

---

# Santa Barbara Papers in Linguistics

---

VOLUME 15

Proceedings from the seventh  
Workshop on American Indigenous Languages

April 30-May 2, 2004

Lea Harper and Carmen Jany,  
Editors

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

## FOREWARD

It is with great pleasure that we present the proceedings of the seventh Workshop on American Indigenous Languages (WAIL 2004). In continuing a tradition begun with the student discussion group on Native American Indian 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 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.

Jeanie Castillo

Carmen Jany

Salomé Gutierrez

Kirk Miller

Lea Harper

Petra Shenk

Dan Hintz

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 Discourses 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
Volume 12:	<u>Recent Studies in Empirical Approaches to Language</u>	\$15.00
Volume 13:	<u>Proceedings from the fifth Workshop on American Indigenous Languages</u>	\$15.00
Volume 14:	<u>Proceedings from the sixth Workshop on American Indigenous Languages</u>	\$15.00
Volume 15:	<u>Proceedings from the seventh Workshop on American Indigenous Languages</u>	\$15.00

# CONTENTS

FOREWARD .....	1
LACANDON COLOUR TERMS .....	3
<i>Suzanne Cook</i>	
NOMINALIZATION AS A QUESTION FORMATION STRATEGY IN TUCANOAN .....	9
<i>Dmitry Idiatov and Johan van der Auwera</i>	
EXTERNAL POSSESSION IN HUEHUETLA TEPEHUA .....	25
<i>Susan Smythe Kung</i>	
SOME TONOGENETIC PROPERTIES OF UPRIVER HALKOMELEM .....	40
<i>J. C. Brown</i>	
VOICE IN IQUITO: IS THERE AN INVERSE? .....	49
<i>Lynda de Jong Boudreault</i>	
A PHONOLOGICAL SKETCH OF THE YUCUNANY DIALECT OF MIXTEPEC MIXTEC .....	61
<i>Mary Paster and Rosemary Beam de Azcona</i>	
GRAMMATIZATION OF TENSE IN NAVAJO: THE EVOLUTION OF <i>NT'ÉÉ</i> . .....	76
<i>Melvatha Chee, Evan Ashworth, Susan Buescher, Brittany Kubacki</i>	
INTRODUCING SAN AGUSTIN MIXTEPEC ZAPOTEC. ....	91
<i>Rosemary G. Beam de Azcona</i>	
THEORETICAL ISSUES IN THE REPRESENTATION OF THE GLOTTAL STOP IN BLACKFOOT .....	106
<i>Tyler Peterson</i>	



## Lacandon Colour Terms

Suzanne Cook  
University of Victoria

### Introduction

This paper considers colour terms in Lacandon, a Yucatec language spoken in the eastern state of Chiapas, Mexico. I will show how it uses morphological processes similar to those found in the other Mayan languages to extend its distinctive set of basic colour terms: black, white, red, yellow, and blue/green. Basic colour terms are lexemes that cannot be analysed into smaller morphemes, do not have compositional meanings, are psychologically salient, and are applicable to a wide range of objects (Berlin and Kay 1969). These terms are often discovered using the stimuli provided by Munsell colour chips. The Munsell framework limits the number and range of sensory attributes to only three dimensions of contrast: hue, brightness and saturation. If it is used as a tool for colour term elicitation it cannot help to discover color categories that discriminate degrees of brightness and saturation, or other sensory dimensions such as texture, transparency, lustre, illumination, location, duration, and fluctuation. As well, the lexemes that encode these other sensory dimensions are typically polymorphemic, so such colour terms are excluded from the basic colour system.

### Colour terms in Maya languages

Colour is an important focus in Mayan cultures. For example, in Tzeltal, Hunn (1977:89) found that the five basic colour terms figure in 217 of 586 animal naming responses of composite forms he recorded, e.g., *cahal nuk' toht* [*cahal* “red” + *nuk'* “neck” toht “robin”] (89). It has been shown that Yucatec, Itzaj, Mopán, and Tzotzil (in Bricker 1999, Hofling and Tesucún 1997, Ulrich and Ulrich 1976, and Haviland 1981) extend the basic Mayan colour palette to include terms that express intensity, texture, opacity, relative size, discreteness, and texture. These languages use compounding of colour terms with other roots and the affixation of special colour suffixes to create words that make these distinctions. Some of these languages use reduplication to express degrees of saturation. For example:

Yuc. *k'áan-hep'-é7en* “deep yellow” < *k'áan* “yellow” + *hep'* “tighten, squeeze” + *-é7en* (Bricker 1999: 285)

Itz. *k'än-jop'-e7en* “bright yellow (flame of fire)” < *k'än* “yellow” + *jop'* “light” + *-e7en* (Hofling p.c., cited in Bricker 1999:290)

Mop. *chun-tuj-e7en* “yellowish, golden” < *chun* “yellow” + *tuj* “unspec. meaning” + *-e7en* (Ulrich and Ulrich 1976:79)

Tzo. *k'an-kep-an* “yellow (bunch of bananas)” < *k'an* “yellow” + *kep* “bunch” + *-an* (Haviland 1975:172)

### Lacandon

Lacandon uses similar morphological processes to extend its basic colour palette to include terms that express degrees of saturation and luminosity, translucence, texture, shape, and location or position.

The consultants used in this study were Sakhol Garcia Paniagua, male, aged 23 and Marta Trujillo Gonzales, female, aged 20. Both are bilingual in Spanish and Lacandon.

### The system of colour categorization in Lacandon

#### Basic colour terms

chak~chäk “red”

k'än “yellow”

yaax “blue/green”

7ek' “black”

sak~säk “white”

#### Reduplicated colour terms (express degree of saturation)

Very saturated: chäk-chäk “red-red, very red” (note: not as red as ne chäk “very red”)

Less saturated: chä-chäk “sort of red”

The same CVC- and CV- patterns occur with the other basic colour roots.

#### Compounds (main colour term + lexical root + -e7en colour adjective)<sup>1</sup>

ch'ay “hanging?” (cf. ch'uy “hang”)

chäk=ch'ay-e7 “red (fruit on tree: e.g. hanging)”

yaax=ch'ay-e7 “green (fruit on tree)”

hay “spread out, extended; thin (transparent?)”

chäk=hay-e7 “red (fruits) spread out (on the ground)”

säk=hay-e7 “clear (sky)”

yaax=hay-e7n “s.t. thin green”

k'än=hay-e7 “yellow (ripe fruit)”

chäk=hay-e7 “s.t. red (spread out on the ground, like tomatoes)”

7ek'=hay-e7 “black (seeds) spread out (on the ground, like chankala7 ‘false ginger’ seeds)”

säk=hay-e7 “s.t. white spread out (like the vomit of a sick child)”

**k'äl “lock, close”**

chäk=k'äl-e7 “red inside the bottom of the eye lid”

yaax=k'äl-e7 “all blue sky (with no clouds)”

**pos “pale, colourless”**

yaax=pos-e7 “pale blue” (B25 on Munsell colour chart)

säk=pos-e7 “off-white”(C on Munsell colour chart)

**puk’ “dissolve, mix”**

yaax-puk'-e7 “blue-purple”(E32 on Munsell colour chart)

**til “dull?” (cf. til [used in colour compounds to express dull colours] (Bricker 1998:276))**

chak=til-e7 (D6, D36-40) [deep coral, pink moving toward orange]

yaax=til-e7 (C20, C21) [aquamarine]

**t'in “shine, reflect” (refers to luminous colour)**

sak=t'in-e7 “brilliant, shimmer (?)”

chak=t'in-e7 “shiny red; luminescent red”

7ek'=t'in-e7 “shiny black, glossy black”

yaax k'En=t'in-e7 [shiny green yellow sample]<sup>2</sup>

yaax sak=t'in-e7 “brilliant green”

k'än=t'in-e7 “luminescent yellow”

7ek' sak=t'in-e7 “brilliant black”

7ek'=t'in-e7 “luminescent black”

sak-sak=t'in-e7 “brilliant (luminescent) white”

yaax=t'in-e7 “luminous green (like fire light, or the colour of spirits)”

**yul “smooth”**

chäk=yul-e7 “smooth red (like the ground where a forest once stood)”

ne yaax=yul-e7 [dark green smooth sample]

yaax=yul-e7 [dark green felt sample]

**sum “hang?” (cf. sum “rope” (Hofling 1997:568))**

7ek'=sum-e7 “s.t. hanging black (e.g. fruit)”

yaax=sum-e7 “s.t. hanging green (e.g. fruit)”

**hal “bright, glow, blaze”**

chak=hal-e7 “red-hot coals”

yaax=hal-e7 “glowing green (like the colour of the cross-section of a mineral)”

yaax=hal-e7 “luminescent green”

**hup “light; when the fire starts to burn, grow”**

chak=hup-e7 “red (clouds at sunrise/sunset)”

k’än=hup-e7 “yellow (clouds at sunset/sunrise)”

säk=hup-e7 “brilliant white (of a rock face)”

**yub “?”[medium gray tone]**

7ek’yub-e7 “not very black” (G on Munsell chart)

chak=yub-e7 “not very red” (E1) [medium-tone orange-red]

yaax=yub-e7 “not very green”(G20, C21) [medium green, light aquamarine]

chak ek’yub-e7 “medium black red”(I1) [burnt sienna]

k’än=yub-e7 “dark yellow” (D9, D10) [mustard]

sak=yub-e7 “dark white” (B)

sak ek’yub-e7 “darker white” (F)

**pal “lacking colour?” (cf. pil [indicating lighter shades] (Bricker 1998:216))**

sak=pal-e7 “pale (face)”

yaax sak=pal-en [light gray-green sample]

**wol “round, circle”**

yaax=wol-e7 “black eye, bruise”

**lah “all, completely”**

chak=lah-e7 “red (rash)”

**ul “small and clustered (on a bush)?” (cf. ul “fatten” (Hofling 1997:651))**

chäk=ul-e7 “red all around, bunches of red (fruits), e.g., tomatoes on a bush”

**ch’ul “wet”**

yaax=ch’ul-e7 “wet green”

**Comparative data.** Lacandon has colour compounds related to compounds in other Yucatecan languages and Tzotzil, a Cholan-Tzeltalan language. However, the specified colors may not be the same.

**Lacandon**

chak=hup-e7 “red (clouds)”

chak=hal-e7 “red-hot (coals)”

ya7ax=hal-e7 “luminescent green”

**Itzaj**

chäk=jop-e’en “red (flames of fire, clouds)”

chäk=jol-e’en “red (flames of fire, clouds)”

ya’ax=jol-e’en “blue,green (flames of fire), green (clouds)”

**Lacandon**

sak=pal-e7 “pale (face)”

ya7ax=puk'-e7 “blue-purple” (E32)

ya7ax=til-e7 (C20-21)

chak=til=e7 “almost red; not v. sat.”

chak=yul-e7 “smooth red”

sak=pos-e7 “off-white” (C)

ya7ax=pos-e7 “pale blue”(B25)

**Yucatec**

sak=pil-é7en “gray, pale (person from fright)”

yá7ax=puk'-é7en “faded green”

yá7ax=til-é7en “olive green”

chak=til-é7en “dull red”

chak=yul-é7en “raw red”

sak=pos-é7en “whitish”

yá7ax=pos-é7en “gray-green”

**Lacandon**

ya7ax=puk'-e7 “blue-purple” (E32)

k'an=hay-e7 “yellow (ripe fruit)”

**Tzotzil**

yax=puk'-an “clear (water)”

k'an=hay-an “transparently yellow”

There are lexical roots in Lacandon that cannot be compounded with colour, but which can enter into colour compounds in Yucatec and/or Itzaj:

ch'uy “suspend”

hats' “hit, beat, whip”

hep' “tight at waist”

chup' “swell”

p'uch “thrash, beat”

7el “burn”

p'ox “break out in a rash, raised bumps on the skin (not red)”

mäk' “lick, eat fruit”

**Conclusions**

Mayan languages use special morphological processes to extend basic colour categories. Lacandon uses these processes, extending its colour palette to encode shape, texture, translucence, relative size and discreteness of objects, and position. The encoding of these perceptual dimensions in colour terminology is productive, identifying categories with contrasting features of colour, other than brightness, hue, and saturation. Luminosity, for example, is an important dimension in Mayan colour categorization. This and other dimensions could never be discovered using only the Munsell colour system as an elicitation tool.

When comparing related data from different languages, the colours identified are often quite different even when the morphemes are the same. That is, the systems for creating descriptive colour terms match, but the referent colours often do not.

Because similar morphological processes of colour identification and classification exist in the Yucatecan languages and in Tzotzil, which is a Cholantzeltalan language, this must be an old pattern in Mayan.

More research needs to be carried out on colour terms in Lacandon, to determine the number and range of lexical roots involved in colour categorization. This will uncover the features that are important in the Lacandon color system.

### Notes

<sup>1</sup> Lexical roots are from a number of categories such as N, V, A.

<sup>2</sup> This and following colours that include the term **sample** were elicited with the help of color fabric samples that uncover sensory attributes outside the Munsell system.

### References

- Berlin, Brent, and Paul Kay. 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley and Los Angeles: University of California Press.
- Bricker, Victoria R. 1999. Color and Texture in the Maya Language of Yucatan. *Anthropological Linguistics* 41, 3: 283-307.
- Haviland, John B. 1981. *Sk'óp Sotz'leb: El Tzotzil de San Lorenzo Zinacantán*. México: Universidad Nacional Autónoma de México.
- Hofling, Charles Andrew, and Félix Fernando Tesucún. 1997. *Itzaj Maya-Spanish-English Dictionary/Diccionario Maya Itzaj-Español-Ingles*. Salt Lake City: University of Utah Press.
- Hunn, Eugene S. 1997. *Tzeltal Folk Zoology: The Classification of Discontinuities in Nature*. New York: Academic Press.
- Laughlin, Robert M. 1975. *The Great Tzotzil Dictionary of San Lorenzo Zinacantan*. Smithsonian Contributions to Anthropology No.9. Washington, D.C.: Smithsonian Institution.
- Ulrich, E. Matthew, and Rosemary Dixon de Ulrich. 1976. *Diccionario Bilingue: Mopán y Espanol/Español y Maya Mopán*. Guatemala: Instituto Lingüístico de Verano.

# Nominalization as a question formation strategy in Tucanoan<sup>1</sup>

Dmitry Idiatov & Johan van der Auwera  
Center for Grammar, Cognition, and Typology  
University of Antwerp

## 1. Introduction

This paper examines a typologically unusual question formation strategy, found in the Tucanoan languages of South America. Both polar questions and information questions are taken into consideration, but the focus is on the former. In the following discussion we will show that these languages exhibit a historical and semantic relationship between nominalization (or “deverbalization”) and question formation strategy. We will argue that the latter evolved from the former.

The Tucanoan language family is traditionally subdivided into two groups, Eastern and Western. The Eastern Tucanoan languages are spoken in the north-western Amazon Basin in the state of Vaupés, Colombia, and in the state of Amazonas, Brazil. The Western Tucanoan languages are spoken in south-western Colombia and in the neighbouring regions of Ecuador and Peru. Grammatical descriptions of most of the Tucanoan languages of both the Eastern and Western branches have been consulted for the present study. However, the discussion will be mainly confined to data on the following languages: Kaye (1970) and Miller (1999) on Desano (Eastern), Morse & Maxwell (1999) and Ferguson et al. (2000) on Cubeo (Eastern), Kinch & Kinch (2000) on Yurutí (Eastern), Johnson & Levinsohn (1990) on Secoya (Western), Cook & Criswell (1993) and Gralow (1993) on Koreguaje (Western).

One cross-linguistically rather unusual feature in Tucanoan verb-subject agreement should be mentioned. Besides the ordinary agreement of the verb in gender and number with third person animate subjects following a three-way distinction between masculine, feminine, and plural, there is a fourth subject agreement morpheme, which is used for agreement with singular inanimates and also for agreement with first and second person singular or plural (Miller 1999:64; Morse & Maxwell 1999:40; Barnes 1999:214). This subject agreement morpheme is generally called the “non-third person” in descriptions of the Tucanoan languages, implying that the subject can be everything except the third person animate (sometimes except the third person plural as well).

## 2. Question formation & nominalization

### 2.1. Eastern Tucanoan

#### 2.1.1. Desano

Polar questions are marked in Desano by the interrogative marker *-ri*, which functions as a suffix that occurs at the end of the main verb and replaces the person, number, and gender subject agreement suffixes that would have occurred there. It follows the suffixes marking evidentiality and tense-aspect.

- (1)     *yí*     *tígõ*           *gahi-bākã-ge*           *ârĩ-ku-bõ*  
1SG older.sister other-town-LOCATIVE be-ASSUM-3SG.F  
‘My sister is at another town’ (Miller 1999:49)<sup>2</sup>

- (2) *igo segundo curso ii-bāsĩ-ku-ri*  
 3SG.F second course do-ABILITATIVE-ASSUM-Q  
 ‘Can/ May she do the second course?’ (Miller 1999:129)

Information questions are formed by using the interrogative pronouns, which either begin the sentence or directly follow the topic, and by the interrogative marker *-ri* known from polar questions. As can be seen from the following examples, the subject pronoun can be omitted, no doubt when the context is clear enough, both in polar and information questions.

- (3) *dōã-re ã̃-yu-ri* (4) *ba-gĩ ii-ri*  
 who-SPC see-HSY-Q eat-SG.M do-Q  
 ‘Whom did they see?’ (Miller 1999:131) ‘Are you eating?’ (Miller 1999:33)

The interrogative suffix *-ri* is quasi-identical in form with the inanimate deverbaling suffix. They differ only in their nasalization pattern: the interrogative suffix seems to be inherently oral (see Kaye 1970:150, example 4a) whereas the deverbaling suffix seems to be unspecified for nasalization (see Kaye 1970:152, example 7a).

In this respect, it seems appropriate to make some general remarks on nasalization and nasal spreading in the Tucanoan languages before going on with the presentation of data. Normally, the Tucanoan languages have six phonemic vowels, all of which can be phonemically nasal or oral. However, some affix vowels are unspecified for nasality. They can be nasalized under the influence of intrinsically nasal vowels. Thus, in Cubeo “[n]asalization spreads from a nasal vowel to the right across morpheme boundaries onto vowels which are unmarked for nasality, providing only vowels unmarked for nasality and nasalizable consonants intervene. [...] Nasal spreading is blocked by an inherently oral suffix. [...] Nasal spreading is blocked by word boundaries” (Morse & Maxwell 1999:8).

It should also be mentioned that the nasalization pattern of a given suffix does not seem to be very stable cross-Tucanoan. On the one hand, the nominalizing suffix at issue (also called “deverbaling”, “infinitive” or “participle” marker in different grammars) is unspecified for nasalization, *-ri/ -rĩ* in Desano but is inherently oral in Barasano (Gómez-Imbert & Jones 2000). On the other hand, the interrogative marker seems to be inherently oral in Desano but unspecified for nasalization in Barasano (Jones & Jones 1991:14). In other words, the nasalization patterns of the nominalizing suffix and interrogative marker in Desano turn out to be the reverse of their Barasano cognates.

The inanimate deverbaling suffix is used for inanimate action nominalizations, which almost always occur with one or more arguments, and generally function as the object of the matrix clause, but it can also be the subject (see examples in Miller 1999:130, 143). It can also be used for inanimate patient nominalizations (5) and instrument nominalizations (6).

- (5) *gasiru kôã-bu-ri-ru* (6) *wi-ri-ru*  
 canoe throw.away-POTENTIAL-NOM-CLS fly-NOM-CLS  
 ‘the canoe to be thrown away’ (Miller 1999:144) ‘airplane’ (Miller 1999:144)

Suffixes of tense, mood, aspect, direction, and negation may occur between the verb root and the deverbaling suffix. However, just like the interrogative suffix *-ri*, the deverbaling suffix *-ri* replaces the subject agreement suffixes. One could therefore claim that an interrogative sentence is just as nonfinite as a nominalization. There seem to be only two structural differences between them. First, the verb phrase marked by the purely

deverbalizing *-ri* can be followed by a noun classifier, thus reinforcing its nominal function. Second, the verb phrase marked by interrogative *-ri* is normally characterized by the absence of any higher clause materials.<sup>3</sup>

### 2.1.2. *Cubeo*

In *Cubeo* polar questions and information questions are marked in the same way as in *Desano*, namely by means of an interrogative suffix on the main verb. The subject pronoun does not seem to be normally omitted. No information on the intonation patterns of either polar questions or information questions in *Cubeo* was found in the sources available.

In *Cubeo* several interrogative suffixes are said to exist. The choice between them is dependent on the person of the subject, the tense, and the aspect. However, according to Morse & Maxwell (1999:20), “there seems to be some variability, perhaps between generations, in the choice of these suffixes”. Below we reproduce a slightly adapted table from Morse & Maxwell (1999:21). It summarizes “the most common patterns” of the usage of interrogative suffixes.<sup>4</sup>

Person	Near Future	Indefinite Future	Present	Recent Past/ Durative, Nonrecent Past/ Habitual	All other tenses and aspects
1	<i>-xi</i>	?			
3 animate	<i>-RI</i>	<i>-bA</i>			
3 inanimate				<i>-RI</i>	
2	<i>-RI</i>			<i>-Rĩ</i>	

TABLE 1. Interrogative suffixes in *Cubeo*

Before proceeding to a discussion, some remarks on nasal spreading and “vowel harmony” in the Tucanoan languages should be made. As to nasal spreading, it was already partly discussed in section 2.1.1. Here, we would like to draw attention to the fact that besides the spreading of nasalization from left to right, which is very productive in all the Tucanoan languages, in some languages (e.g., *Desano*, *Barasano*, *Yurutí*) there is also a nasal spreading from right to left, even though it occurs only in a small closed set of constellations. Thus, certain *Barasano* suffixes acquire specification for nasality from four nasal person number suffixes on their right (Jones & Jones 1991:16). It is also worth reminding ourselves of the fact that the nasalization pattern of a given suffix does not seem to be very stable cross-Tucanoan.

According to Waltz & Wheeler (1972:173), “a tendency to vowel harmony is evident in Proto Tucanoan”, though reflexes of this process in the present-day Tucanoan languages are rather irregular. Judging from the available descriptions, vowel harmony does not seem to be a productive phonological process anymore, yet examples of distant vowel assimilation can still be found. In *Desano* (Miller 1999:18) there is a kind of “vowel harmony”, involving vowel lowering, affecting a limited number of verb suffixes. It is interesting to note that, first, this vowel assimilation is regressive, i.e., it acts from right to left contrary to the normal left-to-right direction of nasal spreading, and second, the suffixes instigating this vowel assimilation are the subject agreement suffixes.

We can now return to Table 1. In the second person, two variants, *-Rĩ* and *-RI*, are found. For most of the core tense/aspect constellations the form of the question marker in the second person is *-Rĩ*. Why should that be the case? The reason, we dare to propose, is

that when a question is asked about the interlocutor's actions the verb ending in a question marker is regularly followed by a second person subject pronoun, which is *bĩ* for the second person singular and *bĩxã* for the second person plural.

- (7) *xabo-ki-Rĩ* *bĩ*  
 rule-NONFUTURE.NOM.SG.M-Q 2SG  
 'Are you the chief/ headman?' (Morse & Maxwell 1999:23)

Thus, what we see here is that the form of the interrogative morpheme, *-Rĩ*, is a result of right-to-left nasal spreading and "vowel harmony". The absence of analogous assimilations for other persons might be accounted for by the pragmatic salience, in questions, of the second person interlocutor. The fact that this assimilation is not present in all of the second person tense aspect forms might be explained by the hypothesis that assimilation at least occurs in the highly frequent and possibly archaic core of the (second person) tense-aspect system.

As far as the interrogative markers *-xi* (first person + Near Future) and *-ba* (second person + Indefinite Future/ Present, first person + Present) are concerned, there seems to be some explanation for the fact why it is only in these tenses and persons that the regular marker *-RI/ -Rĩ* is replaced, but, this being irrelevant for the present discussion, it will suffice to say here that they are etymologically different from the *-RI/ -Rĩ* suffix and apparently represent a rather recent innovation.

It should also be noted that the interrogative suffix *-RI* seems to be a kind of default question marker, because it "may also be used with the other combinations of person/tense [= other than those mentioned in Table 1 above – our addition], except for the first person of the future tenses" (Morse & Maxwell 1999:21). So, etymologically there turns out to be only one interrogative suffix in Cubeo, and it is *-RI*.

As far as nominalizers in Cubeo are concerned, we find an inanimate nominalizer *-RI*, used to form concrete deverbal nouns.

- (8) *kĩrãbĩ* *ki-RI-kũ* 'houseboat' (Morse & Maxwell 1999:86)  
 house live-NOM-CLS

The result of all this is that in Cubeo, as in Desano, the question marker and the inanimate nominalizer are formally identical.

### 2.1.3. Yurutí

In Yurutí polar questions and information questions are marked in the same way as in Desano and Cubeo. No information on the intonation patterns of either polar questions or information questions in Yurutí was found in Kinch & Kinch (2000).

Below we reproduce a slightly adapted table from Kinch & Kinch (2000). It summarizes the patterns of the usage of interrogative suffixes in present and past tenses. As to the future, there are three declarative future forms in Yurutí but there is only one interrogative future form based on the combination of one of these three future markers and the interrogative marker *-ri*.

	Visual	Nonvisual	Deductive ("indicial"), Reportative, Inferential
Present			<i>[-ga]-ri</i>
Past	<i>-ri/ -re</i>		<i>[-ju]-ri</i>

TABLE 2. Interrogative suffixes in Yurutí (based on Kinch & Kinch 2000)

Kinch & Kinch (2000:472) explicitly state that this suffix is intrinsically oral.<sup>5</sup>

Let us now turn to the inanimate nominalizer in Yurutí. In Yurutí, as in many other Tucanoan languages, the nominalizers often amalgamate with other morphemes to accommodate tense/aspect distinctions. In order not to overload the presentation with an analysis of different morphological and morphonological peculiarities that are rather irrelevant for the topic at issue, the discussion will be confined to the inanimate nominalizing suffixes used in Present tense.

Countable		Uncountable
SG	PL	
<i>-ri</i> + CLS	<i>-re</i> + PL.CLS	<i>-re/ -ri</i>

TABLE 3. Present tense inanimate nominalizers in Yurutí (based on Kinch & Kinch 2000)

The exact nasalization pattern of this nominalizer is not completely clear in Kinch & Kinch (2000), due to certain inconsistencies in the nasalization marking of this suffix in the examples in their text. However, because in the examples in the Swadesh list appended to Kinch & Kinch (2000) the suffix *-re* used to mark the “infinitive” forms of the verbs is consistently marked as oral in oral environments and as nasal in nasal environments, we are inclined to consider the inanimate nominalizer as being unspecified for nasality.

Although both the form *-re* and the form *-ri* of the inanimate nominalizer seem to be formally identical to the interrogative suffix, it is only the *-ri* form that is relevant here. A comparison with other Eastern Tucanoan languages indicates that the form *-re* of the inanimate nominalizer goes back to a combination of the nominalizer *-ri* and a general inanimate classifier. The latter is found, among others, in Barasano as *-se* and in Tatuyo as *-hel/-yel/-e*. Moreover, in Barasano we can also find the same fusion *-re* between the nominalizer and inanimate classifier (for more details see Gómez-Imbert & Jones 2000).

As a result, even though the interrogative marker and the inanimate nominalizer happen to be completely formally identical in Yurutí, it is only the *-ri* forms of these morphemes that are structurally identical as well.<sup>6</sup> As to the *-re* form of the interrogative marker, see 3.2.

## 2.2. Western Tucanoan

### 2.2.1. Secoya

The categories marked on an independent verb in Secoya are comparable to those that we find in many Eastern Tucanoan languages, namely gender, number, and person of the subject, tense/aspect, evidentiality status and mood (declarative, interrogative, imperative). The exact patterning of the categories at issue differs however in interesting aspects. For the peculiarities that are of immediate importance for the topic of our discussion, that is the relations between interrogative suffixes and nominalization markers in Tucanoan, we will permit ourselves to deviate a bit from the model we have been using up till now in our presentation of data.

Let us first consider the Secoya subject agreement suffixes used in declarative mood. There are two paradigms of such suffixes distinguished according to evidentiality status, or “perspectiva” as Johnson & Levinsohn (1990) call it in Spanish. There are two such perspectives in Secoya: “perspectiva involucramiento”, by means of which the speaker “afirma su participación o presencia física en los acontecimientos que describe” (Johnson & Levinsohn 1990:66), and “perspectiva de separación”, by means of which the speaker

“indica desconocimiento o falta de testimonio directo de los acontecimientos que describe” (Johnson & Levinsohn 1990:69). The subject agreement suffixes of the “modo declarativo, con perspectiva involucramiento” for Present and Immediate Past tenses are given in Table 4 below; for convenience sake, we will call them “set I suffixes”. Analogous suffixes for Remote Past and Future tenses are not cited here, for they do not add any new relevant information.

	Present	Immediate Past
3SG.M	-hi	-pi  -hi?i
3SG.F	-ko	-ko  -ko?i
“the rest” (NON3SG:AN)	-ji	-wi  -i

TABLE 4. Set I suffixes (based on Johnson & Levinsohn 1990:66)<sup>7</sup>

As to the nasalization patterns of these and all other suffixes given in this section, all suffixes beginning with *w*, *j*, *h* or *ʔ* and unmarked for nasality allow for nasal spreading and are thus unspecified for nasality; for the rest the nasalization pattern is as marked (Johnson & Levinsohn 1990:21).

