began to sing. Then one lying at *mehleu* [“the corner farthest from both entrance and exit] said, ‘Ah, you have begun to sing! You are always doing that; I do not like it.’ It was Sturgeon, Kohko, who said this, for he wished to sleep.”

b. k'e-**syiigo'oogic-k'** neki' kito ckey-ik'
 2-make.noise.**INT.INT**-SG I want sleep-1SG
 “I don't like this singing [*literally* ‘You always make noise’], I want to sleep.”

The quotations in (30a-b) represent two versions of the same text, one recorded in English and one in Yurok. Sturgeon is quoted as saying “You are always doing that” in the more expansive text in (30a), which makes it clear that the real sense in (40b) is “You are always singing”.

Finally, in (31) we present a more uncertain example (*terum-* “speak”).

(31) a. noohl 'o **teru'm**
 then LOC **talk.3SG**
 “(Then he goes down to the river.) Then he talks; (first he kicks it where they will build the dam; then there he speaks across the river)”

b. ki nowor noohl**tegerum** ki 'wes'onah
 FUT be.light then **talk.INT** the sky
 “(As long as they are at work on the dam, so long you will not sleep.) All night long you will talk to the world. (It will always be so.)”

c. 'o ko hegol-e'm wistu' ki soo **tijgrwrmr**
 LOC TEMP tell-2SG and so FUT thus **speak.INT.INT**
 “Then you instructed me how to pray.”

d. nekah co **tijgrwrmaari'n-e'm**
 us IMPV **speak.INT.INT**-2SG
 “You must speak to us (when you are afraid of this).”

The verb *terum-* was known to Robins only in its Intensive form, but a full dossier of related forms can be displayed from older sources. The simple form in (31a) is used in a context where speaking is merely one of a series of actions. Within the same text, the Intensive form is used to express repeated talking all night long in (31b). The double Intensives in (31c-d) unfortunately show additional derivational morphology, but the underlying root is the same and the context in both examples is habitual or characterizing.⁵

To summarize this section, we have added several examples to a dossier of evidence supporting the view that double Intensives may express characteristics, habits, or situations that are repeated “always”, as opposed to situations that are simply repeated. Simple event-external repetition is expressed by the single Intensive form.

Our interpretation of the Yurok double Intensive may help answer a problem of mor-

phology and semantics in an entirely unrelated language. Like Yurok, Latin has a pair of distinct morphological devices for expressing event-internal and event-external repetition: an Iterative suffix expressing event-internal repetition; and a Frequentative suffix expressing event-external, multiple-occasion repetition (Garrett & Slatin in preparation). Typical Frequentatives include *canitāre* “sing repeatedly” (*canere* “sing”) and *scripitiāre* “be in the habit of writing” (*scribere* “write”); Iteratives include *conflictāre* “strike frequently, buffet” (*confligere* “collide”), *dūctāre* “lead around” (*dūcere* “lead”), *clāmitāre* “shout repeatedly” (*clāmāre* “shout”), and *volitāre* “fly about” (*volāre* “fly”). As these forms show, the Latin Iterative suffix has the two productive allomorphs *-t-* and *-it-*. Their distribution and history need not be discussed here; what is crucial is that the Frequentative suffix *-tit-* appears to be compounded from the two Iterative suffixes. Historically, as sketched in (32), the Latin Frequentative seems to be the Iterative of an Iterative.

- (32) HYPOTHETICAL DERIVATION OF THE LATIN FREQUENTATIVE
- | | | |
|--------------|------------------------------------|------------------------------------|
| veh-ere | “carry” | (cf. tmool- “shoot”) |
| vec-t-āre | “carry around” (Iterative) | (cf. tmegool- “shoot often”) |
| vec-t-it-āre | “carry habitually” (Frequentative) | (cf. tmiigool- “shoot habitually”) |

The question raised by this analysis is like the question posed by the Yurok double Intensive: why precisely does doubling a repetition marker produce a multiple-occasion repetition marker?