Table 5 below summarizes the subject agreement suffixes used on an independent verb in anything else than the “modo declarativo, con perspectiva involucramiento” or the imperative mood. In other words, these suffixes are used in the case of “perspectiva de separación” of the declarative mood and, interestingly enough, in the case of interrogative mood for both polar questions and information questions. For convenience sake, we will call these suffixes “set II suffixes”. The attention of the reader should also be drawn to the fact that set II suffixes are distinguished not by tense as in the case of the “modo declarativo, con perspectiva involucramiento” but by aspect: we are dealing here with perfective aspect (“aspecto finalizado, perfectivo”) and imperfective aspect (“aspecto no-finalizado, imperfectivo”). As a result, we will, for instance, have two Remote Past tense forms, a perfective one and an imperfective one (the Remote Past tense itself is marked by the suffix *-aʔ* preceding the subject agreement suffixes).

	Imperfective	Perfective
3SG.M	-#-i -u	-ki
3SG.F	-o	-ko
“the rest” (NON3SG:AN)	-je	-de -te

TABLE 5. Set II suffixes (based on Johnson & Levinsohn 1990:69)

Note that the forms of the imperfective aspect subject agreement suffixes most likely result from lenition of the initial consonants of the perfective forms with subsequent deletion of the lenited consonants for the 3SG.M and 3SG.F forms. A similar lenition in the forms of the perfective aspect subject agreement suffixes must have been blocked by a completive aspect (“aspecto completo”) marker *-si*, cf. (31). Although the latter disappeared afterwards itself, there is enough evidence both comparative and language-internal in favour of this hypothesis. Compare, for instance, the perfective non-agentive nominalizer *-seʔe* (<\**-si* + *-je*) in Table 6 below;<sup>8</sup> see also (Wheeler 1992a; Wheeler 1992b) and note 10.

In the case of “perspectiva de separación”, the set II suffixes are followed by an additional suffix *-jã* (9).

- (9) *jãse waʔi-pi jeha-dã tõ mẽ-jẽ-jã*  
 toucan PL-SUBJECT ground-to fall go.down-NON3SG:AN.IPFV-SEP  
 ‘The toucans have fallen on the ground’ (but the speaker did not see them falling)  
 (Johnson & Levinsohn 1990:70)

Polar questions are marked by set II suffixes only, as in (10) from (Johnson & Levinsohn 1990:71); no information on the intonation of polar questions is given by (Johnson & Levinsohn 1990). Information questions, as in (11) from (Johnson & Levinsohn 1990:72), are marked by set II suffixes followed by a suffix *-ʔdĩ* which marks DOUBT (“duda”) and is probably best translated as ‘perhaps, maybe’.<sup>9</sup>

- (10) *mõũ jure* (11) *hero-pi tidi-o-ʔdĩ*  
 to.fish-M.SG.IPFV now where-from grumble-3SG.F.IPFV-DOUBT  
 ‘Is he fishing now?’ ‘Where is she grumbling from?’

Interestingly, polar question verb forms in Secoya can also be used with a non-interrogative meaning, namely to express a feeling of probability or uncertainty (Johnson & Levinsohn 1990:71), as in (12). In addition, a negative suffix *-mãʔ* can be used in order to increase “la incertidumbre del evento” (Johnson & Levinsohn 1990:71), as in (13, 14). An adverb *jekẽ* ‘perhaps, maybe’ or the suffix *-ʔdĩ* DOUBT can also be added (Johnson & Levinsohn 1990:71), as in (14).

- (12) *hãhẽ paʔi-to, jiki hopo koʔi-de*  
 so be-if IPL.EXCLUSIVE half turn.back-NON3SG:AN.PFV  
 ‘If it were like this, we had probably turned back halfway’ (Johnson & Levinsohn 1990:71)
- (13) *haʔje paʔi-mãʔ-ki*  
 near be-NEG-3SG.M.PFV  
 ‘He may be nearby’ (Johnson & Levinsohn 1990:71)
- (14) *kũĩ-mãʔ-ki-ʔdĩ*  
 bite-NEG-3SG.M.PFV-DOUBT  
 ‘He will possibly bite (her/him)’ (Johnson & Levinsohn 1990:71)

As to Secoya nominalizing suffixes, as will readily be observed from comparison of Tables 5 and 6, they show a striking resemblance to the evidential/interrogative suffixes of the set II. This resemblance is so remarkable that it was explicitly stated by Johnson & Levinsohn (1990:69) themselves. It is worth mentioning in this respect that, despite the fact that nominalizing and interrogative suffixes bear a strong formal resemblance in most Tucanoan languages, it is explicitly mentioned only in one grammar, i.e., that of Secoya by Johnson & Levinsohn (1990).

	Agentive		Non-agentive	
	Imperfective	Perfective	Imperfective	Perfective
M.SG	<i>-#i/-u</i>	<i>-ki</i>	<i>-je</i>	<i>-seʔe (&lt;*-si + -je)</i>
F.SG	<i>-o</i>	<i>-ko</i>		
PL	<i>-o waʔi</i>	<i>-ko waʔi</i>		

TABLE 6. Nominalizing suffixes (based on Johnson & Levinsohn 1990:35)<sup>10</sup>

Note that the agentive F.SG marker is in fact best described as “singular non-masculine” agentive nominalizer because it is also used for agentive nominalizations when the nominalization is inanimate, and as a part of the agentive plural nominalizer for both animate (M/F) and inanimate nominalizations. As to the non-agentive nominalization, no distinction is made between animate/ inanimate and singular/ plural.

The term non-agentive (“no-agentivo”) implies that we are dealing with “una construcción abstracta o no activa” (Johnson & Levinsohn 1990:37), see (15, 16, 17).

- (15) *jã-õ* *siwa-je-té* *kuasa-hi*  
 that-NOM.A.F.SG.IPFV be.happy-NOM.NA.IPFV-OBJ think-3SG.M.PRESENT  
 ‘He is thinking about her happiness’ (Johnson & Levinsohn 1990:37)
- (16) *kuepe-se?e-pi* *ko?a-hi* *ji?i-de*  
 get.drunk-NOM.NA.PFV-because.of be.bad-3SG.M.PRESENT 1SG-OBJ  
 ‘I feel bad (now) because I got drunk (then)’ (Johnson & Levinsohn 1990:37)
- (17) *jo?o-je* ‘[the fact, the event of] being done’ (Wheeler 1992b)  
 do-NOM.NA.IPFV

Note that nominalizations on *-je/-se?e* seem to have an “abstract” interpretation primarily in the case of intransitive verbs, whereas “non-agentive” interpretation appears to be most natural in the case of transitive verbs. As far as we can judge, the same seems to hold true for the cognate suffixes in other Western Tucanoan languages as well, and probably even in Eastern Tucanoan as examples (23, 24) from Desano in 3.1 illustrate.

Finally, it should be noted that besides the nominalizing suffixes given in Table 6, it is also possible in Secoya to use just ordinary (inanimate) classifiers in a nominalizing function, as illustrated in (18). In fact, the same can be done to various extents in many other both Eastern and Western Tucanoan languages as well.

- (18) *a?si do?a-wi-de wẽ-wĩ*  
 hurt gnaw-CLS-OBJ tie.up-NON3SG:AN.PAST  
 ‘I tied up my finger which has a painful bite’ (Johnson & Levinsohn 1990:77)

Note that being inanimate, such nominalizations are naturally most often non-agentive. But, at the same time, they rather inevitably have concrete semantics. So, as a rule, they are non-agentive and concrete. On the other hand, elsewhere in Secoya the term non-agentive implies, as the reader might remember, that we are dealing with “una construcción *abstracta* o no activa” [italics ours] (Johnson & Levinsohn 1990:37). Agentive nominalizations in turn are regularly concrete. In other words, nominalizations in Secoya turn out to be organized in two overlapping and to a certain extent correlated oppositions, that between agentive and non-agentive nominalizations and that between concrete and abstract nominalizations. The former opposition seems, however, to be more morphologically prominent in Secoya.

Let us now briefly subsume the most important moments of what has been said about Secoya in this section. First, in Secoya, the same forms are used in three different functions: (a) to mark subject agreement and non-visual evidentiality, (b) to mark subject agreement and interrogative mood, (c) to mark subject agreement and doubt. Second, these markers are the same as, or at least clearly related to, the nominalization markers. Third, unlike in Eastern Tucanoan languages, there is no category of general inanimate nominalizers in Secoya; instead, the abstract/non-agentive nominalizer *-je/-de/-te* is used.

Furthermore, there is enough evidence to believe that among the three it is the *-te* allomorph that is the original one. Note that the marker *-te* bears a striking resemblance to the inanimate nominalizer/ interrogative suffix *-ti* which we find in Eastern Tucanoan languages.

### 2.2.2 Koreguaje

In Koreguaje nominalizations, we find an opposition between concrete nominalizations (marked by M, F markers and CLS), on the one hand, and abstract inanimate nominalizations (marked by a suffix *-tʃe*)<sup>11</sup>, on the other hand (see Cook & Criswell 1993:30). The latter mean something like ‘the fact, the event of doing/ being/ undergoing *V*’, (19, 22), or ‘something (unspecified for classification) that does/ is/ undergoes *V*’, (20, 21). Note that (21, 22) are nominalizations of transitive verbs and, comparably to what we find in Secoya (17), they are non-agentive.

- (19) *tʃiʃi paʔi-tʃe* ‘my life’ (Cook & Criswell 1993:28)  
 1SG live-NOM.INAN:ABSTRACT
- (20) *kuʔa-tʃe*  
 be.bad-NOM.INAN:ABSTRACT  
 ‘something bad/ bad thing/ bad things’ (Cook & Criswell 1993:28)
- (21) *tʃoʔo-ha-tʃe*  
 do-POTENTIAL-NOM.INAN:ABSTRACT  
 ‘something that will be done’<sup>12</sup> (Cook & Criswell 1993:30)
- (22) *tʃoʔo-tʃe* ‘[the fact, the event of] being done’<sup>13</sup> (Wheeler 1992b)  
 do-NOM.INAN:ABSTRACT

As in Secoya, nominalizers and interrogative suffixes in Koreguaje bear a striking formal resemblance to each other (cf. Cook & Criswell 1993:30, 56). Moreover, they are also very similar in form to their Secoya counterparts. However, they are not reported to be used as markers of subject agreement and doubt or subject agreement and non-visual evidentiality.

As far as the intonational patterns of interrogative utterances in Siona are concerned, Rodríguez González (2000) cite the following information from Herrera (1991:31-38). In polar questions, the pitch contour can be described as middle-low-high. In information questions, it is high-middle-low and in declarative utterances it is middle-low-low.

### 3. From a nominalizer to an interrogative suffix

Above, we have tried to show that in both Eastern and Western Tucanoan languages question markers and the inanimate/ abstract inanimate/ non-agentive nominalizers are formally most similar, if not identical. Furthermore, in some Western Tucanoan languages the same markers are used to express doubt or uncertainty as well as non-visual evidentiality. Now we will argue that the nominalization use is the origin of the interrogative use. We will try to provide some detailed argumentation for the etymological link between the nominalizing and interrogative suffixes in the present section.

In 3.1, we will consider nominalizations in a broader cross-linguistic context. In particular, our attention will be devoted to independent non-finite predications in the languages of the world. We will show that such predications are often used to express certain evidential and modal meanings. With this knowledge, we will return to the

Tucanoan languages in 3.2 where a possible etymological link between the nominalizing and interrogative suffixes will be proposed.

### 3.1. Nominalizations, non-finite predications and the meanings they may convey

We have seen earlier that in Tucanoan languages questions share an important formal feature with nominalizations (or “deverbalizations”): they are both non-finite (or maybe better, they are both less finite than declaratives) because they both lack the normal subject agreement suffixes. Some most obvious examples of this non-finiteness can be found among the Eastern Tucanoan languages, such as Desano (2.1.1).

Being inherently non-finite, a nominalized predication, like (23) or (24) in Desano, is pragmatically “incomplete”, that is, it cannot normally function as a main clause.

- |      |   |      |  |
|------|---|------|--|
| (23) | <i>ba-ri</i><br>eat-NOM<br>'something to eat' (Miller 1999:141) | (24) | <i>i yiʔi buʔe-ri-a</i><br>this 1SG study-NOM-PAST<br>'what I studied' (Miller 1999:141) |
|------|---|------|--|

However, under certain conditions, a noun phrase can become an utterance, provided an appropriate context and an appropriate intonation. For instance, we can use the noun *milk* as an assertion or (giving it an “unfinished” intonation pattern) as a question. Such utterance uses of nouns or nouny entities, like infinitives, may sometimes get associated with a certain conversational implicature. This pragmatic inference may in its turn get conventionalized. Consider, for example, the case of the German infinitive-cum-imperative such as illustrated in (25).

- (25) *aufstehen* ‘to stand up’ (infinitive) vs. ‘stand up!’ (imperative, authoritative order).

Example (26) represents another instance of an independent use of infinitive in German and, at the same time, of the *ing*-form in its English translation.

- (26) *Ich und morgen heimgfahen!*  
1SG and tomorrow go.home  
'Me going home tomorrow!?' [What nonsense!]' (Zaefferer 1990)

In fact, in many languages of the world verbal forms are attested that can function both as finite (i.e., as the predicate of an independent or main clause) and as non-finite (i.e., as an attribute of a noun phrase or as a dependent predicate showing most often also certain morphosyntactic features characteristic of nouns). Kalinina (1999:214) proposes to interpret such forms not as simply two different but homonymous forms, as is ordinarily done, but as one “acategorical” form characterized by a single invariant grammatical meaning. She gives examples of such uses for Lithuanian (Indo-European, Baltic), Nenets (Uralic, Samoyedic), (Siberian) Eskimo (Eskimo-Aleut), several North Caucasian languages (particularly, Bagvalal and Tsakhur), and many Altaic languages (particularly, Nanai and Nogai). Below we will cite some of the examples given in Kalinina (1999), which we consider to be relevant for our discussion.

In Lithuanian, one speaks about a so-called “indirect mood”, which is characterized by the use of a participle (of any tense) in the position of an independent predicate. “The use of the indirect mood indicates that either the speaker is not sure enough about what s/he is saying [see (27)] or that the speaker is not an eyewitness of what s/he is saying [see (28)]” (*Grammatika litovskogo jazyka* 1985:231; cited via Kalinina 1999:215, our translation).

- (27) *užkastieji*                      *pinig-ai*              *degā*  
 bury.PARTICIPLE.PAST money-PL burn.PRESENT.PARTICIPLE.PL  
 ‘The buried money seems to be burning/ is probably burning’ (*Grammatika litovskogo jazyka* 1985:234; cited via Kalinina 1999:215)
- (28) *čia sako*                      *mišt-ai*              *buwē*  
 here say.PRESENT.3SG forest-PL be.PAST.PARTICIPLE.PL  
 ‘Here, he says, there used to be forests’ (*Grammatika litovskogo jazyka* 1985:234; cited via Kalinina 1999:215)

In Nenets, a participle in the position of an independent predicate may be used to express an evidential meaning of the reportative or hearsay kind. In Eskimo, the situation is close to that of Nenets. One speaks about an “inferred evidential”. An example from (Siberian) Eskimo:

- (29) *najg'an-ən'*                      *ukigu-ma-lg'i*                      *tagitu-k*  
 mountains-INSTRUMENTAL go.down-PAST2-PARTICIPLE.PROCESSUAL mist-ABS  
 ‘From the mountains, obviously, the mist came down’ (Menovschikov 1960:205; cited via Kalinina 1999:216, the Eskimo example is transliterated from a Cyrillic alphabet)

In a rather closely related Central Alaskan Yupik Eskimo participle constructions used as independent main clauses “signal vividness and sometimes exclamatory force ... forming a paradigmatic set with the other main clause moods, the indicative, interrogative, and optative” (Woodbury 1985).

Here, we would like to draw attention to the important fact that in languages like Lithuanian, Nenets and Eskimo the non-finite form used as an independent predicate expresses not only different kinds of evidentiality but also uncertainty of the speaker. The latter is especially relevant for our discussion. There is obviously nothing strange in this link between uncertainty and evidentiality. On the one hand, when you are not an eyewitness of the event, you cannot normally be completely sure about all of the event’s circumstances. On the other hand, when you are not completely sure about all of the event’s circumstances, this is most likely because you have not witnessed everything properly yourself, otherwise you would have been more sure about the event.

From this perspective it should come as no surprise that the same markers are used in such Western Tucanoan languages as Secoya to express both nominalizations and doubt or uncertainty as well as non-visual evidentiality.<sup>14</sup> Of course, in Tucanoan, the situation is more complex still, for in these languages strikingly similar, if not identical, markers are also used as interrogative suffixes.

### 3.2. Back to Tucanoan

We can now go back to the Tucanoan interrogative suffix at issue. In section 2, we have tried to show that in the Tucanoan languages question markers and nominalizers are formally most similar, if not identical. Furthermore, in some Western Tucanoan languages, of which Secoya (2.2.1) is probably the clearest example, the same markers are used to express doubt or uncertainty as well as non-visual evidentiality. In 3.1, we have demonstrated that in various languages of the world nominalizations can be used as independent predicates and that when they are used this way they rather often express, among other things, uncertainty of the speaker.

Since a possibility of existence of a link from nominalization to the expression of doubt has been established, all we need is one further step, from the expression of doubt to the expression of a question. This now is actually an easy step. Such scenario is attested elsewhere in the world. Consider, for instance, the case of Warrwa (a Nyulnyulan language of Western Australia). In Warrwa, the polar question is marked by an enclitic *-wada*, which is normally attached to the first word of the clause:

- (30) *juwa-na-wada mingkany*  
 you-ERGATIVE-YNQ you:hit:him  
 ‘Did you hit him?’ (McGregor 1994:36)

According to McGregor (1994:37) “it is quite likely that *-wada* is cognate with the particle *wadi* ‘perhaps’ in Gunin”, a neighbouring Wororan language. The development at issue simply represents a conventionalization of the same implicatures as the ones that an indirect speech act like (31) can have in English.

- (31) – *He probably wants to play.*  
 – *No, I think he is too tired for this now.*

Furthermore, note that expressions of nominalizations, questions and doubt can all be regarded as non-assertive, as merely presenting propositions. This is strongly reminiscent of what Zaefferer (1990) calls “presentative mode”, a kind of sentential modality even more basic than the assertive one. According to Zaefferer (1990), “the function of the presentative would be simply to represent a proposition without any specification of the role it is to play in the discourse, especially without any commitment to its truth, falsity, pleasantness, or whatever, except that it is presented and therefore accessible as a possible topic for the ongoing discourse”. And indeed, this is exactly what expressions of doubt, questions and nominalizations do. The only difference between nominalizations and the rest is that nominalizations are not inherently predicative. However, as demonstrated in 3.1, this discrepancy can be overcome rather easily in discourse and the result of this can consequently get conventionalized.

Interestingly, among the few examples of presentatives that Zaefferer (1990) gives we find an interrogative, namely the Huichol (Uto-Aztecan) interrogative. The Huichol interrogative is also discussed by Palmer (1986). Palmer (1986:33) speaks about a “curious situation in Huichol ... where the form used as a question is often the unmarked form, though Huichol also has a marker for the ‘assertive’ mood” and proposes to consider this form as a kind of weak declarative “merely presenting propositions”. Contrary to Palmer, Zaefferer (1990) claims that the best analysis would be to regard this form as a “straightforward presentative, which naturally under certain circumstances lends itself to interrogative use”. So, being the most neutral mode just making a given proposition “accessible as a possible topic for the ongoing discourse”, presentative can have different interpretations depending on the situation. In other words, the concrete interpretation of a presentative form relies heavily on the context (both linguistic and extra-linguistic), most likely also intonation, accompanying gestures, etc. Note that this is, in fact, very reminiscent of what we find in Secoya.

Now we can return to the (non-visual) evidentiality, mentioned in the beginning of this section as another meaning the markers of doubt, questions and nominalizations can express in languages like Secoya. Given a strong functional link between evidentiality and uncertainty already highlighted in 3.1, it is not surprising to see evidentiality in this row.

An evidential expression can easily be interpreted as non-assertive because the speaker has all the rights to avoid asserting something he has not witnessed himself.

Finally, note that whereas in the Western Tucanoan languages we find a whole paradigm of interrogative markers, in the Eastern Tucanoan languages we find only one interrogative marker. The interrogative marker these two systems apparently have in common has the form *\*-te* in Western Tucanoan and *\*-ti* in Eastern Tucanoan. Interestingly, the latter is also the same as the inanimate nominalizer and the former is the same as the (inanimate) abstract (and/or non-agentive) nominalizer. All these similarities taken together cannot be a coincidence. Furthermore, if we remember (cf. 2.2.2) that the term “(inanimate) abstract” actually means ‘something (unspecified for classification) that *V*-s’ or, what is more important, ‘the fact, the event, the process, the instance of *V*-ing’, we realize that it is exactly this abstract nominalizer that is best suited for the role of a presentative marker which would simply present “the fact, the event, the process, the instance of *V*-ing” into the discourse without asserting it.

The two morphemes, *\*-te* and *\*-ti*, are clearly cognate. However, due to space limitations we cannot demonstrate this in detail within the present paper (for more details, see Idiatov & van der Auwera, ms.). Thus, it will suffice to say that we believe the original form to be *\*-te* and that in Eastern Tucanoan a shift from *e* to *i* took place due to some systemic factors and probably a contact interference from the neighbouring Arawak languages. Yurutí (2.1.3) seems to be the only Eastern Tucanoan language that has preserved the older form of the interrogative suffix next to the new one.

As a conclusion, everything said in the present section can be subsumed under the following claim. An abstract nominalizer, meaning ‘the fact, the event of *V*-ing’ or ‘something that *V*-s’, must have existed in Proto Tucanoan. Nominalizations marked by this morpheme were also used as independent utterances and functioned then as presentatives. In particular, they expressed doubt and questions.

## 5. Conclusion

In the present article, we have argued that the Tucanoan languages exhibit a historical and semantic relationship between nominalization (or “deverbalization”) and question formation. We have also tried to demonstrate that the latter evolved from the former via expression of doubt. The hypothesis we have advanced relies on the assumption of a possible polyfunctional use of nominalizations and the existence of a common non-assertive feature shared by nominalizations, expressions of doubt and questions and appeals to the notion of “presentative” mode.

Finally, it can be mentioned that the proposal tracing the interrogative marker back to a nominalizer might throw an interesting light on the very fact that in the relevant Tucanoan languages the interrogative marker co-occurs with question words. According to, for instance, the Clause Typing Hypothesis of Cheng (1997:17), in the languages of the world there tends to be only one intrinsically interrogative marker per clause. There is no co-occurrence restriction between question words and nominalization markers. The fact that the interrogative marker easily combines with question words could thus be seen as deriving from the former’s non-interrogative origin.

## Notes

<sup>1</sup> This paper was written as part of the GOA (Geconcerteerde Onderzoeksactie) “Mood and Modality” project of the University of Antwerp. We would like to thank Mark Van de Velde and Elsa Gómez-Imbert for their comments on some of the previous versions of this work. For bibliographical help we are grateful to Mily Crevels, Elsa Gómez-Imbert, and Jim Shephard.

<sup>2</sup> Abbreviations: A – Agentive, ABS – Absolutive, AN – Animate, ASSUM – Assumed evidential, CLS – Classifier, F – Feminine, HSY – Hearsay evidential, IPFV – Imperfective, M – Masculine, NA – Non-agentive, NEG – Negative, NOM – Nominalizer, OBJ – (direct) Object, PFV – Perfective, PL – Plural, Q – Question marker, SEP – “Perspectiva de separación” (in Secoya, 2.2.1), SG – Singular, SPC – Specifier, STV – Stativizer, YNQ – “Yes/No” (= polar) question, NON3 – Non-third person.

<sup>3</sup> It should be mentioned that none of the sources on Desano (available to us) gives any information as to whether the utterance containing the verb phrase marked by interrogative *-ri* is characterized by a special intonation pattern or not.

<sup>4</sup> Morse & Maxwell (1999:9-11) use the capital *R* below to indicate “an archiphoneme ambiguous between /r/ and /d/”. Upper case letters are furthermore used for vowels unmarked for nasality, i.e., vowels that can be oral or nasal depending on their (morpho)phonological environment.

<sup>5</sup> We have however encountered several (misprinted?) examples in their text, with the interrogative suffix marked with a nasalization diacritic as if it had undergone nasal spreading.

<sup>6</sup> The nasalization patterns not being taken into consideration.

<sup>7</sup> It should be noted that we have transcribed all Secoya examples we give in the text from the practical Spanish-based orthography used by Johnson & Levinsohn (1990) into phonological notation in conformity with the rules established by Johnson & Levinsohn (1990:23).

<sup>8</sup> The reader might however notice that the lenition was blocked (at least partially) only for the *-de/-te* NON3SG:AN.IPFV suffix of set II, but not for the perfective non-agentive nominalizer *-seʔe* (<\*-*si* + *-je*). This can probably be explained by paradigmatic factors. Specifically, note that whereas the former suffix and other (non-lenited) perfective suffixes, *-ki* and *-ko*, directly group together in one perfective subparadigm as opposed to the imperfective subparadigm, the latter suffix groups first of all with the (lenited) incomplete non-agentive nominalizer *-je* in order to form the non-agentive subparadigm as opposed to the agentive subparadigm.

<sup>9</sup> Compare its Siona cognate *-ʔnē* as found in *judé-ʔnē* ‘perhaps now’ (Wheeler 1970:148).

<sup>10</sup> It is the non-lenited perfective nominalizers that should be considered as basic forms here, comparably to the set II perfective markers of Table 5 already discussed. As a kind of indirect evidence, one can adduce Secoya nominal human gender suffixes, *-ki* M.SG, *-ko* F.SG, and *-ko waʔi* PL (Johnson & Levinsohn 1990:26), which have unvoiced initial consonants free from any possible additional morphonological conditioning.

<sup>11</sup> The notation *-tʃe* is phonological and in conformity with the phonological analysis by Rodríguez González (2000). Cook & Criswell (1993), as well as Wheeler (1992b), write this morpheme as *-ʃe* which equals /dʒe/. In the Koreguaje examples in this text, we use *tʃ* instead of *ʃ*.

<sup>12</sup> In all probability, also ‘the fact, the event of being done in the future’.

<sup>13</sup> In all probability, also ‘something that is being done’.

<sup>14</sup> One can also mention here several Eastern Tucanoan languages. For instance, in Carapana, according to Metzger (1981:13), one can indicate “duda por medio de una pregunta”. Note also that in Retuarã the nominal neuter gender suffix *-a/-ka* which can also function as inanimate nominalizer is sometimes used on independent verbs either in order to background an event in the discourse or to “reduce the responsibility” of the speaker for the state of affairs expressed in the clause (Strom 1992:192-193, 198-200). Interestingly, in another Eastern Tucanoan language

Cubeo, we also find a *-ka* suffix, but Morse & Maxwell (1999:24) gloss it as “dubitative” and define (1999:190) it as a “discourse morpheme” expressing “narrator doubt”. It is certainly conceivable that the Cubeo dubitative suffix *-ka* is a cognate with the Retuarã inanimate nominalizing suffix *-ka*. At least there is a semantic connection, for doubt is also the expression of a reduced speaker responsibility. It is of a different subtype, an epistemic one, whereas in Retuarã the reduced speaker responsibility of the second pragmatic use concerns the state of affairs itself.

#### References

- Barnes, Janet (1999). Tucano. In Dixon, R.M.W. & Alexandra Y. Aikhenvald (eds.), *The Amazonian Languages*, 207-226. Cambridge: Cambridge University Press.
- Cheng, Lisa Lai-Shen (1997). *On the typology of wh-questions*. New York: Garland.
- Cook, Dorothy M. & Linda L. Criswell (1993). *El idioma koreguaje (tucano occidental)*. Bogotá: Instituto Lingüístico de Verano, Editorial Alberto Lleras Camargo.
- Ferguson, Judith, Cari Hollinger, Linda Criswell & Nancy L. Morse (2000). El cubeo. In González de Pérez, María Stella & María Luisa Rodríguez de Montez (eds.), *Lenguas indígenas de Colombia: Una visión descriptiva*, 357-372. Bogotá: Instituto Caro y Cuervo.
- Gómez-Imbert, Elsa & Stephen Hugh Jones (2000). Introducción al estudio de las lenguas del Piraparaná (Vaupés). In González de Pérez, María Stella & María Luisa Rodríguez de Montez (eds.), *Lenguas indígenas de Colombia: Una visión descriptiva*, 321-356. Bogotá: Instituto Caro y Cuervo.
- Gralow, Frances L. (1993). *Un bosquejo del idioma koreguaje*. Bogotá: Instituto Lingüístico de Verano.
- Grammatika litovskogo jazyka [A grammar of Lithuanian]* (1985). Vilnius.
- Herrera, Gloria Elena (1991). *Manual de pronunciación para hablantes koreguajes, basado en el análisis contrastivo a nivel fonológico de las dos lenguas*. Tesis de Licenciatura en Filología e Idiomas. (ms.), Universidad Nacional de Colombia.
- Idiatov, Dmitry & Johan van der Auwera (ms.). Polyfunctionality of nominalizations: from nominalizations to doubts and questions in Tucanoan.
- Johnson, Orville E. & Stephen H. Levinsohn (1990). *Gramática secoya [basada en un borrador elaborado por Alva Wheeler]*: Cuadernos Etnolingüísticos, 11. Quito: Instituto Lingüístico de Verano.
- Jones, Wendell & Paula Jones (1991). *Barasano syntax: Studies in the languages of Colombia*, 2. Dallas: Summer Institute of Linguistics & University of Texas at Austin.
- Kalinina, Ye.Yu. (1999). Ob omonimii i invariantah [On homonymy and invariants]. In Rakhilina, E.V. & Ya.G. Testelec (eds.), *Tipologija i teorija jazyka: Ot opisanija k ob'jasneniju. K 60-letiju A.Ye. Kibrika [Typology and Theory of Language: From Description to Explanation. For the 60th anniversary of A.Ye. Kibrik]*, 212-223. Moscow: "Jazyki russoj kul'tury".
- Kaye, Jonathan D. (1970). *The Desano verb: Problems in semantics, syntax and phonology*. Ph.D. thesis, Columbia University.

- Kinch, Rodney & Pamela Kinch (2000). El yurutí. In González de Pérez, María Stella & María Luisa Rodríguez de Montez (eds.), *Lenguas indígenas de Colombia: Una visión descriptiva*, 469-488. Bogotá: Instituto Caro y Cuervo.
- McGregor, William (1994). *Warrwa: Languages of the world/ Materials*, 89. München & New Castle: LINCOM Europa.
- Menovschikov, G.I. (1960). *Eskimoskij jazyk. Posobie dlja peduchilisch [The Eskimo language. A manual for normal schools]*. Leningrad.
- Metzger, Ronald G. (1981). *Gramática popular del carapana*. Bogotá: Instituto Lingüístico de Verano.
- Miller, Marion (1999). *Desano Grammar: Studies in the Languages of Colombia 6: Summer Institute of Linguistics Publications in Linguistics 130*. Dallas, Tex.: Summer Institute of Linguistics.
- Morse, Nancy L. & Michael B. Maxwell (1999). *Cubeo Grammar: Studies in the Languages of Colombia 5: Summer Institute of Linguistics Publications in Linguistics 130*. Dallas, Tex.: Summer Institute of Linguistics.
- Palmer, F.R. (1986). *Mood and Modality*. Cambridge: Cambridge University Press.
- Rodríguez González, Sandra Patricia (2000). Estudios sobre la lengua koreguaje. In González de Pérez, María Stella & María Luisa Rodríguez de Montez (eds.), *Lenguas indígenas de Colombia: Una visión descriptiva*, 199-214. Bogotá: Instituto Caro y Cuervo.
- Strom, Clay (1992). *Retuarã Syntax: Studies in the languages of Colombia*, 3. Dallas: Summer Institute of Linguistics & University of Texas at Austin.
- Waltz, Nathan E. & Alva Wheeler (1972). Proto Tucanoan. In Matteson, Esther, Alva Wheeler, & Frances L. Jackson (eds.), *Comparative studies in Amerindian languages*, 119-149. The Hague: Mouton.
- Wheeler, Alva (1970). *Grammar of the Siona language, Colombia, South America*. Ph.D. thesis, University of California.
- (1992a). Comparaciones lingüísticas en el grupo Tucano Occidental. In Levinsohn, Stephen H. (ed.), *Estudios comparativos: Proto Tucano*, 17-54. Bogotá: Instituto Lingüístico de Verano.
- Wheeler, Margaret (1992b). Sufijos nominales en el Proto Tucano Occidental. In Levinsohn, Stephen H. (ed.), *Estudios comparativos: Proto Tucano*, 54-76. Bogotá: Instituto Lingüístico de Verano.
- Woodbury, Anthony C. (1985). Noun phrase, nominal sentence, and clause in Central Alaskan Yupik Eskimo. In Nichols, Johanna & Anthony C. Woodbury (eds.), *Grammar inside and outside the clause: some approaches to theory from the field*, 61-88. Cambridge: Cambridge University Press.
- Zaefferer, Dietmar (1990). On the coding of sentential modality. In Bechert, Johannes, Giuliano Bernini, & Claude Buridant (eds.), *Toward a Typology of European Languages*, 215-237. Berlin & New York: Mouton de Gruyter.

# External Possession in Huehuetla Tepehua\*

Susan Smythe Kung  
University of Texas at Austin

## 1. Introduction

External possession constructions are those constructions in which (i) a possessor of a possessed nominal (the possessum) is a core argument of the clause, not simply a dependent of the possessum, and (ii) the possessor is not included in the argument frame of the lexical verb root. External Possession has also been referred to as possessor raising and possessor ascension (Payen and Barshi 1999). The purpose of this paper is to describe two morphosyntactic strategies used in Huehuetla Tepehua to convey external possession and to compare the two Tepehua strategies with similar phenomena in three other Mesoamerican languages: Papantla Totonac, Oluta Popoluca, and Tzotzil.

## 2. Huehuetla Tepehua Background Information

Huehuetla Tepehua, which is a member of the Totonacan language family, is an under-documented, endangered, indigenous language that is spoken by approximately 1,000 people in Huehuetla, Hidalgo, in northeastern Mexico.

Huehuetla Tepehua is a polysynthetic, head-marking (Nichols 1986) language with two sets of person markers. Nominative person markers index subjects of transitive and intransitive verbs, (1a) and (1b) respectively, while accusative person markers index objects (1c). Independent pronouns, such as *yuch* in (1c), rarely occur.

- (1) a. maa jantu ta'uputun palay ju xa'akanit  
maa jantu ta-'u-putun palay ju xa-'akanit  
EVI NEG 3PL.SUBJ-eat-DESID anymore DET UNP-flesh  
'They didn't want to eat the meat anymore.'

[T0020: 032]

- b. tacha'an ju pulasiyaa  
ta-cha'an ju pulasiyaa  
3PL.SUBJ-arrive.there DET police  
'The police arrived.'

[T0055: 027]

- c. maa yuch laktitaymay ju t'akunin  
maa yuch lak-titayma-y ju t'aku'-nin  
EVI 3SG.PRN 3PL.OBJ-chase-INC DET woman-PL  
'It [the snake] chases after the women.'

[T0003: 005]

An absence of person-marking morphology on the verb indicates one or more third person singular arguments, depending on the transitivity of the verb, as seen in (2), where neither the third person singular subject nor the third person singular object is marked on the verb.

- (2) ch'ach'ilh           ju       lapanak       ju       xk'iw  
 ch'ach'i-lh           det     lapanak       det     'ix-k'iw  
 adorn-CMPL       DET    person       DET     3POS-tree  
 'The person decorated his/her alter.'

[Tpwllex: ch'ach'i]

The word order of the core clausal arguments is pragmatically determined (Mithun 1992), while the word order within particular phrases is relatively fixed; for example, the word order in a genitive phrase is shown in (3).

- (3) [DET POS-possessum (DET) possessor]

In a genitive phrase, the possessum must be affixed with a possessive prefix that indicates the person of the possessor, as seen below in (4). Furthermore, the possessum must be preceded by a determiner, while the possessor is optionally preceded by a determiner; thus, the example in (4a) has only one determiner, while the example in (4b) has two.<sup>1</sup>

- (4) a. [ju 'ix-chaqa' Kuulax]  
 DET 3POS-house Nicholas  
 'Nicholas' house'

[MNB15: 53]

- b. maaxtulh                           ju   lapanak   [ju 'ixqélhchat   ju   k'iw]  
 maa-xtu-lh                           ju   lapanak   ju   'ix-qélhchat   ju   k'iw  
 CAUS-be.external-CMPL   DET   person     DET   3POS-chicle   DET   tree  
 'The man extracted the tree's chicle.'

[Tpwllex: maaxtu]

Huebuetla Tepehua person marking follows a primary/secondary object pattern in which (i) the patient or theme argument is marked on the verb in a monotransitive clause (5a) and (ii) the dative or benefactive argument is marked on the verb in a ditransitive clause (5b) (Dryer 1986).<sup>2</sup> Thus, both clauses in (5) have a first person primary object. In (5a), the sole object is the primary object of its clause. In (5b), the first person benefactive object is the primary object and is marked on the verb, while the patient argument *kinkuxtaa* 'my sacks' is the secondary object.

- (5) a. jantu k'i'ut'i  
 jantu ki-'u-t'i  
 NEG 1OBJ-eat-2SUBJ  
 'Don't eat me!'

[T0058: 030]

- b. kimaqtamaqstalh [ju kinkuxtaa] ju Carmelo  
 ki-maqtamaqsta-lh ju kin-kuxtaa ju Carmelo  
 1OBJ-help.lift-CMPL DET 1POS-sack DET Carmelo  
 'Carmelo helped me lift my sacks.'

[MNB15: 52]

Huehuetla Tepehua has a class of reduced nouns that refer to body parts (see the Appendix) and that can be incorporated onto the verb (Mithun 1984). Each member of this class has a corresponding full-form, freestanding noun. (6a) shows a clause without noun incorporation, and (6b) shows a corresponding clause with noun incorporation.

- (6) a. ch'aqay ju **xmaka'**  
 ch'aqay-y ju 'ix-**maka'**  
 wash-INC DET 3POS-**hand**  
 'She<sub>i</sub> washes her<sub>i</sub> hands.'

- b. **makch'aqalh**  
**mak-ch'aqa-lh**  
**hand-wash-CMPL**  
 'She<sub>i</sub> washed her<sub>i</sub> hands.'

[Tpwlex: ch'aqay]

Finally, the presence of a benefactive/malefactive primary object argument in a ditransitive clause requires the affixation of an applicative morpheme onto the verb (7). The applicative suffix *-ni* increases the valency of the verb, thus allowing the extra clausal argument.<sup>3</sup>

- (7) a. waa kintalhijunilh p'ulhnan  
 waa kin-ta-lhijun-ni-lh p'ulhnan  
 FOC 1OBJ-3PL.SUBJ-order-APPL-CMPL first  
 'They ordered [a drink] for me first.'  
 [T0066: 052]
- b. 'alin ju 'ani' ju lhiistak laka tii xakijuniy  
 'alin ju 'ani' ju lhiistak laka tii xa-ki-jun-ni-y  
 there.is DET here DET guard PREP road PAST-1OBJ-tell-APPL-INC  
 "Here there is someone who guards the road," he would tell me.'  
 [T0022: 050]

In (7a) the transitive verb *lhijun* 'order' is suffixed with *-ni*, which licenses the first person benefactive argument, marked on the verb by *kin-*. In (7b) the matrix verb *jun* 'tell' is a transitive verb that is made ditransitive by the addition of *-ni*, which again licenses the first person benefactive argument that is marked on the verb by *ki-*.

### 3. External Possession Constructions

Huehuetla Tepehua utilizes two morphosyntactic strategies to convey external possession. The first strategy involves the affixation of the applicative suffix *-ni* to the verb stem (section 3.1), and the second strategy involves body part incorporation (section 3.2).

#### 3.1 Applicative External Possession Constructions

As mentioned above, the addition of the applicative suffix *-ni* to a transitive or intransitive verb increases the valency of the verb and references a benefactive or malefactive argument, as seen above in (7). When an applicative clause also contains a possessive phrase, then one potential interpretation of the clause is that of external possession, in which the possessor is also the benefactive or malefactive argument of the clause. In (8) the proper noun *Lola* is the possessor of the patient noun *tzi* 'daughter', as well as the malefactive argument of the derived, ditransitive verb. Since Huehuetla Tepehua is a primary object language, *Lola* (being the malefactive argument) acts as the primary object of the clause, while the possessed noun *'ixtzi* 'her daughter' functions as the secondary object.

- (8) Saanilh [ju 'ixtzi']<sub>SO</sub> [ju Lola]<sub>PO</sub> ju lapanak  
 saa-ni-lh ju 'ix-tzi' ju Lola ju lapanak  
 hit-APPL-CMPL DET 3POS-daughter DET Lola DET man  
 'The man hit Lola's daughter (and this affected Lola [BEN/MAL]).'  
 [Tpwllex: saa]

Due to the absence of morphology that indicates third person singular on the verb, the only empirical evidence for arguing that the example in (8) is an external possession

construction (and, thus, that *Lola* is a clausal argument) is the fact that a verb suffixed by *-ni* occurs in the same clause as a genitive phrase. However, examples in which the primary object of a ditransitive clause is first or second person are more revealing, as seen in (9), where the patient/secondary object of the ditransitive clause is the genitive phrase *ju kinherencia* ‘my inheritance’. The possessor of the patient is co-referential with a first person primary object, which is indexed on the verb by the prefix *ki-*, making this a clearer example of external possession than the example in (8).

- (9) Ju kimpay kimokoonilh [ju kinherencia]<sub>SO</sub>  
 ju kim-pay ki-maqaw-ni-lh ju kin-herencia  
 DET 1POS-father 1OBJ-leave-APPL-CMPL DET 1POS-inheritance  
 ‘My father left me [BEN] my inheritance.’  
 [Tpwxlex: maqaw]

The fact that the external possessor is marked on the possessum as well as on the verb in Tepehua is not problematic. According to Payne and Barshi (1999), double-marking is common cross-linguistically in external possession constructions (p. 3).

The external possession reading is available only for two types of applicative clauses in Huehuetla Tepehua: transitive clauses in which the patient argument is possessed, shown in (8) and (9), and unaccusative clauses in which the theme argument is possessed, shown in (10) below.

In (10) the intransitive, unaccusative verb *paataju* ‘fall’ is made transitive by the addition of *-ni*. The possessed noun *xakanit* ‘their flesh’ is the subject/theme of the clause, while *lapanak* functions as a malefactive primary object argument of the clause.

- (10) maa nanaa xpaatajuniy [ju xakanit] [ju lapanak]<sub>PO</sub>  
 maa nanaa 'ix-paataju-ni-y ju xa-'akanit ju lapanak  
 EVI much PAST-fall-APPL-INC DET UNP-flesh DET person  
 ‘Flesh fell off of the people (and this affected them [MAL]).’<sup>4, 5</sup>  
 [T0057: 011]

The applicative form of external possession is also found in Tlachichilco Tepehua, a sister of Huehuetla Tepehua, as seen below in (11).

- (11) ki-xka-ni-y ki-maka:  
 1OBJ-hurt-DAT-IMPF 1POSS-hand  
 ‘My hand hurts to me.’  
 (Cf., xka-y ki-maka:, ‘My hand hurts.’)  
 [Watters 1988: 139 (171)]

In the example above, the first person possessor of the theme argument, *maka:* ‘hand’ is co-indexed on the verb as the primary object and fills the role of malefactive argument of the clause.

It should be noted that when taken out of context, most sentences (from both narratives and elicitation) can have several interpretations (especially when the arguments are all third person singular) because of Huehuetla Tepehua's flexible word order. Thus, other possible readings for the examples in (8) and (10) are given in (8') and (10'), respectively. None of these has an external possession reading.

(8') 'The man hit [Lola's daughter] (and this affected someone else).'  
'[Lola's daughter] hit the man (and this affected someone).'6

(10') The people's flesh fell off (and this affected someone else).'

All of the readings in (8') and (10') have a benefactive or malefactive argument that is not overtly named.

### 3.2 Body Part External Possession

The second type of external possession construction in Huehuetla Tepehua is found when a body part is incorporated into a verb, and the possessor of the body part functions as an argument of the clause. This is the same as Mithun's (1984) Type II Noun Incorporation, and it serves to foreground the possessor of the body part. In (12) below, the semantic theme *laka-* 'body' is prefixed onto the verb, thus foregrounding its possessor—*lapanák* 'people'—and allowing the possessor to serve as the subject/theme argument of the unaccusative clause.

(12) maa waa **lakap'**uch'ilh ju lapanak  
 maa waa **laka-p'**uch'i-lh ju lapanak  
 EVI FOC **body-rot-CMPL** DET person  
 'The people body-rotted.'<sup>7</sup> or  
 'The people's bodies rotted.'

[T0057: 010]

Like Mithun's (1984) Type II Noun Incorporation, body part external possession in Huehuetla Tepehua does not change the valency of the clause, as seen below in (13).

(13) a. ch'ilhch [ju lapanak]<sub>SUBJ</sub> [ju 'ixkawaayuu]<sub>PO</sub>  
 ch'i-lh-ch ju lapanak ju 'ix-kawaayuu  
 tie-CMPL-PUNC DET person DET 3POS-horse  
 'The man tied his horse.'

- b. **katuch'ilhch** [ju lapanak]<sub>SUBJ</sub> [ju xwaakax]<sub>PO</sub>  
**katu-ch'i-lh-ch** ju lapanak ju 'ix-waakax  
**ear-tie-CMPL-PUNC** DET person DET 3POS-cow  
 'The man horn-tied his bull.'  
 'The man tied his bull's horns.'

[Tpwllex: ch'i]

The sentence in (13a) is a simple monotransitive clause in which the agent *lapanak* 'person' is the subject and the patient *'ixkawaayuu* 'his horse' is the object; it is *not* an example of external possession. The sentence in (13b) is an equivalent external possession construction in which the semantic patient *katu-* 'ear' is prefixed onto the verb, allowing its possessor to serve as the grammatical patient argument. Like (13a), the clause in (13b) is monotransitive with an agentive subject and a patient object.

Whereas the applicative external possession strategy allows only the possessor of a patient or theme to be encoded as a direct dependent of the verb, the body part external possession strategy allows the possessor of a theme (12), a patient (13b), or a locative (14) to function as a core argument.

In (14), the directional prefix *lhii-* derives a ditransitive stem from the monotransitive root *paxa-* 'to bathe' and, thus, licenses a third (locative) core argument. The body part *ch'an-* 'foot' is prefixed onto the verb stem and does not have core argument status, while its possessor, *xtzi'* 'his daughter', acts as the locative argument onto which the coffee was spilled.

- (14) **lhiih'ampaxaa** [ju kapen] [ju Miikii] [ju xtzi']<sub>LOC</sub>  
**lhii-ch'an-paxa-V** ju kapen ju Miikii ju 'ix-tzi'  
**DIR-foot-bathe.VT-CMPL** DET coffee DET Miguel DET 3POS-daughter  
 'Miguel foot-spilled coffee on his daughter.'  
 'Miguel spilled coffee [on his daughter's foot]<sub>LOC</sub>.'

[Tpwllex: paxa]

Body part external possession is also found in Tlachichilco Tepehua, as seen below in (15).

- (15) a. **xka-y** ki-maka:  
 hurt-IMPF 1POSS-hand  
 'My hand hurts.'
- b. **k-mak-xka-y**  
 1SUB-hand- hurt-IMPF  
 'My hand hurts.'

[Watters 1988: 223 (346a-b)]

The intransitive sentence in (15a) is not an example of external possession; *maka*: ‘hand’ serves as the theme argument and the first person possessor is its dependent. The sentence in (15b) is the corresponding sentence with body part external possession in which the body part has been incorporated onto the verb and its first person possessor is now the subject/theme argument of the intransitive clause.

#### 4. EP in other Mesoamerican Languages

Cross-linguistically, external possession is quite common, and many languages in Mesoamerica exhibit this phenomenon, including Papantla Totonac (Totonacan), Oluta Popoluca (Mixean), and Tzotzil (Mayan).

##### 4.1 Papantla Totonac

Papantla Totonac is a member of the Totonacan language family and a cousin of Tepehua, so it is not surprising that it exhibits the two types of external possession that are found in Huehuetla Tepehua: the applicative type and the body part type (Levy 1999, 2002). However, Papantla Totonac has an additional type of external possession construction that—like body part external possession—involves the incorporation of nouns onto verbs (Levy 1999).

The applicative external possession construction increases the valency of the verb in both Tepehua, shown above in (8), (9), (10), and (11), and in Papantla Totonac, shown below in (16).

- (16) a. Juan<sub>i</sub> cuku-lh kuxtú nak ix<sub>i/j</sub>-lhi:tay (Pedro<sub>j</sub>)  
 Juan began-CPL weed LOC 3POS-furrow (Pedro)  
 ‘Juan<sub>i</sub> began to weed in his<sub>i/j</sub> furrow (of Pedro<sub>j</sub>)’
- b. Juan<sub>i</sub> cuku-*ni*-lh kuxtú nak ix<sub>i/j</sub>-lhi:tay (Pedro<sub>j</sub>)  
 Juan began-BEN-CPL weed LOC 3POS-furrow (Pedro)  
 ‘Juan<sub>i</sub> began to weed in his<sub>\*i/j</sub> furrow (for Pedro<sub>j</sub>)’

[Levy 2002: 178 (3c-d)]

The verb in (16a) is not affixed with the benefactive suffix *-ni*, and it has no benefactive argument. The sentence in (16b) is the externally possessed counterpart in which the verb carries the applicative suffix and *Pedro*, the possessor of the furrow, is the benefactive argument.

While both monovalent and bivalent verbs allow the applicative external possession construction in Tepehua, only bivalent verbs allow it in Totonac. The applicative external possession reading is available for the possessor of the patient and locative arguments in Totonac, and for the possessor of the patient, theme, and possibly the agent arguments in Tepehua. Finally, the applicative argument is affected (either positively or negatively) by the action of the verb in both languages (Levy 1999, 2002).

Papantla Totonac has two types of external possession construction that involve noun incorporation of a ‘part’: Type 1 and Type 2 (Levy 1999). Totonac’s parts class is

similar to Tepehua's body part prefixes: both are highly productive, bound morphemes that are "visible segmentations of a pre-existing whole" (p. 327) and that demonstrate sound symbolic alternations (see the Appendix for Tepehua examples).

The Totonac external possession construction that involves Type 1 Noun Incorporation, shown below in (17), is equivalent to the body part external possession constructions in Tepehua, shown above in (12), (13b), (14), and (15b).

(17) a. k-a-ma:-náw                      ka:-chuka:kti:-ni-ya:-n                      min-cha':xpá:-n  
1:INCL-go-PRG-PL:PRG    1:PL-amputate-BEN-ICPL-2:OBJ    2:POSS-leg-NMLZR  
'We are going to amputate your leg.'

b. k-a-ma:-náw                      ka:-cha':xpa:-chuka:kti:-ya:-n  
1:INCL-go-PRG-PL:PRG    1:PL-leg-amputate-ICPL-2:OBJ  
'We are going to leg-amputate you.'

[Levy 1999: 325 (1c-d)]

In (17a) the possessed body part *cha':xpá* 'leg', is the patient argument of the clause; in (17b) the body part has been incorporated onto the verb, and its second person possessor acts as the primary object and patient argument of the clause.

This construction is not valency-increasing in either language; rather it serves to foreground the possessor in both languages in that the incorporated nominal indicates the location of the verb's action. While Tepehua allows the possessor of the body part to function as the patient, theme, or locative argument of the clause, Totonac allows the possessor of the part to function only as the patient (Levy 1999).

The Totonac external possession construction that involves Type 2 noun incorporation, shown in (18), is not found in Huehuetla Tepehua.

(18) a. ka-wili-'                      tlamank                      nak                      kim-pa:xtú:-n!  
IMP-sit:CAUS-2    pot                      LOC                      1:POSS-side-NMLZR  
'Put the pot beside me!'  
(Lit: 'Put the pot by my side!')

b. ka-kim-pa:xtu:-wili-'                      tlamank!  
IMP-1:OBJ-side-sit:CAUS-2                      pot  
'Put the pot beside me!'

[Levy 1999: 326 (2a-b)]

In this construction, incorporation is from a locative phrase that contains a genitive phrase that involves a relational noun. The incorporated nominal expresses the reference point for the verb's action (rather than the *location* of the verb's action, as seen in Type 1 noun incorporation). Finally, the verb which hosts the incorporated element is restricted to one of Papantla Totonac's four existential verbs: *wi*: 'sit', *ya*: 'stand', *ma*: 'lie', *maka* 'hang' (Levy 1999).<sup>8</sup>



In (20a), *pu?pu* ‘belly’ is the theme/subject of the intransitive clause. In (20b), *pu?pu* has been incorporated onto the verb, and its first person possessor serves as the theme/subject of the clause.

In Oluta Popoluca the possessor of the incorporated noun may function as the agent, patient, theme, instrument, or locative argument, while in Tepehua it may function only as the patient, theme, or location. The two languages also differ with regard to the type of noun that can be incorporated. While Tepehua allows incorporation of body parts only, Oluta Popoluca allows anything inalienable (body parts, kinship terms, parts of a whole, and bodily secretions, such as blood) to be incorporated (Zavala 1999).

### 4.3 Tzotzil

Tzotzil has only one type of external possession—the applicative type (Aissen 1987), shown below in (21), where the second person possessor of the nominal ‘father’ is also an object argument of the clause, as evidenced by the second person absolutive prefix on the verb.

- (21) l-a-j-nup-be            ta    be    l-a-tot-e  
 CP-B2-A1-meet-IO    on    road    the-A2-father-CL  
 ‘I met your father on the road.’

[Aissen 1987: 126 (3)]

The applicative external possession construction in Tzotzil differs significantly from the corresponding construction in Huehuetla Tepehua in three significant ways. First, in Tzotzil the external possession reading is available for transitive verbs only, while in Tepehua, it is available for both transitive and intransitive verbs. Second, the external possession reading is available in Tzotzil only for the possessor of the patient argument, whereas it is available in Tepehua for the possessor of a patient, theme, or locative argument. Finally, in Tepehua, the applicative external possession construction emphasizes the affectedness of the applicative possessor argument; in other words, the applicative argument is affected either positively or negatively by the action of the verb. This is not the case in Tzotzil, where the applied argument may or may not bear some thematic relationship (additional to ‘possessor of the patient’) to the clause (Aissen 1987).

### 5. Summary

Two types of external possession construction are found in Huehuetla Tepehua: an applicative type and a body part incorporation type. With regard to the applicative type, an external possession reading is available for an applicative clause that has a genitive phrase, though the external possession reading is not the *only possible* reading for that clause. In the external possession reading, the possessor is the benefactive or malefactive argument of the verb. Finally, the external possession reading is available

only for the possessor of the patient argument of a transitive verb and for the possessor of the theme argument of an unaccusative verb.

With regard to the body part external possession construction in Huehuetla Tepehua, a body part is incorporated into the verb and the body part's possessor functions as a clausal argument. Under this type of external possession, incorporation of the body part does not change the valency of the clause, but it does foreground the possessor of the body part. Body Part external possession allows the possessor of the body part to function as one of the following core arguments: the patient, the theme, or the locative.

Unanswered questions about external possession constructions in Huehuetla Tepehua include: (i) Does the set of possible applicative external possession constructions include a case in which the possessor of the agent functions as the benefactive argument of the clause, such as [*Lola's daughter*] *hit the man (and this affected Lola)*? (ii) Is there an applicative external possession reading available for a clause with four participants such as *I gave the children your toy for you*? (iii) Is body part external possession possible for agents and/or instruments in Huehuetla Tepehua (like noun incorporation external possession is in Oluta Popoluca)? (iv) Is it possible for applicative external possession and body part external possession to co-occur in the same clause (e.g., *You horn-tied my bull for me.*)? (v) Where do the two external possession constructions found in Tepehua fit in the cross-linguistic hierarchies of external possession constructions? Answers to these questions will be revealed only through further fieldwork on Huehuetla Tepehua.

## Appendix

Table 1: Huehuetla Tepehua Body Parts

Reduced Noun	Full Noun	Gloss
'aq- ~ 'ak- ~ lakapaa-	'aqtzulh	head
'aqx- ~ 'akx- ~ 'aqxp'in-	'aqxp'un	shoulder, upper back, flat surface
ch'an- ~ tz'an-	ch'aja'	foot, leg, paw
ka- ~ kinka-	kinkati'	tip, point
katu-	'aqaxqolh	ear
kik- ~ kilhtu- ~ qelh-	kilh	mouth, edge
kinka-	lhiixin	nose
laka- ~ laqa- ~ laq-	lakatunaj	body
laq- ~ lak- ~ lakpuu- ~ laq(a)puu-	'ukxpu'	face
laq- ~ lak- ~ lakpuu- ~ laq(a)puu-	laqchulh	eye
mak- ~ maq-	maka'	hand, arm
muuntz'a- ~ muunti	muuntz'an	forehead
muusa-	muusan	groin
paka-	pakapu'	armpit
qaq- ~ kaq-	laqxtan	cheek
qaq- ~ kaq-	kaalhtz'an	jaw
staa-	puulakan	back
tan-	tampuu	stomach
tan-	tankilhak	chest
tii-	tiimus	lower back, buttocks, and hips

## Notes

\* The fieldwork seasons that provided these data were funded by a Fulbright García-Robles Fellowship, an NSF Dissertation Research Grant, and the Project for the Documentation of the Languages of Mesoamerica (which, in turn, was funded by NSF and National Geographic). I am deeply indebted to all of my Huehuetla Tepehua consultants, especially Nicolás Viguera Patricio, Antonio Viguera Huerta, Micaela Santiago Plata, Fidela Sevilla García, Juana Gutiérrez Patricio, and Laurencio Viguera Patricio. I want to thank Roberto Zavala for all of his assistance, direction, and advice, without which, this work could not have been done. I would also like to thank Paulette Levy, Nora England, Judith Aissen, Spike Gildea, Mark Brown, Alexis Palmer, and Lynda de Jong for their comments on and suggestions for this work; it goes without saying that all mistakes are my own. I use a practical orthography to represent the data; the symbols that differ from the IPA symbols are the following: VV = long vowel, C' = glottalized consonant, ' = glottal stop, x = /s/, tz = /ts/, ch = /tʃ/, lh = /ʎ/, y = /j/, and j = /h/. The Abbreviations are as follows: A A-Person Markers, ABS Absolutive, ANIM Animate, APPL Applicative, APPL2 Benefactive/Malefactive Applicative, B B-Person Markers, BEN Benefactive, CAUS Causative, CL clitic, CMPL Completive, CP completive aspect, CPL completive, DAT Dative, DESID Desiderative, DET Determiner, DIR Directional, DO Direct Object, EVI Evidential, FOC Focus, ICPL Incompletive, IMP Imperfective, IMPF Imperfective, INC

Incomplete, INCH Inchoative, INCL Inclusive, INV Inverse, IO Indirect Object, LOC Locative, NEG Negative, NMLZR Nominalizer, OBJ Object, PAST Past, PL Plural, POS Possessor, POSS Possessor, PREP Preposition, PRG Progressive, PRN Pronoun, PUNC Punctual, SG Singular, SUB Subject, SUBJ Subject, UNP Unspecified possessor, VT Transitive verb, 1 = first person, 2 = second person, 3 = third person.