Our preliminary answer, at least for Yurok, is that habitual readings may emerge from iteration of repeated-aspect markers through a form of pragmatic strengthening. The literal meaning of a double Intensive should be that a type of event happens repeatedly over multiple periods of observation in which it might happen. In such circumstances it is likely that the event type generally happens with some consistency: it is habitual or typical. We suggest that the development of a habitual sense in double Intensives may represent a grammaticalization of this pragmatic association.

5. Conclusion. We have argued that the Yurok Intensive is an event-external repetition marker, and have shown how its semantics interact with verbs of different aspectual classes and with negation, and in the double Intensive, to produce an apparent variety of meanings. We have also suggested some ways in which our findings relate to the behavior and description of pluractional markers in other languages.

There are some questions still to be resolved in future research, and other areas we plan to investigate. These include the interaction of the intensive with other aspectual markers, such as the perfect; the semantics of deverbal nouns with single and double Intensive marking; and more explicit characterization of the semantics of the Intensive marker – in particular, considering whether all of the uses of the Intensive can be accounted for within a single general meaning.

Appendix: Sources of Examples

Field notes are cited by notebook and page; e.g. K 40.20 = Kroeber notebook 40, p. 20.

5a: T. T. Waterman Yurok field notebook A, p. 106 (cf. Waterman 1923: 376)

5b: Domingo of Weitchpec, “Buzzard’s Medicine”, 1907 (K 75.26); tr. Kroeber (1976: 313)

5c-d: R 107, 57

6a: Domingo of Weitchpec, “Turip Young Man and His Dogs”, 1906 (K 67.12); tr. Kroeber (1976: 310)

6b: Mary Marshall, Death purification medicine, 1906 (K 66.19)

- 7: **R** 110
 10: Domingo of Weitchpec, "Buzzard's Medicine", 1907 (**K** 75.25); tr. Kroeber (1976: 313)
 11: Mary Marshall, "Coyote Tries to Kill the Sun", 1927 (Sapir 2001)
 14a: Pekwan Doctor, "Gambling with Kapuloyo", 1907 (T. T. Waterman Yurok field notebook 83, p. 18); cf. Kroeber (1976: 385)
 14b: **R** 127
 14c: Florence Shaughnessy, "The First Salmon Rite at Wehlkwew", 1951 (**R** 172-73)
 15a: Mabel Brantner, "Wohpekumew and the Salmon", 1951 (**R** 162-63)
 15b: Captain Spott, Medicine song against rough water, 1906 (**K** 67.47)
 15c: T. T. Waterman field notebook B, p. 29
 16: Spott & Kroeber (1942: 234)
 17a: Robert Spott apud Kroeber (1960: 996)
 17b: **R** 127
 19a: Juanita, Brush Dance medicine, 1907 (**K** 76.7)
 19b: Florence Shaughnessy, "The Young Man from Serper", 1951 (**R** 168-69)
 27a: Waterman (1923: 381)
 27b: Captain Spott, "The Obsidian Cliff at Rekwoi", 1907 (**K** 76.6, 8), tr. Kroeber (1976: 436)
 28a: Florence Shaughnessy, "The First Salmon Rite at Wehlkwew", 1951 (**R** 176-77)
 28b: Florence Shaughnessy, "The Fox and the Coon", 1951 (**R** 164-65)
 29: Juanita, Brush Dance medicine, 1907 (**K** 77.16-17)
 30a: Billy Werk, "A Formula for Salmon", 1907, Kroeber (1976: 250)
 30b: Billy Werk, 'A Formula for Salmon', 1907 (T. T. Waterman Yurok field notebook 84, p. 18)
 31a-b: Mary Marshall, Kepel Fish Dam formula, 1906 (**K** 66.4, 12), ed. Waterman & Kroeber (1938: 75, 76)
 31c: Spott & Kroeber (1942: 240)
 31d: Stone of Weitchpec, Weitchpec Deerskin Dance formula, 1906 (**K** 78.24-32), ed. Buckley (1982: 393)