<sup>1</sup> The number of the possessor is indicated on the possessum by a separate suffix, *-an*, not shown here.

<sup>2</sup> Primary objectivity (Dryer 1986):

	SecondaryObj	PrimaryObj
Ditrans	DO	IO
Monotrans		DO

<sup>3</sup> I have not found a single Huehuetla Tepehua verb root that is inherently trivalent; thus all of the ditransitive clauses shown in these examples have derived verb stems. (N.B., this does not mean that Huehuetla Tepehua does not have any inherently trivalent verb roots.)

<sup>4</sup> In Huehuetla Tepehua *lapának* ‘person’ has an irregular plural form that differs from the singular only in the placement of the accent: *lapanák*. This plural form translates as ‘gente’ (people), which is grammatically singular in Spanish; *lapanák* can be either grammatically singular or grammatically plural. In example (10), it is singular.

<sup>5</sup> The verb *paataju* is a monovalent verb with only one argument, the theme. The example in (i) below shows that the location from which the theme “falls” appears in an oblique prepositional phrase.

(i)	paatajuu	[ju	laka	k’iw]	[ju	lapanak]
	paataju-V	ju	laka	k’iw	ju	lapanak
	fall-CMPL	DET	PREP	tree	DET	person

‘The person fell from the tree.’

[Tpwllex: paataju]

<sup>6</sup> Another *potential* reading for (8) is ‘[Lola’s daughter] hit the man (and this affected Lola [BEN/MAL]).’ I do not know if this is a possible interpretation in Huehuetla Tepehua, and I will need to test this during my next field trip. If this interpretation is possible, this would mean that an external possession reading is also available for the possessor of the *agent* of a transitive verb.

<sup>7</sup> Again the noun *lapanák* is grammatically singular, even though it is glossed as ‘people’ in English.

<sup>8</sup> I have not found any Huehuetla Tepehua examples in which a body part is incorporated onto one of the four cognate existential verbs. Rather, body part external possession occurs only with verbs that are pragmatically prone to act on body parts.

## References

- Aissen, Judith. 1987. *Tzotzil Clause Structure*. Dordrecht: D. Reidel Publishing Co.
- Dryer, Matthew S. 1986. ‘Primary objects, secondary objects, and antitativity.’ *Language* 62(4): 808-845.
- Levy, Paulette. 1999. ‘”Where” rather than “what”: Incorporation of ‘parts’ in Totonac.’ *External Possession*, ed. by Doris L. Payne and Immanuel Barshi, 325-338. Amsterdam: John Benjamins Publishing Co.
- Levy, Paulette. 2002. ‘El applicativo dativo/benefactivo en totonaco de Papantla.’ *Memorias del VI Encuentro Internacional de Lingüística en el Noroeste*, ed. by Zarina Estrada Fernández and Rosa María Oriz Ciscomani, 175-194. Hermosillo, Sonora, Mexico: Departamento de Letras y Lingüística, Universidad de Sonora.
- Mithun, Marianne. 1984. ‘The evolution of noun incorporation.’ *Language* 60(4): 847-894.

- Mithun, Marianne. 1992. 'Is basic word order universal?' *Pragmatics of Word Order Flexibility*, ed. by Doris L. Payne, 15-61. Amsterdam: John Benjamins Publishing Co.
- Nichols, Johanna. 1986. 'Head-marking and dependent-marking grammar.' *Language* 62(1): 56-119.
- Payne, Doris L. and Immanuel Barshi. 1999. 'External possession: What, where, how, and why.' *External Possession*, ed. by Doris L. Payne and Immanuel Barshi, 3-29. Amsterdam: John Benjamins Publishing Co.
- Watters, James K. 1988. *Topics in Tepehua Grammar*. Doctoral Dissertation, UCLA.
- Zavala, Roberto. 1999. 'External possessor in Oluta Popoluca (Mixean): Applicatives and incorporation of relational terms.' *External Possession*, ed. by Doris L. Payne and Immanuel Barshi, 339-372. Amsterdam: John Benjamins Publishing Co.

## Some Tonogenetic Properties of Upriver Halkomelem\*

J.C. Brown

University of British Columbia

### 1. Introduction

Upriver Halkomelem is a Central Coast Salish language, spoken around Chilliwack, British Columbia; there are currently two speakers remaining. Upriver Halkomelem exhibits tonal or accentual phenomena whereas the other Salish languages don't (including related dialects Cowichan and Musqueam). Although very in-depth descriptive accounts exist, the ultimate problem lies in accounting for the tone or accents<sup>1</sup>, and trying to explain how they developed.

The goal of this paper is to characterize the Upriver Halkomelem pattern as an accentual system (as opposed to a tonal system). I propose that given the robust stress system already in existence in the language, and the light functional load that non-stress accent seems to play, these factors indicate that this system is still very early in its development. Granting this early stage of emergence, cases will be discussed where the tonal properties of the language are already distinctive in some areas, including some possible syntactic environments.

### 2. Data

In comparing the Upriver Halkomelem dialect with the Cowichan and Musqueam dialects, the generalization can be drawn that glottal stops in coda position delete, which often leaves the previous vowel lengthened and accented (see Elmendorf & Suttles 1960).

	<u>Musqueam/Cowichan</u>		<u>Upriver Halkomelem</u>	
(1)	síʔlə	'grandparent'	sí:lə	'grandparent'
(2)	həlíʔ	'alive'	həlí:	'alive'
(3)	t'áʔx <sup>w</sup>	'white fir'	t'á:x <sup>w</sup>	'white fir'
(4)	iséʔlə	'two'	isé:la	'two'
(5)	qáʔ	'water'	qá:	'water'
(6)	ʔ'aʔtθé	'that (it's that one)' <sup>1</sup>	ʔ'a:tθé:	'that (it's that one)'
(7)	sp'áʔəmʔ	'smoke' <sup>2</sup>	sp'á:ʔəm	'smoke'
(8)	χáʔqən	'otter'	χá:qəl	'otter'
(9)	χ <sup>w</sup> əχ <sup>w</sup> əyíʔye	'fly' <sup>3</sup>	χ <sup>w</sup> əχ <sup>w</sup> əyá:yε	'fly'
(10)	q <sup>w</sup> séʔyən	'gill net' <sup>4</sup>	q <sup>w</sup> sé:yəl	'gill net'
(11)	láʔθən	'dish'	lá:θəl	'dish'
(12)	q <sup>w</sup> áʔpəʔp	'devil's club'	q <sup>w</sup> á:pəʔp	'devil's club'

The Upriver Halkomelem pattern is not one of complete deletion of glottal stops, however, as glottal stop is retained in onset position:

(13)	ʔéšx <sup>w</sup>	‘hair seal’	ʔé:šx <sup>w</sup>	‘hair seal’
(14)	ʔeyʔiləm	‘clear up’	ʔeyi:ləm	‘clear up’
(15)	q <sup>w</sup> əʔáp	‘crabapple’	q <sup>w</sup> əʔáp	‘crabapple’
(16)	ʔíməθ	‘grandchild’	ʔí:məθ	‘grandchild’

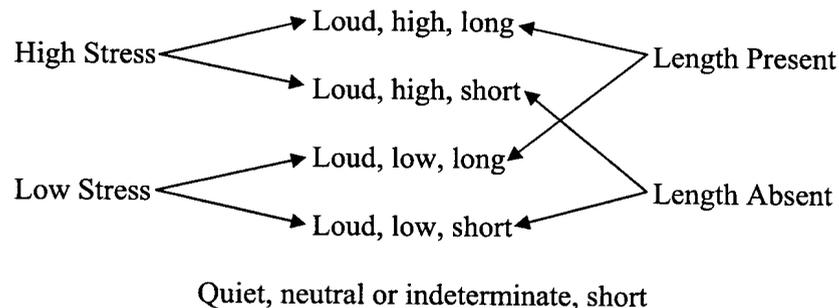
### 3. Previous Descriptive Accounts

To my knowledge, the only substantial accounts of tone in Halkomelem are Elmendorf & Suttles (1960) and Galloway (1991, 1993). These descriptive accounts will be discussed below.

#### 3.1. Elmendorf & Suttles (1960)

Elmendorf & Suttles provide an early account of the tonal properties of the language by contrasting it with neighboring Cowichan and Musqueam dialects. Their descriptions can be generalized as follows: features of length, high/low, and loudness give certain combinatorial possibilities, some of which are present as acoustic features associated with vowels in Upriver Halkomelem.

(17) Acoustic features associated with vowels:



According to Elmendorf & Suttles, this derives something of a three-way contrast in the language between high, low, and neutral stress. Of special note is Elmendorf & Suttles' speculation that the high and low stresses compose a system of pitch accents.

#### 3.2. Galloway (1991, 1993)

As described by Galloway (1991, 1993), there are two marked and one unmarked tone in Upriver Halkomelem. The marked high tone has three allophones, while the low tone has no allophony. As Galloway states: "The Upriver dialects of Halkomelem have three degrees of phonemic pitched stress or phonemic tone: / / high pitch-stress or high tone, / / mid pitch-stress or mid tone, and (unmarked) low pitch unstressed or low tone." (1993:38).

- (18) H allophones  
 i) / / → [ <sup>6</sup> ] (extra high level)

- ii) / / → [ <sup>ˆ</sup> ] (high-falling)
- iii) / / → [ <sup>ˆ</sup> ] (high level)

(19) L (no complex allophony)

Galloway goes on to describe these allophones:

“/ / → [ <sup>ˆ</sup> ], loud stress with high and level pitch, which seems to be about the musical interval of a sixth above the unstressed low pitch. This allophone occurs only on short vowels immediately preceding a weakened word boundary” (1991:1).

“/ / → [ <sup>ˆ</sup> ], loud stress with high-falling pitch which starts about the musical interval of a fifth above unstressed low pitch or low tone and falls to low pitch. This allophone occurs on long vowels in word-final syllables and in free variation with [ <sup>ˆ</sup> ] on the last long vowel in the word in nonfinal syllables” (2).

“/ / → [ <sup>ˆ</sup> ], loud stress with high and level pitch varying between about a fifth and a fourth musical interval above low pitch; this allophone occurs elsewhere” (2).

“/ <sup>ˆ</sup> / has no complex allophony; it is level pitch (about a third above unstressed low pitch) with loudness ranging from moderate to loud” (2).

What is immediately striking about Galloway’s descriptions are his great attention to detail, making note of the closest musical intervals and the complexity in allophonic variation.

With these descriptive accounts outlined, we can next move on to some possible explanations for the genesis of the tonal accents.

#### **4. Possible Explanations**

One of the earliest attempts at explaining the tonogenetic properties of Upriver Halkomelem is provided by Elmendorf & Suttles (1960). Elmendorf & Suttles suggest three possible explanations. They suggest that either the accent is an archaism, that it is an innovation or internal pattern change connected with other phonologic changes, or that the accentual phenomenon is an innovation through contact (possibly with Interior Salish languages). Each of these possibilities will be considered below.

##### **4.1. Archaism**

If the accentual phenomenon is an archaism, this would require the Upriver Halkomelem tonal patterns to be older than neighboring dialects and/or languages. This is extremely unlikely since no other Salish language retains any non-stress accentual properties.

Furthermore, the loss of glottal stop (which remains in the Cowichan and Musqueam dialects) can be considered a textbook trigger to tonogenesis (cf. Hombert 1978 and see section 4.4 below), indicating this is a very recent development.

It is also interesting to note some of Swadesh's (1952) remarks on the phonological changes that have taken place in the diachrony of the Halkomelem dialect continuum. Swadesh notes that the replacement of all /n/ segments by /l/ must be innovative in Upriver Halkomelem because the Nanaimo dialect retains both segments. Although this is somewhat tangential to the topic of the stress/pitch phenomena in Upriver Halkomelem, the same principles must be true. What is furthermore striking about Swadesh's article is the absence of any discussion of tonal phenomena in the Upriver dialect, which indicates that the emergence of such a system may be as recent as Swadesh's fieldwork. The brevity of Swadesh's treatment of Halkomelem in a much larger pan-Salish perspective, however, leaves one only to speculate on such matters.

#### **4.2. "Contact Innovation"**

Elmendorf & Suttles suggest that the system of pitch accents may be an innovation through contact with Interior Salish languages. They cite evidence to suggest that this is a possibility in the form of Interior Salish loans in Halkomelem which make use of phonological segments not in the Halkomelem inventory. Again, this seems an unlikely possibility as there are no other Salish languages (Interior or Coast) with non-stress accentual properties.

#### **4.3. Areal Feature**

Another possibility, related to contact innovation but not discussed by Elmendorf & Suttles, is that the accentual properties are an areal feature. Under this view, the phonologically conditioned tones would be a property shared by many related or unrelated neighboring languages. Observations by scholars on tonogenesis such as Matisoff (1973:87-89) indicate that this is an entirely possible scenario cross-linguistically.

Upriver Halkomelem is neighbored by Interior Salish languages, which don't make use of these features (see above). Another possible candidate for sharing an areal feature would be Nicola, but Nicola went extinct early in the century. Since the Musqueam and Cowichan forms haven't been affected, nor the Interior Salish languages which are also neighboring the Nicola area, this possibility must be ruled out.

#### **4.4. Internal Innovation**

The strongest possible explanation is that it is an internal innovation (see Elmendorf & Suttles 1960). The patterns in (1-16) above are consistent with patterns of tonogenesis attested in the literature (cf. Hombert 1978, Ohala 1973, Matisoff 1973, Mithun 1999:25, Thurgood 2002, Yip 2002). Although the exact mechanisms by which high and low accents emerged and behave by are not well understood, it seems clear that internal innovation is the most plausible general explanation for their emergence. This will be the assumption for the duration of the paper.

### 5. Tone vs. Accent

The next immediate question concerns what type of system Upriver Halkomelem actually is: a tone or accentual system. I have been freely referring to the Upriver Halkomelem pattern as an accent system, and will further clarify this usage below.

Mithun (1999) has noted that there are 2 types of ‘tone’ system prominent in North America: tone systems and pitch accent systems. Under the traditional viewpoint (Trubetzkoy 1969), tonal systems are paradigmatic in nature, allowing for single syllables to contrast in value for tone. In contrast, accent systems are conceived of as syntagmatic, whereby an accented position is such because it stands in a relation to a non-accented position.

It has further been noted that there is not always a clear distinction between stress- and non-stress accent or tone (McCawley 1970, 1978). In order to determine what type of system Upriver Halkomelem is, the ‘tonal’ properties of the language, and the contrastiveness of tones must be explored. This includes looking at minimal pairs (or the ‘functional load’ that tone plays) and some facts concerning optionality. Furthermore, inconsistencies between fieldworkers will be addressed. The possibility of tone playing an emerging role in the morphosyntax and the predictability of tone will conclude this section.

#### 5.1. Functional Load

One possible diagnostic for what the tonal elements are doing is to look at what kind of functional load it carries. Languages closer on the continuum to stress- or pitch-accent will have very few minimal pairs, or tone will carry a low functional load (Beckman 1986). This seems to be the case for Upriver Halkomelem, which has probably less than 5 minimal tone pairs.

Minimal Pairs:

(20)	[q <sup>w</sup> æ:l]	‘mosquito’	
	[q <sup>w</sup> ə:l]	‘to speak’	(Galloway 1991:4)

Furthermore, there is an unequal distribution of high vs. low tones, which may indicate the influence that stress has on the quality of the tones. The prosody of the language also plays an important role in terms of how tone would be functional. Since Upriver Halkomelem is a polysyllabic language, it is not likely to move from a stage of being atonal to being completely tonal (cf. Matisoff 1973).

#### 5.2. Optionality

Optionality plays another diagnostic role in determining what type of system is at work in Upriver Halkomelem. In many cases high or low accent is optional on words:

(21)	hèm ~ hé:m	‘finished (of a story), over’	(22)
	tè:m ~ té:m	‘shout, yell, holler’	
			(Galloway 1993:306)

- (22) χέ:m ~ χέ:m 'cry, weep'  
 (23) wèy:əlàs ~ wéy:əlàs ~ wéy:ələs ~ wéyələs 'tomorrow'

(Galloway 1991:3)

This optionality suggests that the system at hand is not tonal, as a paradigmatic relation between tones would force a contrast between such forms.

There are also radical differences between the transcriptions of fieldworkers. Elmendorf & Suttles (ES), Jimmy Harris (JH) and Brent Galloway (BG) will often conflict in transcriptions. As Galloway has noted:

“Another difference between Galloway 1977 and ES and JH is our frequent disagreement with stress patterns involving / ˘ /. ES may have ˆ on one word which JH cites as ˘ and I have as ˆ ; another example may show ES ˆ, JH ˆ, BG ˘, or ES ˘, JH ˆ, BG ˆ. It seems likely that at least some of these differences are idiolectal or free variation” (Galloway 1991:3).

There is also variation across speakers<sup>5</sup> (which may or may not play a large role in the variation between fieldworkers). For instance, the pronunciation of a place name is given in (24), and the word for ‘whale’ in (25).

- (24) θ'əwé:lí (BJ) ~ θ'əwè:li (AC) ~ θ'əwè:li 'Soowahlie (place name)'

- (25) qʷəl:əs (prompted, AC) ~ qʷəl:əs (BJ, ES) ~ qʷəl:əs (JH) 'whale'  
 (Galloway 1991:3)

Likewise, there is also variability in vowel length.

- (26) t'él:lk'wəm ~ t'élq'wəm 'warm, lukewarm (of food, drink)'  
 (Galloway 1993:304)

- (27) méqe (AC) ~ mé:qe (BJ) 'snow on the ground' (Galloway 1991:3)

There are also cases where no length occurs where length is predicted, or length occurs where it is not predicted to (see Kava 1972).

### 5.3. Morphosyntax

Despite the low contrastive load in terms of minimal pairs, there are certain morphological and syntactic processes which are sensitive to tone. For example, there are allomorphs which are tone pairs:

- (28) -tá:l vs. -tà:l (260)<sup>6</sup>

(29) -là:mət vs. -lá:mət (257)

Transitivity may also be sensitive to tonal differences. The unergative suffix -əm (with some vowel variation) takes a low tone:

(30) k'wíyè:m 'refuse, be stingy'

While the transitive suffix takes a high tone:

(31) k'wíyé:t 'refuse s.o., s.t.' (Galloway 1993:251)

So even though the functional load carried by tone is relatively low in lexical terms, it may be carrying a larger morphosyntactic load as it emerges. There is some allomorphy based on vowel quality, and there are numerous examples that don't conform to these generalizations, but despite these facts, the morphosyntactic uses of tone seem a rich area for future research.

#### 5.4. Predictability

As Elmendorf & Suttles note, the pitch accents in Upriver Halkomelem are unpredictable based on cognate forms. Furthermore, the tones appear to be unpredictable based on the properties of the glottal stop, as well. It has been claimed that the loss of post-vocalic glottal stop consonants has the exclusive effect of raising pitch on previous vowels (Matisoff 1973:76, Ohala 1973:3, Hombert et al. 1979). In Upriver Halkomelem, the loss of post-vocalic glottal stops not only raises pitch (32-37), but it also lowers it in different environments (34-36) (see also Mithun 1999, where counterexamples to the above claim are provided):

	<u>Musqueam/Cowichan</u>		<u>Upriver Halkomelem</u>	
(32)	iséʔlə	'two'	isé:la	'two'
(33)	qáʔ	'water'	qá:	'water'
(34)	χácaʔ, χáceʔ	'lake'	χácə	'lake'
(35)	sáyʔsəyʔ	'fear (to be afraid)'	sí:sì:	'fear (be afraid)'
(36)	t'éləwʔ	'arm'	t'éləw:	'arm'
(37)	háyʔq <sup>w</sup>	'fire'	hé:yəq <sup>w</sup>	'fire'

#### 6. Conclusions and Prospects for Future Research

Although there are a very few cases where there is a clear contrast with minimal pairs, "tone" is still in the stages of being a manifestation of stress or pitch accent. The patterns discussed above indicate that Upriver Halkomelem has undergone some degree of tonogenesis, and that this pattern of consonant-tone interactions stands as a counterexample to the claim that the loss of final glottal stop produces only pitch-raising effects.

There are numerous areas of the phonology of Halkomelem which prove promising in the sense that they are interrelated with the problem of tone. In particular, one question that must be addressed is how the stress patterns of Upriver Halkomelem differ with respect to neighboring Musqueam and Cowichan. As Thompson (2004) has pointed out, despite the surface similarities between the dialects, the fundamental difference lies in the lengthening of stressed vowels in the Upriver dialect (see also Kava 1972). This leads directly to the most immediate of questions, namely, What are the interactions between stress and tone?

Another question to be addressed is, What are the effects of loss of glottal stop or glottalization on glides and resonants? It is less than clear at this point what effect tonogenesis has had on these phonemes, but figuring that out would certainly place the Upriver patterns in a much more coherent cross-Salish phonological context.

Finally, intensive instrumental studies are needed in order to determine what the acoustic correlates of stress and tone are in Halkomelem. In the spirit of the comparative work of Elmendorf & Suttles (1960), cross-dialectal studies seem the most promising in this regard. In the words of Suttles (1994:387), "A further analysis of Halkomelem vowels may be in order, and we certainly need a study of the development of pitch."

### Notes

\*Thanks to Leora Bar-El, Lisa Bennett, Strang Burton, Bryan Gick, Dan Hintz, Karsten Koch, Tyler Peterson, Douglas Pulleyblank, James Thompson, and Martina Wiltschko. Funding was provided by SSHRC grant #410-2002-0041 awarded to Douglas Pulleyblank (principle investigator). All errors remain the author's.

<sup>1</sup>It is not straightforward how to handle the terminology surrounding 'stress', 'accent' and 'tone'. For much of this paper, I will use 'accent', 'non-stress accent' 'pitch-accent' and 'tone' interchangeably to refer to some grammatical use of pitch, while attempting to stay neutral with regard to what type of overall system (accent vs. tone) is being described. Discussion of whether the language is a pitch accent or tonal system will be discussed in detail in section 5.

<sup>2</sup>The Cowichan form is nip'εʔtθéyʔ.

<sup>3</sup>Cowichan sʔ'éyəq'əmʔ.

<sup>4</sup>Cowichan χ<sup>w</sup>əχ<sup>w</sup>əyáʔyε.

<sup>5</sup>Cowichan q<sup>w</sup>áseʔyən.

<sup>6</sup>BJ= Bob Joe (Chilliwack dialect), AC= Amy Cooper (Chilliwack dialect), ES = Elmendorf & Suttles, JH= Jimmy Harris, BG = Brent Galloway.

<sup>7</sup>Although Galloway notes that á: = ablaut aspect.

### References

- Beckman, Mary E. 1986. *Stress and non-stress accent*. Dordrecht: Foris.  
 Elmendorf, William and Wayne Suttles. 1960. Pattern and change in Halkomelem Salish dialects. *Anthropological Linguistics* 2:1-35.

- Galloway, Brent. 1991. A Salish language with tone and other interesting phonological complexities. Paper presented at 30<sup>th</sup> CAIL, 90<sup>th</sup> AAA Annual Meeting, Chicago, IL.
- Galloway, Brent. 1993. *A grammar of Upriver Halkomelem*. Berkeley: University of California Press.
- Harris, Jimmy Gene. 1960. The phonology of Chilliwack Halkomelem. M.A. thesis, University of Washington.
- Hombert, Jean-Marie. 1978. Consonant types, vowel quality, and tone. In V. Fromkin (ed.), *Tone: a linguistic survey*. New York: Academic Press, pp. 77-111.
- Hombert, Jean-Marie, Ohala, John & William Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55:37-58.
- Kava, Tiiu. 1972. A consideration of historical implications in an idiosyncratic development of vowel length in the Chilliwack dialect. *Working Papers in Linguistics* 4(3):113-126. Department of Linguistics, University of Hawaii.
- Matisoff, James A. 1973. Tonogenesis in southeast Asia. In L. Hyman (ed.), *Consonant types and tone. Southern California Papers in Linguistics* 1:71-96.
- McCawley, James D. 1970. Some tonal systems that come close to being pitch accent systems but don't quite make it. In *Papers from the sixth regional meeting of the Chicago Linguistic Society*, pp. 526-532.
- McCawley, James D. 1978. What is a tone language? In V. Fromkin (ed.), *Tone: a linguistic survey*. New York: Academic Press.
- Mithun, Marianne. 1999. *The languages of Native North America*. Cambridge: Cambridge University Press.
- Ohala, John J. 1973. The physiology of tone. In L. Hyman (ed.), *Consonant types and tone. Southern California Papers in Linguistics* 1:1-14.
- Suttles, Wayne. 1994. Review of Brent D. Galloway's *A grammar of Upriver Halkomelem*. *Anthropological Linguistics* 36:386-389.
- Swadesh, Morris. 1952. Salish phonologic geography. *Language* 28(2):232-248.
- Thompson, James. 2004. Stress in Upriver Halkomelem. Ms, University of British Columbia.
- Thurgood, Graham. 2002. Vietnamese and tonogenesis: revising the model and the analysis. *Diachronica* 19:333-363.
- Trubetzkoy, N.S. 1969. *Principles of phonology*. Translated by Christiane A.M. Baltaxe. University of California Press.
- Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

jcb@interchange.ubc.ca

# Voice in Iquito: Is there an inverse?<sup>1</sup>

Lynda de Jong Boudreault  
University of Texas at Austin

## 1 Introduction: Two Studies

The goal of this paper is to describe two contradicting studies regarding inverse voice in Iquito. The first is a pilot study, in which semantically transitive constructions were evaluated using quantitative analysis to identify voice constructions in Iquito. During the pilot study, promising results indicated that certain constructions that had previously been analyzed structurally as active voice constructions with an OV word order were in fact inverse voice constructions. The results pointed to two adpositions – *hina* a locative meaning ‘in, within’ and *anuura* a directional meaning ‘toward’ – as structural indicators of an inverse voice construction in Iquito, shown in (1).

1. *hina* locative adposition meaning ‘in, within’  
*anuura* directional adposition meaning ‘toward, at’

The second is the follow-up study that yielded results contrary to those of the pilot study. That is, *hina* and *anuura*, the two adpositions, are no longer seen as structural indicators of inverse voice.

## 2 Iquito

Iquito is an indigenous language spoken in the northern Peruvian Amazon in the town of San Antonio on the Pintuyacu River (shown on the map). There are approximately 25 speakers, all of whom are more than 50 years old.

Map 1. Peru: San Antonio del Pintuyacu



### 3 Data

The primary source of data presented here comes from narratives recorded and transcribed during the summers of 2002 and 2003 in San Antonio by me and my colleagues through work on the Iquito Language Documentation Project<sup>2</sup> (ILDLP). The secondary source comes from research done by the Summer Institute of Linguistics in the 1960s in the same village.

### 4 Quantitative Analysis

Quantitative analysis provides an empirical method for evaluating the degree of topicality of any given NP in discourse (Cooreman, 1983). Using this method, the topicality of an NP is measured in terms of *referential distance* (RD) and *topic persistence* (TP). Following Cooreman (1983), referential distance measures continuity in terms of the number of clauses which intervene between the original mention and the new mention. Referential distance looks back in counting an NP's referents. Topic persistence measures the topic prominence based on the number of subsequent mentions of an NP in discourse. Topic persistence looks forward for an NP's referents.

These measures are useful in defining the various voice constructions in that they allow us to evaluate the relations between the agent and patient in those constructions with respect to degree of topicality. Following Cooreman (1987) and Givón (1994), the relations between the agent and patient in the four main voice constructions (active, inverse, passive and antipassive) are illustrated in table 1. For active/direct voice constructions, the agent is more topical than the patient, which retains some level of topicality. Conversely, in an inverse voice construction, the patient is more topical than the agent, but the agent retains considerable degree of topicality. In passive constructions, the patient is more topical than the agent, and the agent is non-topical. In *antipassive* constructions the reverse is true. The agent is more topical and the patient, non-topical.

Table 1.

<i>Voice</i>	<i>Relative Topicality</i>
active/direct	AGT > PAT
inverse	AGT < PAT
passive	AGT << PAT
antipassive	AGT >> PAT

#### 4.1 Methodology

In order to obtain these referential distance and topic persistence measures, I looked at the third person agents and patients in semantically transitive constructions. To obtain the referential distance I looked back to the 3 clauses preceding that of the NP being counted. The NP received: one (1) if it appears in the preceding clause; two (2) if it appears two clauses prior; three (3) if it appears three clauses prior; and greater than three (>3) if it does not appear within the previous three clauses or if it is a new mention

To count the topic persistence of an NP I looked ahead to the 10 clauses following that of the NP being counted. If an NP is mentioned in one of the subsequent 10 clauses

it receives a one (1); in two of the subsequent 10 clauses, it receives a two (2); and so on until up to 10 clauses. The maximum count an NP can have is 10. I calculated the RDs and TPs for both agents and patients in each type of construction I observed initially in the text. Those were: Active constructions with VO word order; active constructions with OV word order; and antipassive constructions<sup>3</sup>.

The percentages are evaluated in terms of the topicality of the NP. For Referential Distance, if an NP receives a value of 1, 2 or 3, this indicates that the NP has a high degree of topicality. A value greater than three indicates low topicality. For Topic Persistence, if an NP receives a value of 0, 1 or 2, this indicates a low topicality. Values greater than 2 (that is, more than two mentions) indicate a high degree of topicality.

## 4.2 Voice and topicality in Kutenai and Akatec

### 4.2.1 Kutenai

Table (2) shows the referential distance measures of agents and patients in Kutenai. If we look at the direct voice constructions, we can see that 89% of agents are highly topical, that is, they have mentions within three clauses preceding it them. Compare that to the 52% of patients that are highly topical.

However, observe that for the inverse voice constructions, 97% of patients were highly topical. This is substantially higher than the 54% for agents. These measures correspond to the topicality relations of agents and patients illustrated in table 1.

**Table 2. Percentage of Agents and Patients with RD values 0-3 (High Topicality) and >3 (Low Topicality) in Kutenai (reformatted from Dryer 1994: 75)<sup>4</sup>**

<i>Voice</i>	<i>Agents</i>		<i>Patients</i>	
	<i>0-3</i>	<i>&gt;3</i>	<i>0-3</i>	<i>&gt;3</i>
Direct	89	11	52	48
Inverse	54	46	97	3

In terms of the topic persistence measures for Kutenai (in table 3), observe that for the direct voice constructions, 71% of the agents are mentioned in more than two out of ten clauses that follow, indicating a high degree of topicality. In contrast, only 38% of the patients have a high level of topicality. Now, the percentages for the inverse constructions show that 77% of the patients are highly topical while only 49% of agents are highly topical. Again these percentages correspond to those illustrated in table 1.

**Table 3. Percentage of Agents and Patients with TP values 0-2 (Low Topicality) and >2 (High Topicality) in Kutenai (reformatted from Dryer 1994: 78-79)**

<i>Voice</i>	<i>Agents</i>		<i>Patients</i>	
	<i>0-2</i>	<i>&gt;2</i>	<i>0-2</i>	<i>&gt;2</i>
Direct	29	71	62	38
Inverse	51	49	23	77

#### 4.2.2 Akatek

Turning to the data for Akatek in tables 4 and 5, we can see these correlations for antipassives and passives. For antipassives in (4a), the percentage for agents with high topicality is 92.9%. The percentage of patients with a low topicality is 100%. The percentages for the passives show the converse. In (c), 81% of patients are highly topical, compared to the 92.6% of agents that have a low degree of topicality.

**Table 4. Percentage of Agents and Patients with RD values 0-3 (High Topicality) and >3 (Low Topicality) in the various Akatek Voice Constructions (Zavala 1997)**

<i>Voice Construction</i>	<i>Agent</i>		<i>Patient</i>	
	<i>0-3</i>	<i>&gt;3</i>	<i>0-3</i>	<i>&gt;3</i>
a. Active direct: Active direct (V-first)	91.5	8.5	54.6	45.4
b. Antipassive: Absolute antipassive	<b>92.9</b>	7.1	--	<b>100</b>
c. Passive: Impersonal passive	7.4	<b>92.6</b>	<b>81.4</b>	18.6
d. Inverse: Agented passive	70.5	29.5	77.4	23.6
Inverse	77.8	22.2	88.9	11.1

The same correlations can be seen in the topic persistence measures in table 5. For antipassives in (b), 100% of the patients show low topicality while 97.6% of agents show high topicality. For passives in (c), 63% of patients are highly topical as compared to the 96.3% of agents which are of low topicality.

**Table 5. Percentage of Agents and Patients with TP values 0-2 (Low Topicality) and >2 (High Topicality) in the various Akatek Voice Constructions (Zavala 1997)**

<i>Voice Construction</i>	<i>Agent</i>		<i>Patient</i>	
	<i>0-2</i>	<i>&gt;2</i>	<i>0.2</i>	<i>&gt;2</i>
a. Active direct:				
Active direct (V-first)	7.8	92.2	59	41
b. Antipassive:				
Absolutive antipassive	2.4	97.6	100	--
c. Passive				
Impersonal passive	96.3	3.7	37	63
d. Inverse:				
Agented passive	29.5	70.5	17.6	82.4
Inverse	25	75	33.4	66.6

Again, these proportions are all commensurate with those illustrated in table 1.

## 5 Iquito Background: Word Order, Voice Constructions and Adpositions

Before going on to discuss the measures for Iquito, I want to provide some background information on word order, the apparent voice constructions and the locative and directional adpositions.

Iquito is nominative/accusative language with two basic word orders: SVO and SOV. Initially, based on elicited data, SOV word order was seen only in future tense constructions whereas SVO was seen everywhere else. In examples (2) and (3), observe the word order for simple constructions is SVO.

### 2. Present tense:

**S                    V                    O**  
 Juaa asaa samuukwaati  
 Juan eat plátano  
 Juan is eating plátano.

### 3. Past tense:

**S                    V                    O**  
 aasi pikuu.kura                    iita.ka  
 rain to wet.PAST                    house.PLU  
 The rain wet the houses (yesterday).

However, in the pair of examples in (4), note the word order for a simple construction in a future tense/aspect is SOV in (4a) but the VO order in (4b) is ungrammatical.

## 4a. Future tense:

	<b>S</b>	<b>O</b>	<b>V</b>
amikaaka	Pedro	Maria	kiwaku.r++
tomorrow	pedro	maria	hug.FUT

Tomorrow Pedro is going to hug Maria.

## 4b Ungrammatical future:

	<b>S</b>	<b>V</b>	<b>O</b>
*amikaaka	ki.niki.r++	nuu	
tomorrow	1s.see.MPA	3sPRO	

tomorrow I'm going to see him

In looking at texts, I observed that OV word order does occur in other non-tense related contexts. At the outset, when I began identifying the voice constructions, I observed an active voice construction and antipassive voice constructions. Within the set of active constructions, both a VO and an OV order were apparent, irrelevant of tense/aspect. Example (5) shows an active voice construction with VO word order.

5.

	<b>V</b>	<b>O</b>	
nu.mahataar+.kura.na		nu.iitipu mayaaha	paapaaha
3s.shake.PAST		3s.string front	fish

He shook his string in front of the fish.

Example (6) shows an active voice construction with an OV word order.

6.

	<b>O</b>	<b>V</b>
p+y++ni	paapaaha	na.casiita.kura.na
all	fish	3p.catch.PAST

All (kinds of) fish he caught.

Examples (7a) and (7b) show an active and an antipassive voice construction pair. Example (7a) illustrates the verb *asaa* 'eat' in an active voice construction.

7a.

	<b>V</b>	<b>O</b>
nu.asaa.riki	paapaaha	yawiini.hina
3s.eat.PAST	fish	dia.LOC

He ate fish every day.

## 6 Text Analysis

### 6.1 The pilot study

In the pilot study, I looked at 433 clauses from two Iquito narratives. The frequency distribution of the voice constructions described above is shown in table (6). 149 of those clauses were semantically transitive. Of those, 73.2% were active with VO word order, 10% were active with OV word order, and 16% were antipassive constructions.

	<i>Voice Constructions</i>	<i>N</i>	<i>%</i>
	active/direct (VO)	109	73.2
	active/direct (OV)	15	10.06
	antipassive	24	16.1
	Total	149	99.36

I calculated the referential distance and topic persistence for the agents and patients of each of those constructions.

In terms of referential distance, the percentages for agents in relation to patients patterned as predicted for the active constructions, as seen in table 3. That is, 70% of agents in the VO constructions were highly topical compared to the 29% of patients that were highly topical.

**Table 7. Percentage of agents and patients with RD values 1-2-3 (High topicality) and >3 (low topicality) in the various Iquito constructions**

	<i>Voice</i>	<i>Agent</i>		<i>Patient</i>	
		<i>0-3</i>	<i>&gt;3</i>	<i>0-3</i>	<i>&gt;3</i>
a.	Active VO (109)	<b>70</b>	30	29	<b>71</b>
b.	Active OV (15)	76	24	36	64
c.	Antipassive (24)	<b>66</b>	34	30	<b>70</b>

However, in terms topic persistence, 40% of patients in the active OV constructions were highly topical compared to the 16% of patients in the active VO constructions. The high topicality of patients as compared to patients in the VO constructions suggested that there might be something more within this set of data. That is, the percentages might vary with regard to the constructions with the adpositions.

**Table 8. Percentage of agents and patients with TP values 0-2 (low topicality) and >2 (high topicality) in the various Iquito voice constructions**

<i>Voice Constructions</i>	<i>Agent</i>		<i>Patient</i>	
	<i>0-2</i>	<i>&gt;2</i>	<i>0-2</i>	<i>&gt;2</i>
a. active VO (109)	46	54	84	16
b. active OV (15)	47	53	60	40
c. antipassive (24)	46	54	92	8

Of the 15 OV constructions, 9 appeared with one of the two adpositions (shown in examples 10 and 11). The data were reevaluated based on the distribution of constructions with adpositions. In terms of topic persistence, the OV constructions without adpositions patterned like the active VO constructions from table 3. The active OV constructions with adpositions patterned more like the inverse constructions shown in tables 2 and 3.

**Table 9. Percentage of Agents and Patients with TP values 0-2 (Low Topicality) and >2 (High Topicality) in active OV constructions with and without adpositions**

<i>Voice Constructions</i>	<i>Agent</i>		<i>Patient</i>	
	<i>0-2</i>	<i>&gt;2</i>	<i>0-2</i>	<i>&gt;2</i>
Active VO	46	54	84	16
Active OV w/out adposition (6)	34	66	88	17
Active OV with adposition (9)	56	44	44	56

The percentages in terms of referential distance showed a similar patterning for active OV constructions without adpositions. The referential distance percentages for OV constructions with adpositions did not however support the position that these constructions were inverse. The percentage for patients that were highly topical was too low compared to that of the agents, compared to the TP values.

**Table 10. Percentage of Agents and Patients with RD values 1-2-3 (High Topicality) and >3 (Low Topicality) in the various Iquito Voice Constructions**

<i>Voice Constructions</i>	<i>Agent</i>		<i>Patient</i>	
	<i>0-3</i>	<i>&gt;3</i>	<i>0-3</i>	<i>&gt;3</i>
Active VO	70	30	29	71
Active OV w/out adposition (6)	65	35	17	83
Active OV with adposition (9)	89	11	44	56

One possibility for this disparity was thought to be the small data sample. But the results from topical persistence measures warranted further investigation. This brings us to the second study.

## 6.2 The follow-up study

The subsequent study added narrative text from two sources. Two of the new texts came from work we did with the ILDP in the summer of 2003. Six of the new texts evaluated came from research done by SIL researchers in the 60s. This tripled the corpus to 1167 clauses (approximately 110 pages of interlinearized text) with 361 semantically transitive clauses. The distribution of the various voice constructions are shown in the following table.

<i>Voice Constructions</i>	<i>N</i>	<i>%</i>
active/direct (VO)	264	73
active/direct (OV) / (inverse?)	25	7
antipassive	67	19
passive	5	1
Total	361	100

The results from the follow-up study cut through any doubts as to the status of inverse constructions in Iquito.

In terms of referential distance (in table 12) the agents in all three constructions (active VO, OV without adposition, and OV with adposition) are significantly more topical than the patients. In fact, the OV constructions with adpositions show a higher number of agents with high topicality than the VO constructions or the OV constructions without adpositions.

**Table 12. Percentage of Agents and Patients with RD Values 1-2-3 (High Topicality) and >3 (Low Topicality) in the various Iquito Constructions**

<i>Voice</i>	<i>Agent</i>		<i>Patient</i>	
	<i>0-3</i>	<i>&gt;3</i>	<i>0-3</i>	<i>&gt;3</i>
Active VO (264)	67	33	31	69
OV without adposition (13)	74	26	32.5	67.5
OV with adpositions (17)	94	6	36	65

Further, with respect to topic persistence (table 13), note that the percentages for agents that are highly topical are all significantly higher than for the patients.

**Table 13. Percentage of Agents and Patients with TP values 0-2 (Low Topicality) and >2 (High Topicality) in the various Iquito Constructions**

<i>Voice</i>	<i>Agent</i>		<i>Patient</i>	
	0-2	>2	0-2	>2
Active VO (264)	43	57	85	15
OV without adposition (13)	29	71	61	39
OV with adposition (17)	59	41	64	35

### 6.3 An additional note: On evidence of language change

I stated previously that six (6) of the texts came from work done in the 1960s. In considering that language change might be a factor in terms the differences in percentages between the pilot study and the follow-up study, observe tables 14 and 15, which show that the measures for agents and patients in the 60s are the same as for agents and patients today. The possibility of language change being a factor is unlikely.

**Table 14. Percentage of Agents and Patients of OV with adposition constructions for ILDP and SIL texts with RD Values 1-2-3 (High Topicality) and >3 (Low Topicality) in the various Iquito Constructions**

<i>Voice</i>	<i>Agent</i>		<i>Patient</i>	
	1 - 3	>3	1 - 3	>3
ILDP	89	11	33	67
SIL	88	12	37	63

**Table 15. Percentage of Agents and Patients of OV with adposition constructions for ILDP and SIL texts with TP values 0-2 (Low Topicality) and >2 (High Topicality) in the various Iquito Constructions**

<i>Voice</i>	<i>Agent</i>		<i>Patient</i>	
	0-2	>2	0-2	>2
ILDP	56	44	55	45
SIL	63.5	37.5	74	25

## 7 Conclusions

The initial conclusion, based on the pilot study, had been that the adpositions *hina* and *anuura* were structural indicators of inverse constructions. Based on further investigation, that no longer seems to be the case. The number of examples available from collected texts is still rather small (25) and perhaps further analysis of text collected in the future will return other results. However, the findings presented here suggest that an inverse voice isn't likely to be found in semantically transitive constructions in Iquito. The locative and directional adpositions are not structural indicators of inverse constructions.

On the other hand, this work did bring about interesting observations with respect to auxiliaries, subordinators and word order, which had not emerged from elicitation but through the scrutiny of the natural discourse data. In fact, in the documented examples from elicited data there is not one occurrence of an auxiliary + *hina* adposition. Examples with the directional adposition *anuura* do appear, however they were treated much differently in isolation during analysis of elicited data and had not been considered part of larger paradigm. Work on auxiliary and subordinator paradigms is currently under way.

In sum, quantitative analysis did not prove fruitful in terms of identifying inverse constructions; rather it achieved the contrary, showing that Iquito does not have inverse constructions. Finally, this approach did provide insight into identifying other types of constructions such as auxiliaries.

### Notes

1. I am grateful to Roberto Zavala for his guidance and to Judith Aissen and Spike Gildea for their suggestions. All errors in this paper are mine alone.
2. This project is made possible by funding from the The Endangered Languages Fund and the Hans Rausing Endangered Languages Project. My gratitude to Christine Beier and Lev Michael for recruiting me to participate in this project and to our team of Iquito speakers and community *linguistas* in San Antonio del Pintuyacu is immeasurable.
3. No passives were clearly identifiable and require further analysis.
4. The percentages for the referential distance of 1 and 2 or 3 as calculated by Dryer (1994) have been collapsed here.

### References

- Cooreman, A. 1983. 'Topic Continuity and the Voicing System of an Ergative Language: Chamorro.' *Topic continuity in discourse. A quantitative cross-language study*, ed. by T. Givón. P. 425-490. Amsterdam: John Benjamins Publishing Company.
- Dryer, Matthew S. 1994. 'The discourse function of Kutenai inverse.' *Voice and Inversion*, ed. by T. Givón, ed. Pp. 65-99. Philadelphia, PA: John Benjamins Publishing Company.
- Gildea, Spike. 1994. 'Semantic and pragmatic inverse: "Inverse alignment" and "inverse voice" in Carib of Surinam.' *Voice and Inversion*, ed. by T. Givón, 187-230. Philadelphia, PA: John Benjamins Publishing Company.
- Givón, T. 1983. 'Topic continuity in discourse: An introduction.' *Topic continuity in discourse. A quantitative cross-language study*, ed. by T. Givón, 1-42. Amsterdam: John Benjamins Publishing Company.
- Givón, T. 1994. 'The pragmatics of de-transitive voice: Functional and typological aspects of inversion.' *Voice and Inversion*, ed. by T. Givón, 3-46. Philadelphia, PA: John Benjamins Publishing Company.
- Thompson, C. 1994. 'Passive and Inverse Constructions.' *Voice and Inversion*, ed by T. Givón, 47-64. Philadelphia, PA: John Benjamins Publishing Company.
- Zavala, R. 1997. 'Functional Analysis of Akatek Voice Constructions.' *International Journal of American Linguistics*. 63: 439-474.

# A Phonological Sketch of the Yucunany Dialect of Mixtepec Mixtec\*

Mary Paster and Rosemary Beam de Azcona  
UC Berkeley

## 1. Introduction

In this paper, we present a sketch of the phonology of the Yucunany dialect of Mixtepec Mixtec in terms of its segmental and tonal inventories as well as some phonological processes that we have observed. The Mixtepec Mixtec language is spoken by 12,000 people (Ethnologue, 2004). The speech community sits on the border between the Mixteca Alta and Mixteca Baja areas. The language is spoken in San Juan Mixtepec and many smaller surrounding towns in the ex-districts of Juxtlahuaca and Tlaxiaco. Previous discussions of Mixtepec Mixtec are found in Pike & Ibach (1978) and Josserand (1983).

Throughout the paper, we refer to the ‘couplet’, which is a common term for the structure of Mixtec roots (see, for example, Josserand 1983), since roots contain exactly two vowels. In Mixtepec Mixtec, root shape is as shown below:

(1) (C)(C)V(C)V

## 2. Segments

### 2.1 Consonants

#### 2.1.1 Consonant inventory

The phonemic consonantal inventory of Mixtepec Mixtec is shown below. The symbols used reflect the practical orthography devised by our Berkeley Mixtec working group. Symbols in parentheses represent phonemes that are marginally used in the language and/or occur only in loanwords.

(2) Consonantal inventory

Manner	Bilabial	Dental	Alveolar	Palatal(ized)	Retroflex	Velar	Labiovelar
Stop/affricate	(p)	t	tz	(ty)	ch	k	kw
Fricative			s		x	(j)	
	v		z			g	
Nasal	m	n		ɲ			
Liquid			l, (r, rr)				
Glide	(w)			y			

#### 2.1.2 Labial variation

We have observed multiple allophones of the /v/ phoneme: [v], [β], and [b]. [v] is the most common, occurring invariably in word-initial position, as in *vílú* [vílú] ‘cat’.

Word-medially, especially following a glottal stop, we have observed the [b] and [β] allophones occurring interchangeably with [v], as in the word *u'và* [uʔbà] ~ [uʔβà] ~ [uʔvà] 'salty'.

### 2.1.3 Phonotactic constraints on consonants

Consonants are restricted to onsets in Mixtepec Mixtec. We have found only one word containing a coda consonant, and it is a loanword: *árròðz* 'rice' < Spanish *arroz*.

Consonant clusters do not occur root-medially. They may occur word-initially or word-medially where prefixes concatenate with roots. The following clusters are attested:

- (3) Nasal + stop or affricate
- |                     |                     |                |                              |
|---------------------|---------------------|----------------|------------------------------|
| /nt/                | [nd]                | <i>ndúchá</i>  | 'goat'                       |
| /nt <sup>m</sup> s/ | [ndz]               | <i>ndzikì</i>  | 's/he walked'                |
| /nt <sup>m</sup> f/ | [nd <sup>m</sup> ʒ] | <i>ncháá</i>   | 'blue'                       |
| /nk/                | [ŋg]                | <i>kungà'á</i> | 'we will talk (to somebody)' |
| /nk <sup>w</sup> /  | [ŋg <sup>w</sup> ]  | <i>ngwii</i>   | 'fox'                        |
- /s/ + stop
- |       |               |                        |
|-------|---------------|------------------------|
| /st/  | <i>stàá</i>   | 'tortilla'             |
| /sk/  | <i>skéta</i>  | 'run!'                 |
| /skw/ | <i>Skwiía</i> | 'Santiago Juxtlahuaca' |
- /ʃ/ + consonant
- |                     |                 |                              |
|---------------------|-----------------|------------------------------|
| /ʃt <sup>m</sup> f/ | <i>xcháà</i>    | 'tomorrow'                   |
| /ʃn/                | <i>Ndìxnúúú</i> | 'Heroica Ciudad de Tlaxiaco' |

### 2.1.4 Loan and marginal consonants

The loan and marginal consonants of Mixtepec Mixtec are as follows. First, we have found /p/ only in loanwords such as *páii* 'rebozo' < Spanish *pañó* 'kerchief'. /t/ also occurs in loanwords, including *lúrrú* 'donkey' < Spanish *burro* 'donkey'. We have found /x/ only in loanwords such as the name *Jwáaà* 'Juan.' /r/ occurs in both Spanish loanwords and in the third person masculine respectful pronoun *-raà*, which may be a loan of unknown origin; Coatlán-Loxicha Zapotec has a third person human informal pronoun *ár*. The phoneme /w/ occurs only in function words, such as the first person plural exclusive pronoun *weè*. Finally, we have found /tʰ/ in only one word: *tyùtza* 'tree'.

## 2.2 Vowels

### 2.2.1 Vowel inventory

The vowel phonemes of Mixtepec Mixtec are shown below. This is a canonical five-vowel system with widely dispersed vowels.

(4) Vowel inventory

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

Each of these vowels has a nasalized counterpart. Although the vowels' duration varies contextually (for example, vowels bearing a contour tone are somewhat longer than vowels bearing a level tone), vowel length is not phonologically contrastive.

### 2.2.2 Vowel distribution

The vowels /i/, /u/, /o/, and /a/ occur regularly throughout the language, but the /e/ vowel is rare, occurring in few words. One instance of /e/ is found in *ve'e* 'house'.

Although each of the vowels shown in (4) has a nasalized counterpart, /õ/, and /eõ/ are rare. Some examples are shown below (nasalization is indicated in our orthography via underlining):

- (5) *nákò'ò* 'let's go!'  
*kwéèè* 'slow'

The nasalized vowels /iõ/, /ã/, and /uõ/ are common, as in the examples below.

- (6) *ii* 'one'                      *iií* 'skin'                      *ĩñù* 'six'  
*Jwáaà* 'Juan'                      *kà'á* 'talk!'                      *tzà'ù* 'fifteen'  
*nùú* 'face'                      *chu'ú* 'spider'                      *tzànu* 'brother's wife'

Interestingly, a majority of monomorphemic words in Mixtepec Mixtec have the same vowel in both syllables. This could be due to a historical vowel harmony process, though no such process exists as part of the synchronic phonology.

There are apparently no diphthongs in the language. In the rare examples containing non-identical adjacent vowels (e.g., *páii* 'rebozo'), the vowels belong to different syllables. We have not found any couplets with three different vowels not separated by consonants; such a couplet would probably have to be analyzed as having a diphthong since, as mentioned above, couplets contain exactly two syllables.

### 3. Suprasegmentals

#### 3.1 Tone

##### 3.1.1 Tone inventory

Mixtepec Mixtec has a three-tone system. The tonal inventory, as exemplified below, includes High (H), Mid (M), and Low (L) tones. In our orthography, H tone is represented by an acute accent, L tone by a grave accent, and M tone is not marked. Since vowel length is not contrastive, in our orthography we use additional vowels as needed to accommodate tone markings for contour tones.

(7)	L	<i>chùù</i>	‘star’	<i>tzàtù</i>	‘spicy’
		<i>ndàà</i>	‘flat’	<i>ĩñù</i>	‘six’
	M	<i>aa</i>	‘yes’	<i>luu</i>	‘little’
		<i>uu</i>	‘yes’	<i>ii</i>	‘one’
	H	<i>ncháá</i>	‘blue’	<i>vilú</i>	‘cat’
		<i>íí</i>	‘hail’	<i>kóní</i>	‘female turkey’

##### 3.1.2 Tone patterns on monomorphemic couplets

Couplets exhibit a maximum of three tones. The observed one-, two-, and three-tone patterns on couplets are shown below.

(8)	L	CVV	<i>chùù</i>	‘star’	<i>ìì</i>	‘nine’
			<i>ndàà</i>	‘flat’	<i>tzà’ù</i>	‘fifteen’
		CVCV	<i>sòkò</i>	‘shoulder’	<i>sùtù</i>	‘priest’
			<i>ùvì</i>	‘two’	<i>ùtzà</i>	‘seven’
	M	CVV	<i>ngwii</i>	‘fox’	<i>ve’e</i>	‘house’
			<i>luu</i>	‘little’	<i>u’u</i>	‘five’
		CVCV	<i>mula</i>	‘mule’	<i>machu</i>	‘mule’
			<i>yachi</i>	‘near’		
	H	CVV	<i>ncháá</i>	‘blue’	<i>nchá’á</i>	‘salsa’
		CVCV	<i>kóló</i>	‘male turkey’	<i>lóchí</i>	‘vulture’
			<i>kóní</i>	‘female turkey’	<i>ndúchá</i>	‘goat’
	LM	CVV	<i>chàa</i>	‘man’	<i>kàa</i>	‘metal’
			<i>ñà’a</i>	‘thing’	<i>sò’o</i>	‘ear’
		CVCV	<i>tzànu</i>	‘brother’s wife’	<i>tika</i>	‘cricket’
			<i>tìtzi</i>	‘stomach’	<i>yùcha</i>	‘river’

LH	CVV	<i>stàá</i> ‘tortilla’	<i>nùú</i> ‘face’
	CVCV	<i>kòó</i> ‘snake’ <i>kùmí</i> ‘four’ <i>tzàtú</i> ‘coffin’	<i>kò’ó</i> ‘plate’ <i>námá</i> ‘soap’ <i>sàví</i> ‘rain’
MH	CVV	<i>yo’ó</i> ‘rope’	<i>che’é</i> ‘cute’
	CVCV	<i>chu’ú</i> ‘spider’ <i>yatá</i> ‘old’	<i>ña’á</i> ‘woman’
HM	CVV	<i>xí’a</i> ‘hawk’	
	CVCV	<i>tzóko</i> ‘possum’ <i>ñáni</i> ‘brother’	<i>tzíka</i> ‘far’ <i>ká’nu</i> ‘big’
HL	CVV	<i>cháì</i> ‘chair’	<i>kwá’à</i> ‘red’
	CVCV	<i>chá’à</i> ‘short’ not attested	<i>nchá’i</i> ‘black’
ML	CVV	<i>yoò</i> ‘drinking vessel’	<i>saà</i> ‘bird’
	CVCV	<i>xitò</i> ‘uncle’ <i>yu’và</i> ‘ice’	<i>tutù</i> ‘paper’ <i>u’và</i> ‘salty’
LML	CVV	<i>xàaà</i> ‘chin’	<i>xìoò</i> ‘dress, skirt’
	CVCV	<i>ndàakù</i> ‘broom’	
LHM	CVV	<i>tzàáa</i> ‘new’	
	CVCV	<i>yòoso</i> ‘metate’ <i>kàása</i> ‘sister’s husband’	<i>yùúti</i> ‘sand’ <i>tííchi</i> ‘avocado’
MLH	CVV	<i>viíí</i> ‘healthy-looking’	
	CVCV	<i>yosòò</i> ‘grassy plain’ <i>tikwàá</i> ‘orange’	<i>ixíí</i> ‘hair’ <i>ikíí</i> ‘squash’
MHM	CVV	<i>Skwiía</i> ‘Santiago Juxtlahuaca’	
	CVCV	<i>kotóo</i> ‘sarape’	
HML	CVV	<i>Jwáaà</i> ‘Juan’ (< Spanish)	<i>páii</i> ‘rebozo’
	CVCV	<i>yúkuù</i> ‘yoke’ (< Sp. <i>yugo</i> ) <i>Mínaà</i> ‘Los Tejocotes’ (< Sp. <i>mina</i> ‘mine’)	<i>ánaà</i> ‘heart’ <i>súkuù</i> ‘high’
HLH	CVV	<i>chííí</i> ‘fingernail’	<i>ííí</i> ‘skin’
	CVCV	<i>kwííí</i> ‘narrow’ <i>xííí</i> ‘hat’	<i>kwííí</i> ‘green’

Note that not all possible combinations of three tones are observed. Some generalizations regarding the gaps in the set of tone patterns on couplets are as follows. First, we do not find patterns with two adjacent instances of the same tone (e.g. \*HHL, where the first syllable is H-toned and the second has a HL falling tone)<sup>1</sup>. We can rule out such patterns via a version of the Obligatory Contour Principle, which bans adjacent identical tone specifications, applying to underlying forms. However, we must also rule out some couplets with multiply linked underlying tones to avoid unattested patterns such as \*HHL arising from underlying HL (where the H tone is linked to both the first and second syllable). This is problematic since we wish to claim that couplets with the same tone on both syllables, e.g. *vilú* ‘cat’, have only a single, doubly linked tone. One way of solving this problem is to assume that tones are not linked underlyingly; they are associated to the couplet via an algorithm associating tones to tone-bearing units one-to-one from left to right. This would prevent underlying /HL/ from surfacing as \*HHL.

Pike & Ibach (1978) report that HM and MH contours do not occur on a single syllable; however, we have found two MHM words in the Yucunany dialect which do exemplify these two contours on a single syllable. *Skwiia* ‘Santiago Juxtlahuaca’ has a MH contour on the first syllable and *kotóo* ‘sarape’ has a HM contour on the second syllable. Thus, although HM and MH contours are rare, they are not banned outright.

In addition to the unattested patterns mentioned above, we have not found the tone patterns LMH, LHL, MLM, MHL, HLM, or HMH on couplets. As far as we can tell, this is not a unified category of tone patterns that can be ruled out via any principled constraint or set of constraints. Therefore, we conclude that these are accidental gaps that may (or may not) be filled during future work on this language.