Notes

* Our work is supported by National Science Foundation grant BCS-0004081 to the University of California, Berkeley (<http://linguistics.berkeley.edu/~yurok>). We use the following non-obvious abbreviations:

ATTR = attributive mood

CIRC / IRR / LOC / TEMP = circumstantial / irrealis / "locative" / temporal preverbal particle

K = field notes of A. L. Kroeber

R = Robins (1958)

Otherwise unattributed data are cited from the lexicon of Robins 1958. We use a transcription scheme identical to that of Robins except that doubling indicates vowel length, apostrophes replace glottal stops, /k^w/ is written *kw*, /t/ is written *h*, and the vowel /ɪ/ is written *r*.

¹ So Waterman (1923: 380); see also Spott & Kroeber (1942: 216), Kroeber (1976: 312, 405), and Robins (1985a: 640, 1985b: 728). Waterman also used the term "frequentative" throughout his unpublished field notes.

² Strictly speaking, Bybee, Perkins & Pagliuca (1994: 127) define "iterative" vs. "frequentative" aspect as repetition on a single occasion vs. multiple occasions; i.e. they contrast (6a-b) vs. (6c).

³ For the medial elements *-pey* and *-sew* see Proulx (1985: 131, 136); for *-owohl-* cf. *ckenowohl-* "talk maliciously" (*cken-* 'be scarce, be few'). For *-onihl-*, "cost" or "pay" seems better justified than "expensive" by the actual data cited by Proulx (1985: 121).

⁴ The example in (38b) is not unlike (5a), an ordinary Intensive which occurs in a habitual context. What is crucial is that *all* examples of the double Intensive have this special extra reading and that the single Intensive only has it sometimes (when something *other* than its Intensiveness prompts it).

⁵ The form *tiigrwrm-* "speak, greet" is clearly the Intensive of **trwrm-*, which is in turn clearly related to *terwm-*. Apart from *r*-coloring diminutivization, the details of the derivation are unclear to us now, as are the further derivational histories of *tiigrwrm-r* and *tiigrwrm-aari'n-*.

Archaeology and Ethnology 35: 143-256.

Waterman, T. T. 1920. Yurok geography. University of California Publications in American Archaeology and Ethnology 16: 177-314.