### 3.2 Nasalization

As mentioned above, all five vowels of Mixtepec Mixtec have contrastively nasalized counterparts, although the mid vowels /õ/ and /eə/ are rare. Below are some minimal and near-minimal pairs differing in nasalization:

- (9) *luu* ‘little’      *uu* ‘yes’  
*ĩí* ‘husband’      *ĩĩ* ‘skin’  
*kàa* ‘metal’      *kà’á* ‘talk!’  
*nàamá* ‘soap’      *ñùmá* ‘wax’

In addition to contrastive nasalization, we find contextual nasalization on vowels following nasal consonants, as also reported by Pike & Ibach (1978). However, in the Yucunany dialect, this contextual nasalization is non-neutralizing, so oral and nasal vowels do contrast following a nasal consonant as seen in the examples below.

- (10) *nàamá* ‘soap’      *ñùmá* ‘wax’  
*ánaà* ‘heart’      *ĩnù* ‘six’

*kóní* ‘female turkey’      *tzànu* ‘brother’s wife’  
*mula* ‘mule’                      *nùú* ‘face’

In couplets, vowel nasalization usually occurs in both syllables or neither, although there are also rare examples of couplets with nasalization only in the second syllable (e.g., *tzànu* ‘brother’s wife’).

### 3.3 Glottalization

Glottalization is contrastive on the first vowel of the couplet, there is apparently no contrastive glottalization on the second vowel of the couplet. Glottalization is sometimes realized as creaky voicing, but more commonly as a glottal stop following the first vowel (as reflected in our orthography, where glottalization is indicated by ’ after the first vowel). Below are some minimal and near-minimal pairs differing in glottalization.

- (11) *ncháá* ‘blue’                      *nchá’á* ‘salsa’  
*kòó* ‘snake’                              *kò’ó* ‘plate’  
*ùvà* ‘sour’                                *ù’và* ‘salty’  
*chùú* ‘chicken’                        *chu’ú* ‘spider’  
*yuvà* ‘wild vegetable’              *yu’và* ‘ice’

## 4. Phonological rules

### 4.1 Tone rules

#### 4.1.1 Gradient Smoothing

In LHH sequences, the first H is lowered to M. This phenomenon was also documented in Chalcatongo Mixtec by Hinton et. al. (1991). As shown below, LH-final roots, when followed by the H-toned 2sg familiar suffix, undergo Gradient Smoothing, so that the final syllable of the root surfaces with M tone.

- (12) *chìí* ‘fingernail’                      *chìí gú* ‘your fingernail’  
*kwíí* ‘narrow/thin’                      *kwíí gú* ‘you are narrow/thin’  
*tikwàá* ‘orange’                              *tikwàá gú* ‘your orange’  
*nàmá* ‘soap’                                      *nàma gú* ‘your soap’

We formalize Gradient Smoothing below. Note that we are assuming that the mora is the tone-bearing unit, though this assumption is not crucial.

- (13)  $\begin{array}{ccc} \updownarrow & \updownarrow & \updownarrow \\ | & \# & | \\ L & H & H \end{array}$

This rule delinks H between L and H. In order for this rule to yield surface [LMH], we must assume that the unlinked H is deleted via Stray Erasure, and that a vowel without a

phonological tone is realized phonetically with Mid tone. However, we do not wish to argue that there is no phonological Mid tone; rather, we claim that there are two representations that correspond to surface Mid tone: phonologically specified Mid, and tonal non-specification.

#### 4.1.2 Low Tone Spreading

When an underlying LH sequence spans two syllables, the L is realized on both syllables, yielding surface [L.LH]. Low Tone Spreading (LTS) is also documented in Chalcatongo Mixtec (Buckley 1991). The effect of LTS can be seen in the 2sg. As shown below, the 2sg clitic surfaces with level H tone when it follows a H or M tone.

(14)	<i>lúrrú</i>	‘donkey’	<i>lúrrú gú</i>	‘your donkey’
	<i>tikwàà</i>	‘orange’	<i>tikwàà gú</i>	‘your orange’
	<i>xá’nú</i>	‘cigarette’	<i>xá’nú gú</i>	‘your cigarette’
	<i>tzàáku</i>	‘corral’	<i>tzàáku gú</i>	‘your corral’
	<i>kàa</i>	‘metal’	<i>kàa gú</i>	‘your metal’

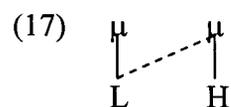
When preceded by a L tone, the 2sg *gú* surfaces with a LH rising tone, as seen below.

(15)	<i>cháì</i>	‘chair’	<i>cháì gú</i>	[t <sup>m</sup> ʃáì ùú]	‘your chair’
	<i>chá’à</i>	‘short’	<i>chá’à gú</i>	[t <sup>m</sup> ʃá’á ùú]	‘you are short’
	<i>ánaà</i>	‘heart’	<i>ánaà gú</i>	[án:áá ùú]	‘your heart’
	<i>yuchì</i>	‘knife’	<i>yuchì gú</i>	[yut <sup>m</sup> ʃì ùú]	‘your knife’

LTS also appears to apply within lexical items. There is no contrast between LH and L.LH in roots, and as shown below, an underlying /LH/ sequence surfaces as [L.LH].

(16)	/kù mí/ →	[kù:m:í]	‘four’	/námá/ →	[nám:áá]	‘soap’
	/tsátú/ →	[tsátùú]	‘box’	/sàví/ →	[sàví]	‘rain’

We formalize LTS as shown in (17).



## 4.2 Segmental rules

### 4.2.1 Nasal Place Assimilation

There is no place contrast in nasals immediately preceding another consonant. The nasal is invariably pronounced at the same place of articulation as the following

consonant. This does not apparently result in any nasal place alternations, but the pattern is robust in the lexicon, as in the examples below.

- |      |          |             |          |                              |
|------|----------|-------------|----------|------------------------------|
| (18) | [ndikaà] | ‘pine cone’ | [ndaakù] | ‘broom’                      |
|      | [nda’á]  | ‘hand’      | [ɲgàù]   | ‘Santa María Tepostlantongo’ |
|      | [ɲgwii]  | ‘fox’       | [nchaa]  | ‘blue’                       |

We posit a rule of Nasal Place Assimilation to account for this pattern. In this and subsequent rules, C and V refer to Root nodes of consonants and vowels, respectively.

- (19)
- 

We have not found any cases of [mp] (or, more likely given the Post-Nasal Voicing Assimilation rule to be discussed, [mb]). However, this is not surprising since as mentioned earlier, /p/ is a marginal phoneme found only in loanwords.

An alternative to this approach would be to claim that there is a prenasalized stop series, so that the word-initial onsets in the examples in (18) are single segments. We have opted to analyze these as complex N+C onsets based on evidence from verbs, which are marked by an *n-* prefix in the completive aspect when the root begins with /t/ or /ts/. In these cases, the onset of the completive form of the verb surfaces as [nd] or [ndz], respectively, and these onsets sound identical to those found in the couplets in (18). Some examples are provided below:

- |      |                 |                 |                  |           |
|------|-----------------|-----------------|------------------|-----------|
| (20) | <i>tii</i>      | ‘I am holding’  | <i>ndii</i>      | ‘I held’  |
|      | <i>tivii yù</i> | ‘I am blowing’  | <i>ndivii yù</i> | ‘I blew’  |
|      | <i>tatzii</i>   | ‘I am eating’   | <i>ndatzii</i>   | ‘I ate’   |
|      | <i>tzi’i yù</i> | ‘I’m dying’     | <i>ndzi’i yù</i> | ‘I died’  |
|      | <i>tzi’ii</i>   | ‘I am drinking’ | <i>ndzi’ii</i>   | ‘I drank’ |

#### 4.2.2 Post-Nasal Voicing Assimilation

In the examples in (20) above, couplet-initial /t/ and /tz/ surface as [d] and [dz], respectively, when preceded by the completive *n-* prefix. There is no contrast between, e.g., [nt] and [nd]; in every combination of a nasal consonant + other consonant, the entire sequence is voiced. We account for this via a Post-Nasal Voicing Assimilation rule, as formalized below.

- (21)
-

[nasal]

### 4.2.3 Sonorant Gemination

Sonorant consonants are audibly lengthened intervocalically in a root, as shown in the narrow transcriptions below.

- (22) *kù mí* [kù:m:ií] ‘four’                      *ù nì* [ùn:i] ‘three’  
*ì ñù* [ìn:yù] ‘six’                                      *kóló* [kól:ó] ‘male turkey’

We have not systematically investigated the difference in duration between lengthened and non-lengthened sonorants, but in a typical token of *vílú* ‘cat’, the duration of medial [l] is 165 ms, compared to 119 ms for the initial [l] in a typical token of *lóchí* ‘vulture’.

To account for this pattern, we posit a rule of Sonorant Gemination that inserts a mora linked to the medial sonorant. In effect, this causes the medial sonorant to become a coda to the first syllable in addition to being the onset of the second syllable.

- (23) 
$$\left( \begin{array}{c} \mu \\ \downarrow \\ (C) \quad V \end{array} \quad \begin{array}{c} \mu \\ \vdots \\ C \\ | \\ \text{[sonorant]} \end{array} \quad \begin{array}{c} \mu \\ \downarrow \\ V \end{array} \right) \text{Couplet}$$

## 5. Morphophonology

### 5.1 Completive aspect allomorphy

Verb stems apparently can only begin with /t/, /ts/, /k/, or /sk/. As shown below, when the stem begins with /k/, the completive aspect is marked by *nì-* and a complex set of tonal alternations for which we do not yet have a full analysis.

- (24) *ká' à yù* ‘I am talking’                      *nìkà' à yù* ‘I talked’  
*kíkuù* ‘I am sewing’                              *nìkíkuù* ‘I sewed’  
*ká' viì* ‘I am reading’                              *nìka' viì* ‘I read (past)’  
*kánaà* ‘I am yelling’                              *nìkanaà* ‘I yelled’

Verb stems beginning in /t/ or /ts/ are marked by an *n-* prefix in the completive form and also a floating L tone prefix, which corresponds to the low tone of the *nì-* allomorph.

- (25) *tíiì* ‘I am holding’                              *ndiì* ‘I held’  
*tíviì yù* ‘I am blowing’                              *ndíviì yù* ‘I blew’  
*tzatzíi* ‘I am eating’                              *ndzàtzíi* ‘I ate’  
*tzí' ì yù* ‘I'm dying’                              *ndzì' ì yù* ‘I died’  
*tzí' iì* ‘I am drinking’                              *ndzì' iì* ‘I drank’

When the stem begins with /sk/, the completive is not marked segmentally but takes the floating L tone prefix.

- (26) *skétaà* 'I am running'      *skèétaà* 'I ran'  
*skánaà ñaà* 'I am throwing it'      *skàánaà ñaà* 'I threw it'

There is a difference in the way that the floating L tone associates in each of these groups. In the /t/- and /ts/-initial group the floating L replaces the stem-initial tone. In the /sk/-initial group the floating L concatenates to the left of the stem-initial tone. At present, we do not have an explanation for why the different stem types exhibit different tonal behaviors in the past tense.

## 5.2 1sg allomorphy and non-homophony

The 1sg is usually marked by a floating L tone that associates to the end of the root. When the root has final H tone, the L associates straightforwardly to the end of the root, yielding a HL falling tone on the final syllable with no further tone changes.

- (27) *nà má* 'soap'      *nàmáà* 'my soap'  
*kwíì* 'narrow/thin'      *kwíìì* 'I am narrow/thin'  
*xínì* 'hat'      *xínìì* 'my hat'  
*vílú* 'cat'      *vílúù* 'my cat'

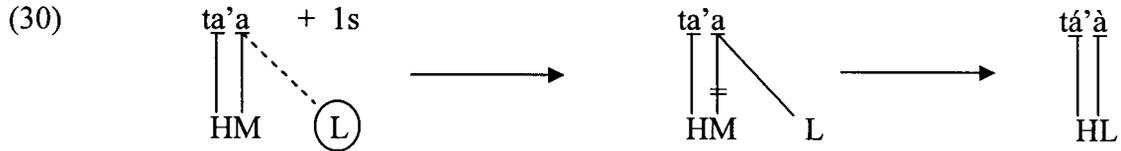
The association of the floating L (indicated by circling) to the end of the root is schematized below.

- (28)  $nà má + 1s \rightarrow nàmáà$
- 

The tonal behavior of M-final roots in the 1sg depends on the preceding tones. On M-final roots where preceding tone is M or H, the floating L tone of the 1sg associates to the end of the root, and the final syllable of the root surfaces with a level L tone rather than the ML falling tone that would have been expected if no tone changes occurred.

- (29) *la'la* 'mucus'      *la'là* 'my mucus'  
*ve'e* 'house'      *ve'è* 'my house'  
*tá'á* 'relative'      *tá'à* 'my relative'  
*xá'nu* 'cigarette'      *xá'nù* 'my cigarette'

Since the final M of the root does not surface in the 1sg, we characterize the tone change as delinking (and subsequent deletion via Stray Erasure) of the final M of the root, as illustrated below.



There is an unusual exception to the above generalization, which is that, unlike other roots ending in HM, if the root has the tone pattern LHM, then the final M of the root is retained in the 1sg, as in the examples in (31).

(31)

<i>yùúti</i>	'sand'	<i>yùútiì</i>	'my sand'
<i>tzàáku</i>	'corral'	<i>tzàákuù</i>	'my corral'
<i>yòóso</i>	'metate'	<i>yòósoò</i>	'my metate'
<i>kàása</i>	'sister's husband'	<i>kàásaà</i>	'my sister's husband'

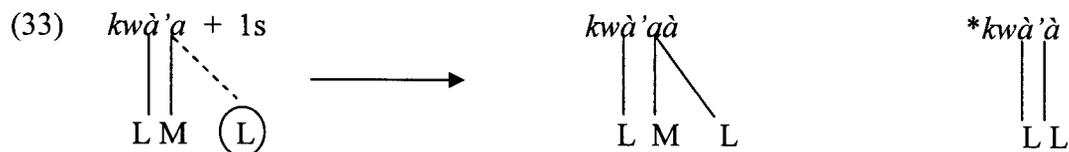
This may relate to the fact that there are no LHL roots. However, it cannot result from a general ban against LHL sequences, since such sequences can be derived.

As shown below, when the root has final M tone and the preceding tone is L, the M is retained in the 1sg.

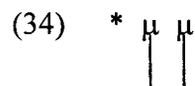
(32)

<i>kwà'a</i>	'man's sister'	<i>kwà'aà</i>	'my sister'
<i>sì'i</i>	'leg'	<i>sì'iì</i>	'my leg'
<i>tìtzi</i>	'stomach'	<i>tìtziì</i>	'my stomach'
<i>kàa</i>	'metal'	<i>kàaà</i>	'my metal'

We schematize this situation below. The floating L tone associates to the end of the root, but the M tone is not delinked. Thus, rather than LL (which might be expected based on the behavior of HM- and MM-final roots), these roots surface with LML tone in the 1sg.



One possible analysis for the above pattern is that the Obligatory Contour Principle (OCP; Leben 1973, 1978, Goldsmith 1976) is active in this language, penalizing (derived) sequences of L tones. Deleting the M would have yielded a sequence of L tones in these words, but this could be blocked by a version of the OCP as formalized below.



## L L

Until now we have discussed the behavior of roots ending in H or M tones only. Roots with final L tone take a different allomorph of the 1sg marker: =yù. Some examples are provided below.

(35)	<i>chá'à</i>	'short'	<i>chá'àyù</i>	'I am short'
	<i>chài</i>	'chair'	<i>chàiyù</i>	'my chair'
	<i>tutù</i>	'paper'	<i>tutùyù</i>	'my paper'
	<i>sòkò</i>	'shoulder'	<i>sòkòyù</i>	'my shoulder'

It is interesting to note that this allomorphy has the effect of preventing homophony between 1sg and unmarked forms. If L-final roots took the floating L tone allomorph, their 1sg forms would be homophonous with the unmarked form. In the dialect of Mixtepec Mixtec described by Pike & Ibach (1978), the 1sg is, in fact, homophonous with the unmarked form of L-final roots. Pike & Ibach give -yù as marking the 1sg polite; our Yucunany consultant does not distinguish polite from familiar in the 1sg.

Pike & Ibach (1978) appear to assume that yù is a suffix, but in Yucunany, yù can occur after noun+adjective, suggesting that it is not a suffix since material can intervene between yù the root. In cases where yù occurs with noun+adjective, the choice of the floating L tone vs. yù allomorph depends on the final tone of the adjective, irrespective of the tones of the noun. As shown below, when the noun has final L tone (as in *tìnà* 'dog'), the selection of the 1sg allomorph is determined by the final tone of the adjective that follows the noun.

(36)	<i>nchá'ì</i>	'black'	<i>tìnà nchá'ìyù</i>	'my black dog'
	<i>yá'à</i>	'gray'	<i>tìnà yá'àyù</i>	'my gray dog'
	<i>ncháá</i>	'blue'	<i>tìnà nchááà</i>	'my blue dog'
	<i>luu</i>	'little'	<i>tìnà luu luù</i>	'my little baby dog'

As expected, when the adjective is held constant (as in the examples below using *nchá'ì* 'black'), the tones of the noun are irrelevant to the selection of the 1sg allomorph.

(37)	<i>kwìkà</i>	'comb'	<i>kwìkà nchá'ìyù</i>	'my black comb'
	<i>chài</i>	'chair'	<i>chài nchá'ìyù</i>	'my black chair'
	<i>ve'e</i>	'house'	<i>ve'e nchá'ìyù</i>	'my black house'
	<i>vílú</i>	'cat'	<i>vílú nchá'ìyù</i>	'my black cat'
	<i>sa'màá</i>	'clothing'	<i>sa'màá nchá'ìyù</i>	'my black clothing'
	<i>lúrrú</i>	'donkey'	<i>lúrrú nchá'ìyù</i>	'my black donkey'

### 5.3 3sg vowel allomorphy

The 3s g familiar also exhibits phonologically conditioned allomorphy. The 3sg has three allomorphs whose distribution depends on the final segment of the root. On *i*-final roots, the 3sg is marked with *-à*.

(38)	<i>sì'ì</i>	'leg'	<i>sì'áà</i>	'his leg'
	<i>kachìi</i>	'cotton'	<i>kachìáà</i>	'his cotton'
	<i>kwíi</i>	'narrow'	<i>kwíáà</i>	'he is narrow'
	<i>tzí'ì</i>	'I am dying'	<i>tzí'à</i>	'she is dying'
	<i>yatzii</i>	'small water dipper'	<i>yatziaa</i>	'his small water dipper'
	<i>xìchí</i>	'nose'	<i>xìchìà</i>	'his nose'

Elsewhere, the 3sg is marked using *ì* or *ñaa*. The *-ì* allomorph occurs with roots ending in *a*, *o*, and *u*, as shown below.

(39)	<i>sáà'má</i>	'clothing'	<i>sáà'mì</i>	'his clothing'
	<i>vàá'a</i>	'bad'	<i>vàá'ì</i>	'it is bad'
	<i>tá'a</i>	'relative'	<i>tá'ì</i>	'his relative'
	<i>nda'á</i>	'hand'	<i>nda'ì</i>	'her hand'
	<i>kò'ó</i>	'plate'	<i>kò'ì</i>	'his plate'
	<i>yo'ó</i>	'rope'	<i>yo'ì</i>	'his rope'
	<i>sò'o</i>	'ear'	<i>sò'ì</i>	'his ear'
	<i>ma tzá'nu</i>	'grandmother'	<i>ma tzá'nì</i>	'his grandmother'
	<i>kù'ù</i>	'woman's sister'	<i>kù'ì</i>	'her sister'

The *-ñaa* allomorph occurs with roots ending in *a*, *o*, *e*, and *u*.

(40)	<i>nchá'á</i>	'salsa'	<i>nchá'á ñaa</i>	'his salsa'
	<i>yosòó</i>	'grassy plain'	<i>yosòó ñaa</i>	'his grassy plain'
	<i>xitò</i>	'uncle'	<i>xitò ñaa</i>	'his uncle'
	<i>xìòò</i>	'dress'	<i>xìòò ñaa</i>	'her dress'
	<i>yoò</i>	'drinking vessel'	<i>yoò ñaa</i>	'his drinking vessel'
	<i>ve'e</i>	'house'	<i>ve'e ñaa</i>	'his house'
	<i>kiché'é</i>	'bone'	<i>kiché'é ñaa</i>	'his bone'
	<i>kwéèè</i>	'slow'	<i>kwéèè ñaa</i>	'he is slow'

The *-ì* and *-ñaa* allomorphs overlap in their distribution, since both can occur with roots ending in *a*, *o*, and *u*. We have not been able to establish a generalization regarding the use of *-ì* vs. *-ñaa*. However, one clear generalization regarding the 3sg (familiar) is that roots ending in /i/ never take the *-ì* allomorph. This has the effect of preventing homophony with the plain form of roots ending in L-toned /i/ and with the 1sg form of underlyingly HM- and MM-toned roots ending in /i/.

## 6. Conclusion

In this paper we have provided a brief description of the phonological system of the Yucunany dialect of Mixtepec Mixtec as we currently understand it. Some of our most interesting results have been in the domain of phonologically based allomorphy. In particular, there are two categories (1sg and 3sg) where allomorphy has the effect of preventing homophony with other morphologically related forms. The exact distribution of these allomorphs and the historical origin of the allomorphy are part of our ongoing research on this dialect. Other phonological topics for future research include the distribution of nasalized vowels and patterns of loanword adaptation.

## Notes

\* We are very grateful to our consultant Tisu'mà X for his time and patience. We also thank Wendianne Eller, Rich Rhodes, and Leanne Hinton for their various contributions to this work. This research has been funded by the Survey of California and Other Indian Languages at UC Berkeley.

<sup>1</sup> We have found a small number of apparent counterexamples which we suspect are either synchronically or historically multimorphemic or borrowings since they have more than two syllables, e.g. *Kanáamá* 'Cañama' (a toponym).

## References

- Buckley, Eugene. 1991. 'Low-Tone spreading in Chalcatongo Mixtec.' *Papers from the American Indian Languages Conferences: Occasional Papers on Linguistics* 16, ed. by James E. Redden, 168-72. Carbondale: Southern Illinois University.
- Goldsmith, John. 1976. *Autosegmental Phonology*. PhD. dissertation, MIT.
- Hinton, Leanne, Gene Buckley, Marv Kramer, and Michael Meacham. 1991. 'Preliminary analysis of Chalcatongo Mixtec Tone.' *Papers from the American Indian Languages Conferences: Occasional Papers on Linguistics* 16, ed. by James E. Redden, 147-55. Carbondale: Southern Illinois University.
- Josserand, Judy Kathryn. 1983. *Mixtec Dialect History*. PhD. dissertation, Tulane University.
- Leben, Will. 1973. *Suprasegmental phonology*. PhD. dissertation, MIT.
- Leben, Will. 1978. 'The representation of tone.' *Tone: A Linguistic Survey*, ed. by Victoria Fromkin, 177-219. New York: Academic Press.
- Pike, Eunice V. and Thomas Ibach. 1978. 'The phonology of the Mixtepec dialect of Mixtec.' *Linguistic and Literary Studies in Honor of Archibald A. Hill, Volume 2: Descriptive Linguistics*, ed. by Mohammed Jazayery, Edgar C. Polomé, and Werner Winter, 271-85. The Hague: Mouton.

## Grammaticization of Tense in Navajo: The Evolution of *nt'éeé*

Melvatha Chee<sup>1</sup>, Evan Ashworth, Susan Buescher, Brittany Kubacki\*  
University of New Mexico

### 1. Introduction

Grammaticization, according to Hopper and Traugott (1993:1) “focuses on how grammatical forms and constructions arise, how they are used, and how they shape the language.” In this paper, we examine a case of grammaticization in Navajo involving the emergence of tense marking. Athabaskan languages are well known for their elaborate system of morphologically marked aspect and aktionsart (Axelrod 1993, Midgette 1995). Data from Navajo suggests that this Southern Athabaskan language is developing a tense system in addition to its system of aspect. We explore here, in particular, the functions of Navajo *nt'éeé* to illustrate its development from a temporal adverb to a tense marker.

We begin our discussion with an explanation of grammaticization. Next we provide an overview of the temporal system of Navajo and references to *nt'éeé* in the scholarly literature on Navajo. This historical perspective details both how the construction was used by speakers of Navajo in the past and also what the leading researchers have thought regarding its meaning and usage. We then proceed to examine examples of *nt'éeé*, as it is found in current usage patterns and discourse style. Much of the data considered here was collected in a 2003 linguistic field methods course at the University of New Mexico, in which author Chee was the native speaker consultant. The rest of the examples come from spontaneous conversation between three to four native speakers, recorded and transcribed by Chee.

---

\*We are sincerely indebted to Melissa Axelrod for her continuous encouragement and support as we worked on this paper. We would also like to thank Joan Bybee for her guidance and knowledge in the area of grammaticization.

<sup>1</sup> Many thanks to Melvatha Chee for providing all the modern Navajo data and giving us the gift of her expertise in collaborating on the analysis.

## 2. Grammaticization

Our discussion of *nt'éeé* in this paper is guided by the theoretical framework of grammaticization. “Reduced to its essentials, grammaticization theory begins with the observation that grammatical morphemes develop gradually out of lexical morphemes or combinations of grammatical morphemes” (Bybee et al. 1994). As concluded in Bybee 1985 and Bybee et al. 1994, one of the most important elements impacting grammaticization is frequency of use. As lexical items increase in frequency, they become semantically generalized and phonetically reduced. With the loss of meaning and substance, these forms can be used by speakers in more contexts. Cross-linguistic evidence of grammaticization cases display common tendencies and regularities, in which frequently used verbs like *go*, *have*, *do*, and *be* become obligatory in certain contexts because of their generality.

Lexical items and constructions all follow a grammaticization path, which tends to consist of four main principles: phonological reduction, morphosyntactic reduction, semantic generalization, and pragmatic extension (Bybee et al. 1994). Phonological reduction is the shortening or loss of phonetic characteristics, as in *have to* becoming [haftə]. Morphosyntactic reduction is the loss of salient morphological or syntactic features; that is, the grammaticizing construction tends to lose its syntactic independence and become more like a clitic or an affix. Semantic generalization is the loss of features of meaning. For example, the verb *will* in English used to mean “to want, to desire.” Because it was frequently used, it lost this meaning of desire and came to mean just ‘willingness,’ and from there, ‘intention.’ The future or prediction meaning was then inferred from this intention meaning, and so *will* generalized from the meaning of ‘desire’ to the English Future tense marker (Bybee et al. 1994: 254-256). The last component of grammaticization is pragmatic extension, in which the lexeme extends to new

contexts. Thus, when a future or predicative meaning was inferred from *will*, the form was extended pragmatically and was used in a wider range of contexts.

It is not only important to look at the process of grammaticization, but the effects as well. One of these effects is layering, in which a construction develops polysemous meanings that reflect the past stages of the grammaticization path. Once again, we return to the example of *will*. Although *will* has Future tense meaning in English, it also retains the meanings of ‘willingness’ and ‘intention’ (Hopper and Traugott 1993:97). These layered meanings are further evidence of the grammaticization path.

### 3. Verbal and temporal system of Navajo

This section examines the structure of the Navajo verb, and how it expresses temporality. According to Young (2000:vii), “The verb occupies a central position in Navajo, and in the Athabaskan languages generally. Most of the noun lexicon, along with adverbs and adjectivals, are verbal in origin or verb related in one way or another.”

The Athabaskan verb is a polysynthetic structure composed of a stem and its prefixes (Axelrod 2004). The stem is usually monosyllabic and is the last syllable of the verb form. It may also have an obligatory prefix known as the *root*. Together, the root and the stem form the *verb base* (Faltz 1998). The verb base can undergo aspectual modifications as well; for example, some verbs have an Imperfective and a Perfective stem, which differ phonologically.

The verbal prefixes are used to mark person, number, aspect, tense, and adverbials. These prefixes generally occur in the same order, as shown in Figure 1.

10	9	8	7	# #	6	5	4	3	2	1	0
post-positions	derivational/ thematic	reflexive reversionary semiliterative iterative	distrib./pl	disjunct boundary	DO	Subj. (3 <sup>rd</sup> )	derivational/ thematic	mode/ conjugation	Subj. (1 <sup>st</sup> & 2 <sup>nd</sup> sg & pl)	voice/ valence	stem

### Figure 1 Navajo Prefix Chart

Some prefixes are obligatory, such as person, tense, aspect, mode, voice/valence (or “classifier”) prefixes (Axelrod 2004). Non-obligatory prefixes tend to be derivational in nature, and thus their “slots” need not be filled in every verb.

Temporal expression in the Navajo verb is marked by the Mode and Aspect prefixes. Mode is used to define the temporal character of an activity, that is if it is incomplete, complete, ongoing, future, potential, customary, or recurrent. Mode is marked by modifications to the verb stem, or in the third position prefix. There are seven Modes in Navajo and these Modes are illustrated below.

(A) The Imperfective describes an action or event that is incomplete. It is usually translated as the English Present Tense: *yáshti*, ‘I talk.’

(B) The Perfective describes an action or event that is completed. It is usually translated as the English Past Tense: *yááhti*, ‘I talked.’

(C) The Future describes an action or event that is impending or in the future. It is derived from the inceptive prefix *di-* and the progressive: *deesháál*, ‘I’ll go/come.’

(D) The Usitative Mode describes an action or event that occurs usually or customarily. For example, compare *yishdlá* ‘I’m drinking it’ with the usitative *yishdlfih* ‘I usually drink it’.