Esther Wood & Andrew Garrett

125

References

- Berman, Howard. 1982. A supplement to Robin's Yurok-English Lexicon. *International Journal of American Linguistics* 48: 197-241.
- Buckley, Thomas. 1980. Monsters and the quest for balance in native northwest California. *Manlike monsters on trial: Early records and modern evidence*, ed. by Marjorie M. Halpin and Michael M. Ames, pp. 152-71. Vancouver: University of British Columbia Press.
- . 1982. Yurok realities in the nineteenth and twentieth centuries. Ph.D. dissertation, University of Chicago.
- Bybee, Joan; Revere Perkins; and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Cusic, David Dowell. 1981. Verbal plurality and aspect. Ph.D. dissertation, Stanford University.
- Garrett, Andrew. 2001. Reduplication and infixation in Yurok: Morphology, semantics, and diachrony. *International Journal of American Linguistics* 67: 264-312.
- Garrett, Andrew, and Patricia Slatin. In preparation. Repetition in Latin grammar: Form, function, and evolution. University of California, Berkeley.
- Hardy, Heather K., and Timothy R. Montler. 1988. Alabama radical morphology: H-infix and disfixation. In honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics, ed. by William Shipley, pp. 377-409. Berlin: Mouton de Gruyter.
- Horn, Lawrence R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
- Kimball, Geoffrey D. 1991. *Koasati grammar*. Lincoln, Neb.: University of Nebraska Press.
- Kroeber, A. L. 1869-1972. A. L. Kroeber papers. Manuscript collection (MSS C-B 925), Bancroft Library, University of California, Berkeley. [Cited from microfilm copy (FILM 2049).]
- . 1911. The languages of California north of San Francisco. *University of California Publications in American Archaeology and Ethnology* 9: 414-26.
- . 1960. Yurok speech usages. *Culture in history: Essays in honor of Paul Radin*, ed. by Stanley Diamond, pp. 993-99. New York: Columbia University Press.
- . 1976. *Yurok myths*. Berkeley: University of California Press.
- Lasersohn, Peter. 1995. *Plurality, conjunction and events*. Dordrecht: Kluwer Academic Publishers.
- Mithun, Marianne. 1988a. Lexical categories and the evolution of number marking. *Theoretical morphology: Approaches in modern linguistics*, ed. by Michael Hammond and Michael Noonan, pp. 211-34. San Diego: Academic Press.
- . 1988b. Lexical categories and number in Central Pomo. In honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics, ed. by William Shipley, pp. 517-37. Berlin: Mouton de Gruyter.
- Newman, Paul. 1990. *Nominal and verbal plurality in Chadic*. Providence, R.I.: Dordrecht.
- Proulx, Paul. 1985. Notes on Yurok derivation. *Kansas Working Papers in Linguistics* 10/2: 101-143.
- Robins, R. H. 1958. *The Yurok language: Grammar, texts, lexicon*. (University of California Publications in Linguistics, 15.) Berkeley: University of California Press.
- . 1985a. The young man from Serper: A Yurok folktale. *Collectanea philologica: Festschrift für Helmut Gipper*, ed. by Günter Heinz and Peter Schmitter, vol. 2, pp. 633-44. Baden-Baden: Valentin Koerner.
- . 1985b. Numerals as underlying verbs: The case of Yurok. *Studia linguistica diachronica et synchronica Werner Winter sexagenario anno MCMSXXXIII*, ed. by Ursula Pieper and Gerhard Stickel, pp. 723-34. Berlin: Mouton de Gruyter.
- Sapir, Edward. 1922. The Takelma language of southwestern Oregon. *Handbook of American Indian languages*, ed. by Franz Boas, vol. 2, pp. 1-296. Washington: Government Printing Office.
- . 1930. Southern Paiute, a Shoshonean language. *Proceedings of the American Academy of Arts and Sciences* 65: 1-296.
- . 2001. Yurok texts, ed. by Howard Berman. *Collected works of Edward Sapir*, vol. 14, Northwest California linguistics, ed. by Victor K. Golla and Sean O'Neill, pp. 1015-38. Berlin: Mouton de Gruyter.
- Spott, Robert, and A. L. Kroeber. 1942. *Yurok narratives*. University of California Publications in American

- Archaeology and Ethnology 35: 143-256.
- Waterman, T. T. 1920. Yurok geography. University of California Publications in American Archaeology and Ethnology 16: 177-314.
- . 1923. Yurok affixes. University of California Publications in American Archaeology and Ethnology 20: 369-86.
- . n.d. Yurok field notes. Item 99, Ethnological Documents Collection of the Department and Museum of Anthropology, University of California, Berkeley. Manuscript Collection (CU 23-1), University of California Archives, Bancroft Library, University of California, Berkeley. [Cited from microfilm copy (FILM 2216, reel 99(1-3)).]
- Waterman, T. T., and A. L. Kroeber. 1938. The Kepel fish dam. University of California Publications in American Archaeology and Ethnology 35: 49-80.
- Xrakovskij, Viktor S. ed. 1997. Typology of iterative constructions. (Lincom Studies in Theoretical Linguistics, 4.) München: Lincom Europa.
- Yu, Alan C. L. 2001. Pluractionality in Chechen. To appear in CLS 37. Chicago: Chicago Linguistic Society.

Department of Linguistics
 1203 Dwinelle Hall #2650
 University of California
 Berkeley, CA 94720-2650

ejwood@socrates.berkeley.edu
garrett@socrates.berkeley.edu