(E) The Iterative Mode is closely related to the Usitative, with the addition of the repetitive prefix *ná-*. It describes an action or event that occurs customarily or repeatedly, with an emphasis on repetition: *ndídááh*, ‘he walks back and forth (as a sentinel). Faltz (1998:15) states that of the two Modes, the Iterative is used more than the Usitative if there is a choice between the two.

(F) The Progressive Mode describes an action that is ongoing. It is usually translated with the English Present Progressive: *yisháál*, 'I'm walking along'. Faltz (1998:16) mentions that many verbs do not actually have a Progressive Mode.

(G) The Optative Mode, which is marked by -ó- or -o- in position VII, describes an action as potential or desired: *nahóltáá' laanaa*, 'I wish it would rain'. The second person optative is used to express a negative command: *ndoó'nééh*, 'don't get up!'

The Aspects express the manner in which an activity or event is carried out over time – whether it happens once, or repeatedly, or at length. There are twelve Aspects that make such distinctions in meaning as: 'I am red' vs. "I turned red"; 'I am eating' vs. 'I swallowed'; 'I am walking about' vs. 'I walked in'. (Young and Morgan with Midgette, 1992). Aspect is marked by stem shape and/or prefixes.

The twelve Aspects in Navajo are:

- (A) Momentaneous, which is an action that takes place at one point in time.
- (B) Continuative, an action or event that extends over an indefinite period of time, or an action without a goal.
- (C) Durative, an action that extends over some time.
- (D) Conclusive, an action or event that is extended but has a definitive endpoint.
- (E) Repetitive, a repeated action or event;
- (F) Semelfactive, a single, extremely telic action, such as coughing.
- (G) Distributive, a marker that describes the distribution of items or actions.
- (H) Reversative, an action involving a turning or reversing of direction.
- (I) Diversative, which describes a "here and there" action.
- (J) Conative, which describes the attempt to perform an action.

(K) Transistional, which describes the changing from one place or state to another.

(L) Cursive, which describes the subject of an intransitive verb or the object of a transitive verb moving along in a straight line. (Young and Morgan 1992)

In this paper, we concentrate primarily on the Imperfective and Perfective Modes and their relationship to tense marking.

#### 4. Previous research on *nt'éeé*

One of the earliest researchers to work on the Navajo language was Edward Sapir, who collected and transcribed a wide range of folktales and historical narrative from Navajo speakers during the 1920's and 1930's. The folktales usually center around the infamous Coyote and his adventures. Sapir (1942) records several occurrences *nt'éeé*, used as a temporal adverb meaning 'then.' In the narrative context, *nt'éeé* signals a new event and tends to occur almost exclusively in sentence-initial position, as in the following example "The First Louse," from Sapir (1942:16).

1. *nt'ée'- ma'i haélyod zini.*

"Then Coyote came running to them they say."

*Nt'éeé* is also referred to as an adverb of past time in Young and Morgan (1987) and Midgette (1995). Young and Morgan (1987) describe *nt'éeé* as well as a "conjunctive", meaning "but, so, and, with the result that, and when." Importantly, they say that the use of *nt'éeé* as a conjunctive always occurs "in a past tense framework" (24). The following two sentences demonstrate this use of *nt'éeé*. We have included the glosses given by Young and Morgan (1987).

2. *Chidí bikee' béésh léi' bee nánishshish nt'éeé' biniká níshiizh.*

'I was poking at the tire with a knife when I poked a hole in it.'

3. *Kintahdéeé' shiít ná'ooltał nt'éeé' chidí bikee'ła'sits'áá'diitaa'.*

'I was speeding back from town when one of my tires went to pieces on me.'

Young and Morgan (1992) define *nt'ée* as a temporal-modal enclitic. These enclitics serve "a wide variety of purposes. Some connote emphasis or other shades of meaning commonly conveyed by intonation in English" (1992: 940). It is worth noting that unlike most of the other twenty temporal-modal enclitics, which are written attached to the verb as a suffix, *nt'ée* is written as an independent word. The semantics of *nt'ée* are said to be equivalent to *was* or *used to be*, and the adverbial use of *nt'ée* described in Young and Morgan (1987) is not noted in the later work. It is also interesting that in both of their example sentences, *nt'ée* occurs in sentence-final position, rather than initial position as in the Sapir (1942) texts.

4. *'ashkii nishlínéédáá' shimá sání bána 'nishkaad nt'ée.*

'When I was a boy I used to herd for my grandmother.'

5. *ánii naashá háádáá't'áá 'ákwíí ghaaí bíjł haashzheeh nt'ée.*

'I used to hunt deer every winter when I was young.'

Faltz (1998) is perhaps the first to mention *nt'ée* as a past tense marker and concludes that it occurs only with verbs of the imperfective mode if they are dependent verbs.

To summarize, *nt'ée* has been described as a temporal conjunction (Sapir 1942 and Young and Morgan 1987), as a temporal/past tense adverb (Young and Morgan 1992, Midgette 1995), and as plain past tense marker (Faltz 1998). These changes in description may mirror changes in the function of the word and are compatible with theories of grammaticization from syntactic to inflectional grams. In Section 3, we will report on the uses of *nt'ée* in current informal speech and argue that they represent a case of grammaticization.

## 5. Evidence for the Grammaticization of *nt'éeé*

As stated earlier, because grammatical morphemes develop from lexical constructions through an increase in frequency, observations of the changes in the form and meaning of the developing morpheme provide evidence for grammaticization. In addition to a reduction in substance, the grammaticizing morpheme is also likely to generalize semantically. Because of the loss of specific meanings, the morpheme will be also be used in a wider range of contexts. This process can be observed in Navajo, as *nt'éeé* undergoes much phonological reduction when produced in continuous speech and is also general enough semantically to be used by speakers in a varying number of contexts. We will now discuss each of these changes in greater detail.

### 5.1 Phonological reduction and the loss of morphosyntactic properties

Although, originally a separate lexeme, *nt'éeé*<sup>2</sup> now appears to have cliticized and often undergoes phonological reduction. According to Bybee (1985), in the process of grammaticization, lexical items (constructions) become free grammatical units (clitics, auxiliaries, etc.) and eventually develop into bound inflectional affixes. This process is accompanied by a reduction in the phonetic substance of the form. According to Hopper and Traugott (1993), as items become more grammatical, they tend to “deategorialize,” or lose their word class characteristics. Examples (6a) and (6b) illustrate this; note that in (6b) *nt'éeé* is fused to the verb and this change is reflected in a resyllabification and a change from [t'] to [d].

6a. *anists'iisí yéédaáá' nashnéé leh nt'éeé'*

6b. *anists'jada nashnélaendée'*

‘When I was young I used to play.’

---

<sup>2</sup> *Nít'éeé'* is the base form; other allomorphs include *nidéé'*, *nt'éeé'*, *ndéeé'*, *ne'é'*, and *ne'*. Our data incorporates nasal hooks as a reflection of the pronunciation of our consultant.

As has been mentioned, verbs in Navajo have affixal marking for number, person, aspect, and aktionsart. *Nt'ée'* is not a part of that system of inflectional marking. When functioning as a past tense marker, it is not always bound to the verb. Its status as a clitic is confirmed by its placement outside of the negative circumfix, as illustrated in example (7). Notice that in (7), the *do* and *da* of the negative surround the verb, but exclude the *nt'ée'*.

7) *dibe nitsaa hahoodzo biyi'di shighanígíí shimá doo bił bée hozin da nít'ée'.*

My mother didn't know I lived in Colorado.

## 5.2 Semantic generalization and pragmatic extension

Hopper and Traugott (1993) and Bybee et al. (1994) demonstrate that grammaticization involves both semantic and grammatical generalization. They point out that semantic generalization includes the development of newer, more abstract meanings as more concrete meanings are lost. From this, we can extrapolate that the item has more functions and a greater range of application. This is accompanied by an increase in token frequency, providing further impetus for the grammaticization process.

Examples (8) and (9) illustrate the use of *nt'ée'* as a past tense marker with imperfective verbs. Because these verbs are not marked as 'complete' it is our hypothesis that *nt'ée'* was used by speakers in this environment as a past tense marker. The addition of the form creates a periphrastic construction that has a completive meaning.

8. *t'ée' bíghah nahal'in nít'ée'*

'it barked all night'

9. *t'ée' bíghah dilkos nít'ée'*

'she coughed all night'

*Nt'éeé* is used with imperfective verbs to create a distinction between present imperfective, in (10a), and past imperfective, as in (10b):

10a. *ádishni'*

'I am blinking'

10b. *ádishni' n ít'éeé'*

'I was blinking.'

There is also a distinction between past imperfective and perfective as illustrated in examples (11a) and (11b). Recall from Section 3 that perfective and imperfective verbs in Navajo are marked, in part, by a change in stem shape resulting from vowel mutation and or suffixation. In order for the verb 'eat' to take *nt'éeé* speakers must use the imperfective stem. This form is illustrated in example (11a), while (11b) demonstrates the same verb with a perfective stem.

11a. *bilisana yiishaandee*

'I was eating an apple.'

11b. *bilasáana yiyá*

'I ate an apple.'

The examples in (12) show perfective, imperfective present, and imperfective past verbs. Note again the change in stem between the perfective in (12a) and imperfective in (12b) and (12c).

12a. *bilasáana nihélgizh*

'I cut the apple' (perfective)

12b. *bilasáana nihéshgéesh*

'I am cutting the apple' (imperfective present)

12c. *bilasáana nihéshgéesh nít'éeé'*

'I was cutting the apple' (imperfective past)

In example (12c), *nit'éeé*' is used to explicitly indicate that the imperfective activity occurred in the past. We see *nt'éeé* here being used to mark a semantic distinction between past and present imperfective.

*Nt'éeé* only infrequently appears with verbs in the perfective, as in example (13).

13. *Oh! Akóó deeshaal niiz'íí ndée'é*

'I was thinking, oh, I will go over there.'

*Nt'éeé* with the verb 'think' in the perfective appears to also carry the meaning of a sudden change of thought. In the examples in (14), however the use of *nt'éeé* with the perfective verb seems to express a middle or resultative meaning. In all of the examples we collected, *nt'éeé* only occurred following the verb.

14a. *Yibeezh nit'éeé'*

'it was boiling' (past imperfective)

14b. *Shibeezh*

'it boiled' (perfective)

14c. *Shibeezh nit'éeé'*

'it was boiled'(resultative/middle)

## 6. The layering of *nt'éeé*

Further evidence to support the claim that *nt'éeé* has grammaticized can be gained from observing the range of occurrences of the morpheme. Although new meanings have been gained from the new contexts in which the form occurs, older, more specific meanings of the form have not yet been lost, despite the growing frequency with which the construction is used in more generalized environments. The grammaticization process has resulted in a layering of the older meanings alongside the newer meanings, a cross-linguistic phenomena common to documented

cases of grammaticization (Bybee et al. 1994). These older meanings give an indication of the stages of the form's development. The following sections discuss some of the older meanings of *nt'éeé*, such as its ability to indicate 'surprise' and mark causation. We also discuss how the construction is used to form conditionals.

### 6.1 Surprise

*Nt'éeé* can be used to indicate an element of surprise in an utterance. When *nt'éeé* is used in this context, it appears as a clause linker, following the verb in the matrix clause. Examples (15) and (16) illustrate this meaning.

15. *ch'iníyáa nít'éeé' chahalheel*

'I walked outside (and all of a sudden) it was dark.'

16. *ha'í'áq nít'éeé' chahalheel*

'The sun rose (but, unexpectedly) it was dark.'

This use was not common in our data and needs to be further investigated in order to determine its relation to the grammaticization path.

### 6.2 Causation

*Nt'éeé* can also be used to indicate a level of causation. Just as was the case with its use as a marker of 'surprise', when used as a causative marker, it occurs as a clause linker, following the verb of the first clause. Furthermore, *nt'éeé* is used in place of the coordinating conjunction *dóó*. Examples (17a) and (17b) demonstrate this use of the morpheme.

17a. *naa'iitlzh nít'éeé' haachá*

'She started crying because she fell down.'

17b. *naa'iitlzh dóó haachá*

'She fell down and started crying.'

Although both *nt'ée* and *dóó* can be used as conjunctions, these forms differ in semantic meaning, as is illustrated in examples (18a) and (18b). These examples demonstrate the use of *nt'ée* to indicate immediacy and surprise (18a) in contrast to *dóó* in example (18b) which does not express these concepts.

18a. *naa'eelts'id nít'ée'*

'it (bottle) fell down (right here) and was broken (not my fault)

18b. *naa'eelts'id dóó sít's'il*

'it (bottle) fell down '(elsewhere) and was broken (I'm just telling you about it)'

### 6.3 The Conditional

One of the most important additions to the analysis of *nt'ée* in Young and Morgan (1992) is the documentation of a conditional tense from the use of *nt'ée* in combination with *doo* and *dooleeł*. This use is illustrated in example (19).

19. *As shibee'eldqoh holóqgo k'ad bíjh haashzheeh dooleeł nt'ée'*

'I'd be hunting deer now if I had a gun.'

This conditional use can be extended to the English glosses *would*, *could*, and *should* when used in a past tense framework, as in example (20) which comes from our modern Navajo data.

20. *Atsá jiiłtsáqadoo nít'ée'*

'I should have seen the eagle.'

This sentence would be the same whether the speaker used *should*, *could*, or *would*. This contrasts with example (21) which uses the perfective mode and thus has no conditional meaning.

21. *atsá yiłtsá*

'I had seen the eagle.'

## 7. Conclusions

An examination of the current function and distribution of the morpheme *nt'éeé* in Navajo has shown that the form has undergone grammaticization and that in doing so, has become phonologically reduced and semantically general. A layering of older meanings is present, which also gives us an indication of the path that the morpheme has traveled along throughout the grammaticization process. Moreover, *nt'éeé* has come to mark past tense within the temporal system of Navajo. The development of a tense marker in Navajo is an important finding, since Athabaskan languages do not typically include tense within their temporal systems.

The appearance of *nt'éeé* in a wide range of contexts and with such high frequency indicates that speakers of Navajo are now using this morpheme productively. The fact that Navajo has had a major change in the temporal system during the last century, despite being threatened by dominant languages, demonstrates the vitality and perseverance of its speakers. Language change happens as a result of language use and the ability of the speakers to express new ideas in their language ensures the transmission of that language to future generations.

These findings also have wider implications for linguistic theory. Our evidence regarding the use of *nt'éeé* supports the cross-linguistic trends that have been found to occur in other cases of grammaticization and language change, whereby semantically general forms extend to new contexts and also lose phonetic substance (Bybee et al. 1994, Hopper and Traugott 1993). The origin of *nt'éeé*, and its original syntactic position have yet to be determined, and would be a good area for future research. A diachronic study of such phenomena might be compared to the development path proposed by Bybee et. al (1994: 105) where resultative grams develop into anteriors before becoming markers of perfective/past. Evidence of such a development path for *nt'éeé*, would further confirm the cross-linguistic tendencies of grammaticization.

## References

- Axelrod, Melissa. 1993. *The Semantics of Time: Aspectual Categorization in Koyukon Athabaskan*. University of Nebraska Press.
- , 2004. Navajo and Athabaskan Eyak. In Philipp Strazny (eds). *Encyclopedia of Linguistics*. Routledge.
- Bybee, Joan. 1985. *Morphology: A Study of the Relation Between Meaning and Form*. Philadelphia: John Benjamins Publishing Company.
- 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.
- Faltz, Leonard M. 1998. *The Navajo Verb: A Grammar for Students and Scholars*. Albuquerque: University of New Mexico Press.
- Midgette, Sally. 1995. *The Navajo Progressive in Discourse: A Study in Temporal Semantics*. New York: Peter Lang.
- Sapir, Edward and Harry Hojier. 1942. *Navajo Texts*. William Dwight Whitney Series. Baltimore: Linguistic Society of America.
- Young, Robert W. 2000. *The Navajo Verb: An Overview*. Albuquerque: University of New Mexico Press.
- Young, Robert W. and William Morgan, Sr. 1987. *The Navajo Language: A Grammar and Colloquial Dictionary*. Albuquerque: University of New Mexico Press. Revision of Young and Morgan 1984.
- Young, Robert W., William Morgan, Sr., and Sally Midgette. 1992. *Analytical Lexicon of Navajo*. Albuquerque: University of New Mexico Press.

## Introducing San Agustín Mixtepec Zapotec

Rosemary G. Beam de Azcona  
University of California, Berkeley

### 1. Introduction & Endangerment status

With this paper I introduce a new language, San Agustín Mixtepec Zapotec (SAMZ). This paper contains background information on SAMZ and its speech community. I highlight SAMZ data for historical and comparative implications within Southern Zapotec (SZ), including both segmental and suprasegmental correspondences. As this is a working paper, I also feature two not-yet fully analyzed but intriguing problems in the synchronic phonology of SAMZ having to do with nasal-obstruent clusters and tonal morphology respectively. The information given in this paper is also especially valuable because of the lack of previous documentation and because of SAMZ's impending death.

SAMZ is spoken in the single town of San Agustín Mixtepec, in the ex-district of Miahuatlán in the Southern part of the state of Oaxaca, Mexico. SAMZ is today only spoken fluently (when spoken at all) by two elderly residents of the town for which the language is named. There are reportedly also two semi-speakers. The information in this paper comes from the youngest fluent speaker, who is now 72 years old. The findings presented here are very preliminary, the result of only three weeks' total fieldwork in 2002 and 2003<sup>1</sup>. More long-term fieldwork is planned for the next three years.

This community apparently agreed to stop speaking Zapotec in 1965. In addition to the undeserved stigma formally assigned to the language at that time, immigration out of SAM has no doubt been a key factor in this language's demise. Modern land disputes have contributed to the town's loss of resources and further loss of population. Yet the exodus out of SAM began much earlier with a dispute that divided the townspeople into two groups, with one group going South to found San Agustín Loxicha (SAL) in the ex-district of Pochutla (see <http://www.laneta.apc.org/rio/loxicha/historia.htm>). Both of these San Agustín towns have stories about the statue of the patron saint of each town being switched. Each version has the opposite town coming out the winner. In the SAM version this switch is the cause of the SAL's population increase and SAM's population decrease.

### 2. History, Geography & Classification

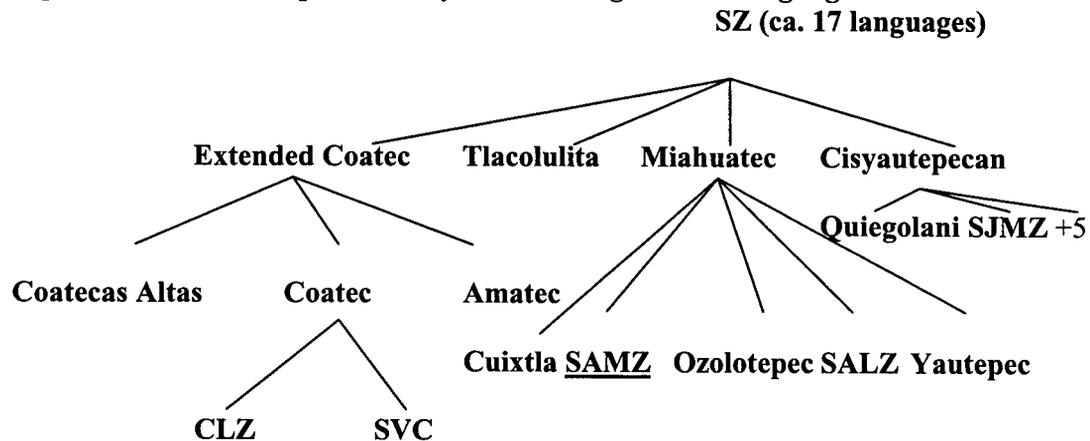
SAMZ belongs to the Miahuatec subgroup of Southern Zapotec (<Zapotecan < Otomanguan). According to Smith Stark (2003), SZ languages are defined by the addition of an /m/ at the beginning of words for animals and other words which take *b*- in other Zapotec languages. This renders the prefix *mb*- in Miahuatec languages such as SAMZ. Miahuatec languages differ from other SZ languages by having *t* and *d* reflexes of the Proto-Zapotec (PZ) phonemes reconstructed as *\*ss* and *\*s* (Kaufman, 2003).

According to Rojas (1950), SAM was founded by people from Miahuatlán, and the linguistic evidence supports this, although the town's enemies say that it was founded by people from Sola de Vega, an excuse for taking land away. Miahuatlán was supposedly founded by a group from Coatlán (Rojas, 1950), an assertion which, if accurate, would suggest that Miahuatec is most closely related to the Coatec group within SZ.

The Miahuatec group of languages gained significant ground with the southward expansion led by emigrants from SAM. The language of SAL, which must have developed from SAMZ, is today spoken vigorously in several towns near the Pacific coast, such as Cozoaltepec, Santo Domingo de Morelos, and many of the towns named Loxicha. Linguistic evidence of a more northerly origin for that southernmost Miahuatec language comes in the form of a borrowed placename. The SALZ name for the town of Santa María Colotepec, is clearly a borrowing from Coatlán-Loxicha Zapotec (CLZ), and specifically from a northern dialect of that language (Beam de Azcona, 2004), one which lies much closer to SAM than to SAL. Thus, SAMZ is a language which historically links the languages of the northern and southern parts of the Miahuatec region.

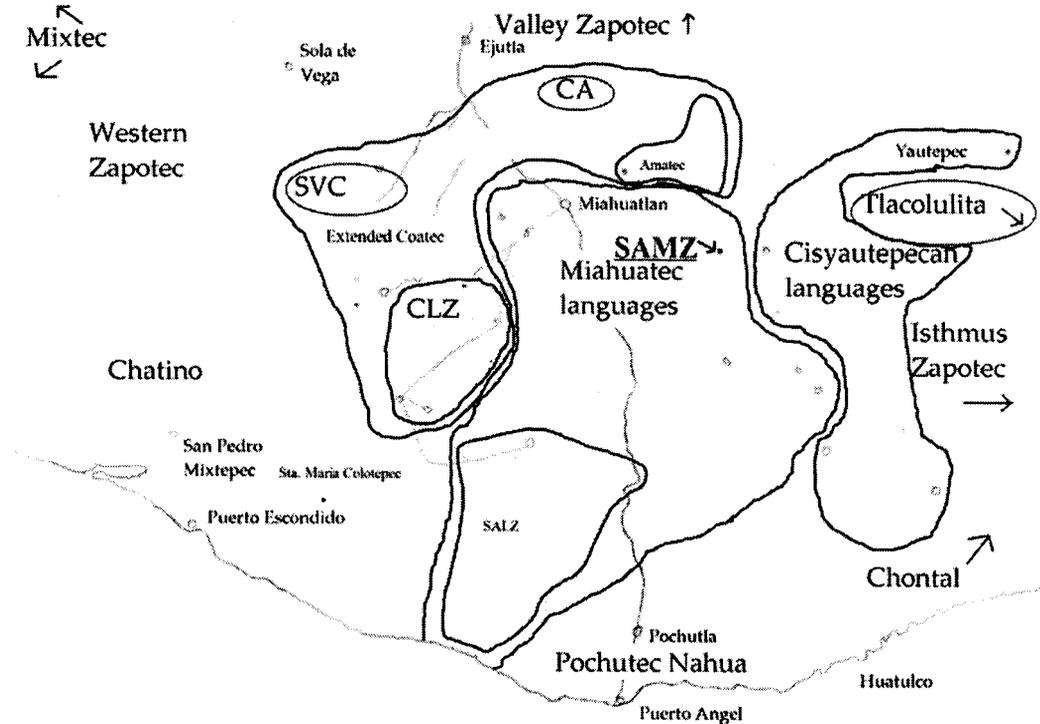
SAM is surrounded by and politically linked to towns where the use of other Zapotec languages is vigorous. SAM is subject to the municipality of San Cristobal Amatlán, where some 86.5% of the residents speak Amatec, an “extended Coatec” language also spoken in several other towns (Smith Stark, 2003). The similarly named language of San Juan Mixtepec Zapotec (SJMZ) also shares a geographic border with SAMZ and belongs to a third subgroup of SZ, Cisyauatepecan. It is more closely related to Quiegolani, Xanica, Xanaguía, Quierí, Lapaguía and the very endangered Xadani Zapotec. Cisyauatepecan languages are defined by Smith Stark as those which have lost the *b* and have only an *m*- prefix for most animal words. Many of the languages mentioned so far are shown in relation to SAMZ in Figure 1 (based on Smith Stark, 2003).

**Figure 1: Southern Zapotec family tree showing selected languages**



Each of the main groups shown in Figure 1 is located geographically, along with SAMZ, and some other individual languages and towns referenced here, in the map in Figure 2. The location of towns in this map is based on a map found in Rojas (1950) with supplementary information from the modern INEGI maps. The linguistic boundaries shown here are based on Smith Stark (2003) and on my own fieldwork with these languages.

**Figure 2: SAMZ and major subgroupings of SZ**



**3. Sounds**

Here I give a brief account of the phoneme inventory. Phonology is covered in 4.

**3.1 Segments**

SAMZ has a six vowel system, as shown in Figure 3.

**Figure 3: SAMZ vowels**

i		u
e		o
ë [ɛ ~ æ]		
	a	

SAMZ has three types of obstruents: voiceless plosives<sup>2</sup>, voiceless spirants and voiced spirants. Voiced plosives do occur phonetically as part of the prenasalized plosive series which I tentatively posit here. My analysis of this series is as yet incomplete and I cover the details of the problem in 4.2.2. The full consonant inventory of SAMZ is shown in the practical orthography in Figure 4, with marginal and loan phonemes in parentheses

Figure 4: SAMZ consonants

	Labial	Dental	Alveolar	Palatal	Retroflex	Velar	Labiovelar	Glottal
Voiceless plosives	p	t	tz	(ky)	ch	k	kw	
Voiceless spirants		th	s		x			(h)
Voiced spirants	b	d				g		
Prenasal-plosives	mb		ndz		ndzh	ng	ngw	
Nasals	m	n				(nh)		
Liquids			r, (rr)	(ly)	l			
Glides	w			y				

### 3.2 Suprasegmentals

SAMZ has three tones: high, low, and rising. The SAMZ low tone is phonetically mid-to-low falling, the same as in the northern dialects of CLZ.

In addition to tone there is a three-way contrast involving glottalization: plain V, rearticulated V<sup>7</sup>V, and checked V<sup>7</sup>. These terms are fairly self-explanatory. Plain vowels lack glottalization. Rearticulated vowels are interrupted by a glottal stop which is followed by an echo of the same vowel (this is phonologically one segment, I do not mean to say that there is a vowel plus a second echo vowel in terms of the number of segments represented). Checked vowels are vowels that end in a glottal stop.

All three tones can occur on plain and checked vowels but only level tones can occur on rearticulated vowels. The pre-glottal stop portion of a checked vowel is longer in duration than the pre-glottal stop portion of a rearticulated vowel, and this is likely the reason why the language's single contour tone does not occur on rearticulated vowels. The possible combinations of tone and vowel type are shown with examples in Figure 5.

Figure 5: Combinations of tone and glottalization in SAMZ

	Plain vowels	Checked vowels	Rearticulated vowels
Low tone	<i>làg</i> 'hoja; leaf'	<i>mbë<sup>7</sup>l</i> 'culebra; snake'	<i>yë<sup>7</sup>ël</i> 'hondura; swimming hole'
High tone	<i>lár</i> 'ropa; clothing'	<i>xí<sup>7</sup>l</i> 'ala; wing'	<i>xí<sup>7</sup>il</i> 'algodón; cotton'
Rising tone	<i>ngǒ</i> 'huevo; egg'	<i>ngǒ<sup>7</sup>n</i> 'buey; ox'	Not attested

## 4. Phonology & Morphology

In this section I provide selected details of SAMZ phonology and morphology, including some interesting but unresolved morphophonological issues.

### 4.1 Possession

There is some interesting tonal morphology involved in marking a first person singular possessor on a possessed noun in this language, described in 4.1.2 below. Details of the morphology of possession marking are also potentially enlightening in a problem

of phonological analysis discussed in 4.2.3. In order to better understand the implications discussed below, I give a basic description of how possession works in this language.

#### 4.1.1 Segmental marking of possession

The morphology and syntax of possessed noun phrases found in SAMZ is fairly typical of Zapotec languages as a whole, though individual languages have variations on these patterns (as in CLZ, see Beam de Azcona, 2004). Possessors follow possessed nouns in these left-headed languages. Zapotec languages typically display a difference between alienable and inalienable possession. As in many of these languages (see, for example, Pickett et al., 1998) in SAMZ alienable possession is marked with an *x*-initial prefix. This pattern affects loanwords as well as native words, as shown in (1). Inalienable possession in SAMZ only differs by the lack of this prefix, as shown in (2), and does not make use of a separate noun or prefix as some languages do (e.g. CLZ).

- |     |              |                                       |               |              |                                   |
|-----|--------------|---------------------------------------|---------------|--------------|-----------------------------------|
| (1) | <i>kũch</i>  | <b>'cochino; pig'</b>                 | <i>xkũch</i>  | <i>lè</i>    | <b>'tu cochino; your pig'</b>     |
|     | <i>kwě'l</i> | <b>'totomoxtle; cornhusk'</b>         | <i>xkwě'l</i> | <i>mě yá</i> | <b>'su totomoxtle; his husk'</b>  |
| (2) | <i>làb</i>   | <b>'huarache; sandal'<sup>3</sup></b> | <i>làb</i>    | <i>lè</i>    | <b>'tu huarache; your sandal'</b> |
|     | <i>bě'l</i>  | <b>'hermana; woman's sister'</b>      | <i>bě'l</i>   | <i>mě yá</i> | <b>'su hermana; her sister'</b>   |

As in other, for example Valley, Zapotec languages (see Suárez, 1973), nouns beginning in lenis consonants in the citation form have root variants with the corresponding fortis consonant after *x*-, as shown in (3).

- |     |             |                       |              |           |                             |
|-----|-------------|-----------------------|--------------|-----------|-----------------------------|
| (3) | <i>bě'l</i> | <b>'llama; flame'</b> | <i>xpě'l</i> | <i>nà</i> | <b>'mi llama; my flame'</b> |
|     | <i>yèt</i>  | <b>'olla; pot'</b>    | <i>xkèt</i>  | <i>lè</i> | <b>'tu olla; your pot'</b>  |

The nasal portion of an animacy prefix is often lost following *x*-, as shown in (4).

- |     |                         |                            |                         |              |   |
|-----|-------------------------|----------------------------|-------------------------|--------------|---|
| (4) | <i>mbèr</i>             | <b>'guajolote; turkey'</b> | <i>xpèr</i>             | <i>mě yá</i> | <b>'su guajolote; her turkey'</b>         |
|     | <i>ngõ<sup>7</sup>n</i> | <b>'buey; ox'</b>          | <i>xkõ<sup>7</sup>n</i> | <i>mě yá</i> | <b>'su buey; his/her ox'</b>              |
|     | <i>ngwěñ</i>            | <b>'dedo; finger'</b>      | <i>xkwěñ</i>            | <i>lè</i>    | <b>'tu dedo; your finger'<sup>4</sup></b> |

#### 4.1.2 Tonal marking of possession by the first person singular

In most Zapotec languages a floating high tone is added to verb roots with 1s subjects and noun roots with 1s possessors (see, for example, Beam de Azcona, 2004; Bickmore & Broadwell, 1998). In CLZ this floating H is added as a suffix:

- |     |               |                           |  |      |    |     |
|-----|---------------|---------------------------|--|------|----|-----|
| (5) | <i>lid</i>    | <b>'casa; home'</b>       |  |      |    |     |
|     | <i>lĩd nâ</i> | <b>'mi casa; my home'</b> |  |      |    |     |
|     |               |                           |  |      |    | CLZ |
|     |               |                           |  | home | 1s |     |

In SAMZ the floating high tone affects both both types of possessed nouns equally.

- |     |              |                      |                  |                        |
|-----|--------------|----------------------|------------------|------------------------|
| (6) | <i>kwě̀l</i> | <i>xkwě̀l mě́ yá</i> | <i>xkwě̀l ná</i> | Alienable possession   |
|     | ‘totomoxtle’ | ‘su totomoxtle’      | ‘mi totomoxtle’  |                        |
|     | ‘cornhusk’   | ‘his cornhusk’       | ‘my cornhusk’    |                        |
| (7) | <i>lě̀</i>   | <i>lě̀ lè</i>        | <i>lě̀ ná</i>    | Inalienable possession |
|     | ‘nombre’     | ‘tu nombre’          | ‘mi nombre’      |                        |
|     | ‘name’       | ‘your name’          | ‘my name’        |                        |

In CLZ when the targeted root’s tone ends high (i.e. has a high or rising tone), the floating high tone morpheme has no effect on the surface form, as shown in (8).

- |     |             |                   |                |                         |     |
|-----|-------------|-------------------|----------------|-------------------------|-----|
| (8) | <i>bě́l</i> | ‘hermana; sister’ | <i>bě́l ná</i> | ‘mi hermana; my sister’ | CLZ |
|-----|-------------|-------------------|----------------|-------------------------|-----|

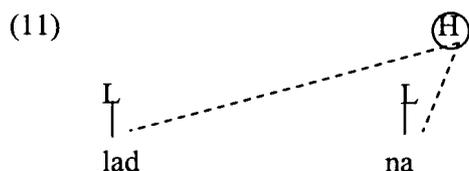
In SAMZ a root whose tone ends high is likewise unaffected by the floating high tone. However, in this case the pronoun itself undergoes a change in tone. Note that this tonal effect on the pronoun must be the result of the same floating high tone involved in first person singular marking and cannot be the result of some kind of tone sandhi caused by the concatenation of a certain type of root and a low-toned CV pronoun because the second person familiar pronoun is unaffected in the same context. In (9) for comparison I give the rising-toned bare root, and the second person and first person singular possessed forms. This behavior of the floating high tone in SAMZ makes it seem more like a clitic than an affix (which is how I’ve analyzed the cognate tonal marker in CLZ) because although tones are necessarily phonologically bound, the domain of this floating high tone seems to be the noun phrase rather than the head noun itself.

- |     |                  |                          |                        |
|-----|------------------|--------------------------|------------------------|
| (9) | <i>lǒs</i>       | <i>lǒs lè</i>            | <i>lǒs ná</i>          |
|     | ‘lengua; tongue’ | ‘tu lengua; your tongue’ | ‘mi lengua; my tongue’ |

The docking of the floating high tone on the first person singular pronoun is consistent, predictable, and easy to explain in the environment just described, an environment in which it would be phonologically impossible for this tonal morpheme to mark the possessed noun. The behavior of the floating high tone with low-toned possessed noun roots, as I described it above seems equally unsuspecting and natural but in reality in that environment it is more complicated. Sometimes when a low-toned root is possessed, the floating H docks on both the root and the pronoun, as shown in (10).

- |      |                |                        |                      |
|------|----------------|------------------------|----------------------|
| (10) | <i>làd</i>     | <i>làd lè</i>          | <i>làd ná</i>        |
|      | ‘cuerpo; body’ | ‘tu cuerpo; your body’ | ‘mi cuerpo; my body’ |

To account for this one could posit a floating high tone that targets all the word-final low tones in the possessed noun phrase, as modeled in (11).



The problem is predicting when both root and pronoun will be affected vs. the root alone. I have considered historical, phonological, and semantic explanations. For example, some SAMZ low tones correspond to CLZ low tone and others to CLZ falling tone, but there are both kinds of correspondences in each group of SAMZ words.

(12)	CLZ	SAMZ	1s possessed	
'sal; salt'	<i>zèd</i>	<i>tèd</i>	<i>těd nà</i>	<b>Head noun only</b>
'nombre; name'	<i>lě^</i>	<i>lě̀</i>	<i>lě̃ nà</i>	
'diente; tooth'	<i>ndě`y</i>	<i>lèy</i>	<i>lěy nă</i>	<b>Pronoun also</b>
'cuerpo; body'	<i>lâz</i>	<i>làd</i>	<i>lăd nă</i>	

The difference cannot be attributed to a difference in alienability of possession. Both types of possessed noun phrases show the same diversity, with the first person singular pronoun taking the floating high tone along with some low-toned possessed nouns but not with others, as shown in (13).

(13)	Gloss	Noun in isolation	1s possessed	Type of possession
	'huarache; sandal'	<i>làb</i>	<i>lăb nă</i>	inalienable
	'nombre; name'	<i>lě̀</i>	<i>lě̃ nà</i>	inalienable
	'guajolote; turkey'	<i>mbèr</i>	<i>xpěr nă</i>	alienable
	'comal; griddle'	<i>sìl</i>	<i>xsìl nà</i>	alienable

Of the documented SAMZ low-toned nouns, those that remain glottalized with 1s possession never allow a tonal change on the accompanying 1s pronoun. Deglottalization with first person singular marking is also found on a small number of verbs in CLZ (Beam de Azcona, 2004), and on alienably possessed nouns in Lachixío Zapotec (Sicoli, 1998). Thus far I have seen deglottalization with first person singular marking on too few roots in SAMZ to tell whether it is predictable or not. The fact that the pronoun is unaffected tonally when the root stays glottalized suggests that the floating high tone first docks onto the possessed noun root and only spreads to the following pronoun if there is no glottal stop following the root vowel to block the spreading. This generalization is based on relatively few examples, however. In (14) I show two nouns with checked vowels, exhibiting the generalization that when there is deglottalization the high tone can spread and when there is not the glottalization blocks any spreading. Note though, that my notion of spreading here is not linear because it is not as though the high tone is simply copied and spreads one pace rightward. The high tone cannot concatenate in front

of a low tone in the same word, here the pronoun, and surface that way because there are no phonologically falling tones in this language (though the low tone does fall phonetically). Since the sequence HL is impossible in the language, the kind of spreading that takes place when the high tone docks on both noun and pronoun may involve tonal metathesis on the pronoun.

- (14) ‘barriga; belly’     $lè^7n$     →     $lěn nǎ$     (deglossalized)  
           ‘papel; paper’     $yè^7tz$     →     $xkè^7tz nà$     (not deglossalized)

There are a small number of, perhaps irregular, low-toned nouns which either show no tonal alternation on the root or have a surface high rather than rising tone when possessed by 1s, and which do not allow tonal alternation on the accompanying pronoun. I show two of these in (15).

- (15) ‘yerno; son-in-law’     $xùs$     →     $xús nà$   
           ‘llama; flame’     $bě^7l$     →     $xpě^7l nà$

For unglottalized low-toned nouns with surface rising tone when possessed by the first person singular, I have not found a phonological predictor for when the pronoun will be affected and when it will not.

- | (16) | Floating H affects root only    |  | Also affects pronoun       |
|------|---------------------------------|--|----------------------------|
|      | $blàn$ ‘carbón; coal’           |  | $ditz$ ‘espalda; back’     |
|      | $kwě^7l$ ‘totomoxtle; cornhusk’ |  | $gòn$ ‘limosna; alms’      |
|      | $lě^7$ ‘nombre; name’           |  | $kwàt$ ‘cachete; cheek’    |
|      | $mbyìn$ ‘pájaro; bird’          |  | $làb$ ‘huarache; sandal’   |
|      | $sìl$ ‘comal; griddle’          |  | $làd$ ‘cuerpo; body’       |
|      | $tèd$ ‘sal; salt’               |  | $lèy$ ‘diente; tooth’      |
|      | $xlě^7$ ‘fruta; fruit’          |  | $lò$ ‘cara; face’          |
|      | $xnì$ ‘luz; light’              |  | $mbèr$ ‘guajolote; turkey’ |
|      | $yù$ ‘tierra; soil’             |  | $mbě^7 kw$ ‘perro; dog’    |
|      | $yùx$ ‘arena; sand’             |  | $mbě^7 l$ ‘pescado; fish’  |

It is unclear at this time whether a phonological solution will present itself or whether these lexical items fall into classes based on their morphological behavior when marked for a first person singular possessor. Nevertheless I include the above information for those who are interested in seeing fresh data from this Zapotec language, which I expect will differ in this respect from other Zapotec languages which some readers will be familiar with.

#### 4.2 Segmental phonology

In this section I give two clear rules of segmental allophony and I explore the evidence for positing a separate series of prenasalized plosive phonemes.

#### 4.2.1 Devoicing

SAMZ has word-final devoicing, as seen in (17).

(17) *låg* [lax̥] ‘hoja; leaf’      *lár* [lax̥] ‘ropa; clothing’

There is also anticipatory voicing assimilation when /β/ in a prefix precedes a voiceless segment in a root, as shown in (18).

(18) *bdǒ* [βðǒ] ‘plátano; banana’    vs.    *bthë* [βθè] ‘epazote’  
*bchú’ux* [βčúʔux] ‘jitomate; tomato’

The *b* in the *mb*- animacy prefix also assimilates to the voicelessness of following consonants.

(19) *mblàn* ‘liebre; jackrabbit’    vs.    *mpsîn* ‘venado; deer’  
*mptîn* ‘ratón; mouse’

The same thing happens with the homophonous class A<sup>5</sup> completive prefix. This prefix is shown before roots with initial voiced and voiceless consonants in (20).

(20) Completive	+	Verb stem	=	Surface form
<i>mb-</i>		<i>-dí’b</i> ‘costurearlo; sew (it)’		<i>mbdí’b</i>
<i>mb</i>		<i>-rò’o</i> ‘salir; go out’		<i>mbrò’o</i>
<i>mb-</i>		<i>-tè’ed</i> ‘aprender; learn’		<i>mptè’ed</i>
<i>mb-</i>		<i>-thà’as</i> ‘aplastarse; get smashed’		<i>mpthà’as</i>

#### 4.2.2 Prenasalized stops and affricates

As defined by Smith Stark (2003), SZ languages have nasal-initial animal words. In all branches but Cisyatepecan these words in fact begin in nasal-obstruent sequences. Besides the animacy prefix, SAMZ (a Miahuatec language) and CLZ (a Coatec language) both have homorganic nasal-obstruent sequences in the habitual and completive aspect prefixes. One might ask whether these sounds are truly clusters or whether they are an independent series of phonemes. In SAMZ these would be /<sup>m</sup>b, <sup>n</sup>dz, <sup>n</sup>j, <sup>n</sup>g, <sup>n</sup>gw/.

There are no voiced stop or affricate phonemes in SAMZ. Under either analysis some synchronic or historical rules must be posited. These include homorganic nasal assimilation, occlusion of voiced fricatives and or voicing of voiceless plosives. In the case of *mb* and *ng* a phonological rule could cause occlusion of /β/ and /ɣ/. This is certainly what happened historically since words which here have *mb* and *ng* are

reconstructed with lenis *\*p* and *\*k* (Kaufman, 2003). *Ngw* results historically from the earlier marker /ko/ cited by Córdova (1578) and reconstructed by Kaufman (2003). The labialization is a reduction of the vowel. Synchronically, the cluster analysis is more problematic for *ndz* and *ndzh* [<sup>n</sup>dz, <sup>n</sup>ʒ] than for *mb*, *ng*, *ngw*. While there are voiced fricatives in SAMZ that correspond to *mb* and *ng(w)*, there are no voiced sibilants. A synchronic analysis would have to involve either voicing of the voiceless affricates *tz* and *ch* or the voicing of the sibilants *s* and *x* with additional stop epenthesis.

If these sounds are underlying clusters we might expect to find evidence that they can result synchronically from the concatenation of the component phonemes and that there can also be synchronic reduction or deletion of one of the component phonemes. There are nouns sharing a derivational relationship which differ only by the initial nasal, e.g. *bē'l* 'llama; flame' and *mbē'l* 'estrella; star.' On the surface these would seem to suggest that *m* is added to form the more animate word, but in fact the animacy prefix is *mb-* and in all the cases I've found the less animate word is *b-*initial so it could be that the *b* of the prefix causes the initial /β/ of the root to delete since [mbβ] does not occur. As seen in (20) above, the *b* of the *mb* prefix is devoiced while the *m* is not, preceding voiceless sounds. Better evidence is the fact that the *b* of the animacy prefix can delete altogether. In SAMZ I have only seen this happen before *k*, as shown in (21).

(21)	Classifier	+	Noun root	=	Surface form
	<i>mb</i>		<i>kĩ</i> 'liendre; nit'		<i>mkĩ</i>

In SAMZ, the habitual aspect marker [<sup>n</sup>ʒ] also reduces before consonant-initial verb roots but instead of reducing to a plain nasal *n* it reduces to a prenasalized stop *nd*. The sequence [nd] without any following sibilant, only occurs in this reduced form of the completive aspect marker and not in any other word in the language that I have seen.

(22)	Habitual	+	Verb root	=	Surface form	SAMZ
	[ <sup>n</sup> ʒ]		- <i>àth</i> 'morir; die'		[ <sup>n</sup> ʒ àθ]	
	[ <sup>n</sup> ʒ]		- <i>bó'o</i> 'sacar; take out'		[ <sup>n</sup> dβó'o]	

Since the cluster analysis is stronger for *mb*, *ng*, *ngw* than for *ndz* and *ndzh*, I look now to the historical origins of the latter two sounds. Instances of [<sup>n</sup>ʒ] in SAMZ are cognate with *nd* in CLZ. There are two historical sources for this consonant (I take PZ forms from Kaufman, 2003), *\*(nV+)l* and *\*(nV)+ty*. Lenis *\*ty* otherwise usually reflects as /t/ in SAMZ, as in the verb 'salir; go out' in (20) above, except when it preceded *\*i*, as in 'squirrel.' This split took place in all SZ languages except in Coatepec proper. In other Miahuatec languages for which I have collected data, (SAL, Santa Cruz Xitla, San Sebastián Río Hondo), the reflex of *\*ty* before *\*i* is /z/, e.g. '7' *\*katyi* > *yaz* and 'hogar; home' *\*lityi* > *liz*. In SAMZ there is no /z/ and instead the reflex is /s/. In some Cisyauatepecan languages, including SAMZ's closest neighbor, San Juan Mixtepec, the reflex of *\*ty* before an earlier front vowel is /dz/, the same as in the SAMZ word for

‘squirrel.’ Instances of [ʰdz] in SAMZ usually correspond to CLZ *nzh* [ʰj̃] and are either reconstructed with \*y or retroflex \*ʂ.

**Figure 6: PZ origins of and CLZ correspondences with SAMZ [ñ] and [ndz]**

PZ	gloss	CLZ	SAMZ
*ni+lana	‘tizne; soot’	<i>ndān</i>	[ʰj̃ān]
*lawo	‘cara; face’	<i>ndô</i>	<i>lõ</i>
*nV+tye?	‘aquí; here’	<i>ndē˘ ~ ndē˘ʔ</i>	[ʰj̃éʔe]
*kwe+tyi:ʔ(s)sa	‘ardilla; squirrel’	<i>ndi˘z</i>	[ʰdziʔit]
*ti+yaka(?)	‘oreja; ear’	<i>nzhâ</i>	[ʰdzàx]
*ʂ-okwa?	‘maíz en grano; dried corn kernels’	<i>nzhöp</i>	[ndzõʔβ]

The sounds *mb*, *ng*, and *ngw* seem to be more recently formed or transparent clusters, while the sounds *ndz* and *ndzh* seem to be less analyzable and more like single segments.

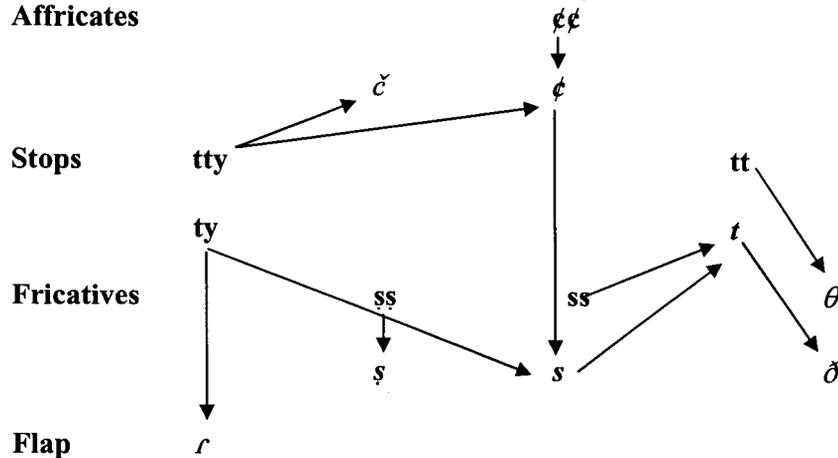
### 5. Historical and comparative data of interest

The segment found initially in the word ‘mouth’ is reconstructed by Swadesh (1947) and Fernández de Miranda (1965) as \**r*, by Suárez (1973) as \**t* and by Benton (1988) and Kaufman (2003) as \**ty*. Most modern Zapotec languages, including SAMZ, have /*r*/ here. In (Beam de Azcona, 2001) I argue that the change from \**ty* to /*r*/ is post-contact and in fact related to contact with Spanish which itself has an alveopalatal flap. Evidence that SAMZ probably did not have /*r*/ at the time of contact with Spanish is found in the loanword *lénkwa* ‘jorobado; hunchback’ < Spanish *renco* ‘limping, gimp.’

Changes in coronal obstruents in Southern Zapotec are interesting and often involve chain shifts. Sounds that have been reconstructed by Kaufman (2003) are given in bold and sounds that exist in modern SAMZ are italicized:

**Figure 7: Coronal obstruents from Proto-Zapotec to SAMZ**

#### Affricates



Most of the changes in Figure 7 took place in other Miahuatec languages too. The only changes from Figure 7 that also took place in Coatec languages are \*šš > š, \*tt > θ, and \*t > ð. The chain shift in which ty > s > t > ð is thus probably a drag chain, beginning before the breakup of Miahuatec and Coatec.

Few attempts have been made to reconstruct tone in Proto-Zapotec or its daughters, mostly due to lack of reliable data. The first attempt was made by Swadesh (1947) and the most recent by Benton (2002). Having studied the tone system of Coatlán-Loxicha Zapotec intensely since 1996, and now embarking on this study of SAMZ, I hope that by comparing the suprasegmental systems of both languages I will be able to make my own contribution to the diachronic study of tone in this family. The system of suprasegmental contrasts found in SAMZ appears more conservative than that of CLZ. There are fewer tones in SAMZ and more phonation type contrasts, something more like the system of languages like Isthmus Zapotec (e.g. see Pickett, 1959). In Figure 8 are the results of a survey of 220 cognates between these two languages. Tonal correspondences are represented in rows and vowel type correspondences are represented in columns. Symbols on the left of a tilde refer to SAMZ and on the right of the tilde to CLZ. Thus, the third box in the second row indicates that there are two words which have low tone in SAMZ corresponding to falling tone in CLZ which also have a checked vowel in SAMZ and a plain vowel in CLZ. Since glottalization acts as a tone in CLZ but can be cognate with glottalization in SAMZ (which is not tonal), it is indicated along both axes.

**Figure 7: Suprasegmental combinations in correspondence---SAMZ ~ CLZ**

	V ~ V	V ~ V <sup>ʔ</sup>	V <sup>ʔ</sup> ~ V	V <sup>ʔ</sup> ~ V <sup>ʔ</sup>	V <sup>ʔ</sup> V ~ V	V <sup>ʔ</sup> V ~ V <sup>ʔ</sup>	Total
L ~ L	31		5		8		44
L ~ F	46		2		3		51
L ~ R	1		2				3
L ~ ?				8		10	18
R ~ R	14		11				25
R ~ F	11		1				12
R ~ L	5		2				7
R ~ H	2						2
R ~ ?		6		2		1	9
H ~ F	10		1		3		14
H ~ H	1				2		3
H ~ ?		3		4		22	29
H ~ R			1		1		2
H ~ L					1		1
<b>Total</b>	<b>121</b>	<b>9</b>	<b>25</b>	<b>14</b>	<b>18</b>	<b>33</b>	

I have not made a reconstruction of tone for the ancestor of these two languages, but certain significant gaps and tendencies can be found in the table above from which we can make some generalizations.

More than 80% percent of CLZ words with ? tone have some type of glottalization in SAMZ. Unglottalized SAMZ words that correspond to glottalized CLZ words have SAMZ H or R tone but not L. Only about 60% of SAMZ checked and rearticulated words have ? tone in CLZ. Regarding SAMZ as conservative, it seems that glottalization was lost in CLZ about 40% of the time, but that the CLZ glottal tone does indeed reflect historical glottalization. Another 20% of words with CLZ glottal tone seem to have developed it not from glottalization but from the high pitch which is found with high and rising tones and also from glottalization.

More than 80% of CLZ low-toned words correspond to SAMZ low-toned words. About 60% of CLZ falling-toned words have low tone in SAMZ. More than 80% of CLZ rising-toned words have rising tone in SAMZ. CLZ high-toned words are rare but never correspond to SAMZ L tone. Low and rising tones are common and are found in both languages, not having changed in 80% of the words surveyed. While just over half of CLZ falling tones seem related to the SAMZ low tone, the other 40% may have resulted from the addition of a low tone following a high tone, for example with post-tonic vowel deletion, in CLZ and Coatec. Of the 60% of CLZ falling tones that correspond to SAMZ low tone, while it is possible that there was a merger in SAMZ, it is equally possible that there was a split in CLZ. The SAMZ low tone is not level at all but falls phonetically, the same as in the most Northern dialects of CLZ. Upstep, floating tones, or other details about the tonal environment in these words may have contributed to the development of a higher falling tone that exists contrastively in CLZ today.

Looking at the same results but with the focus on SAMZ, we find that SAMZ high is most likely to correspond to CLZ high (30%) and glottal (60%) tones. SAMZ rising tone is most likely to correspond to CLZ rising tone (almost 50%) but can correspond with any CLZ tone. SAMZ low tone is most likely to correspond to CLZ falling and low tones (about 40% each).

About 25% of CLZ words in this sample have glottal tone. About 25% of SAMZ words have checked vowels. About 25% of SAMZ words have rearticulated vowels. About 50% of SAMZ words are not glottalized. So, while about half of SAMZ words have some type of glottalization, this number has been reduced by half in CLZ.

Looking at the frequency with which tones occur we find that low is the most common SAMZ tone, making up a little over 50% of this sample. High and rising tones are equally common in SAMZ. This differs from CLZ in which high tone is rare.

## 6. Conclusion

The description of SAMZ will have much to offer toward the comparison with other Zapotec languages and the documentation of the history of the Southern Zapotec region itself. This language has made important contributions to this region and is unfortunately now at the end of its days. The complexity, e.g. of tonal morphology, found in this language will surprise those accustomed to seeing more morphological simplification in dying languages. I see two possible explanations for this. It is possible that forms which were once predictable based on particular generalizations come to be used more haphazardly. However, if this were the case one would expect more free variation than I

have seen so far with the first person possessed forms I have recorded. More likely, the complexity preserved by the last two speakers of this language, survives because this language virtually ceased to be spoken with one fell blow in 1965. If it had ceased to be used in a more gradual way the expected simplifications could have spread through the dwindling speech community. Instead, the only two people on earth who can speak this language fluently acquired it as children when the language was not so endangered and for political reasons stopped speaking it to each other or passing it along to younger generations. The language is almost mummified, resisting change because of lack of use. Yet, the language is not dead yet and the younger generation today is more interested than the generation of 1965. While no children are learning this language, young adults frequently visit the speakers and ask them how to say things in Zapotec. I hope that whatever documentation I am able to produce will also be of value to future generations of people in San Agustín Mixtepec, Miahuatlán, Oaxaca, Mexico.

### Notes

<sup>1</sup> My fieldwork on this language in the summers of 2002 & 2003 was generously funded by the Endangered Language Fund, The Survey of California and Other Indigenous Languages, and the California Indian Language Center. I am also especially grateful to Lázaro Díaz Pacheco & my anonymous SAMZ consultant.

<sup>2</sup> I use the term “plosive” to include both stops and affricates. This may not be a standard usage for this term but it is the one I originally acquired in my acquisition of linguistic jargon and I find it convenient to use this definition when talking about series of Zapotec phonemes that include both stops and affricates.

<sup>3</sup> This noun is obligatorily possessed. To simply refer to any old sandal one would use the compound *yě'l yid* and by contrast that noun is alienably possessed, e.g. *xkě'lyid lè* ‘tu huarache; your sandal.’

<sup>4</sup> Most body parts are inalienably possessed but some extremities are alienably possessed, perhaps because these body parts can come off in accidents or violent incidents during one’s lifetime, while the loss of more core body parts cannot be lost before the end of one’s life. A similar example is the loanword *kól* ‘cola; tail’ in *xkól má<sup>7</sup>* ‘su cola del animal; it’s tail.’

<sup>5</sup> For discussion of class A see Kaufman (1993, 2003) or Beam de Azcona (2004). For different analyses see Bartholomew (1983) or López & Newberg (1990).

### References

- Bartholomew, Doris A. 1983. *Gramatica Zapoteca* in Nellis, Neil and Jane G. Nellis, comps. *Diccionario zapoteco de Juárez: zapoteco-español, español-zapoteco (Ca titsa' qui' ri'u)*. Vocabularios (y Diccionarios) Indígenas ``Mariano Silva y Aceves" 27. Mexico: Instituto Lingüístico de Verano.
- Beam de Azcona, Rosemary G. 2001. Proto-Zapotec coronal stops in Coatlán-Loxicha Zapotec and other Southern Zapotec languages. ms.

- , 2004. *A Coatlán-Loxicha Zapotec Grammar*. Ph.D. dissertation. University of California, Berkeley.
- Benton 1988. Proto-Zapotec Phonology. ms.
- , 2002. A reconstruction of the tone system of Proto Zapotec. ms.
- Bickmore, Lee and George Aaron Broadwell. 1998. High Tone Docking in Sierra Juárez Zapotec. *International Journal of American Linguistics*, vol. 64, no. 1, pp. 37-67.
- Black, Cheryl A. 1994. *Quiegolani Zapotec Syntax*. Ph.D. dissertation. University of California, Santa Cruz.
- Córdova, P. Fray Juan de. 1987 [1578]. *Arte del Idioma Zapoteco*. México: INAH.
- Fernández de Miranda, María Teresa. 1995 [1965]. El Protozapoteco. México: El Colegio de México y el Instituto Nacional de Antropología e Historia.
- Kaufman, Terrence S. 2003. Proto-Zapotec reconstructions. ms.
- López L., Filemón and Ronaldo Newberg Y. 1990. *La Conjugación del Verbo Zapoteco: Zapoteco de Yalálag*. México, D. F.: Instituto Lingüístico de Verano.
- Pickett, Velma B. y colaboradores. 1959. *Castellano-zapoteco, zapoteco-castellano [Vocabulario zapoteco del Istmo]. Vocabularios (y Diccionarios) Indígenas "Mariano Silva y Aceves" 3*. Mexico: Instituto Lingüístico de Verano. ii, 246 pp. [2nd ed., rev. and enlarged (1965), xii, 163 pp.; repub. (1968, 1971)].
- Pickett, Velma B., Cheryl Black, and Vicente Marcial Cerqueda. 1998. *Gramática Popular del Zapoteco del Istmo*. Juchitán: Centro de Investigación y Desarrollo Binnizá A. C. e Instituto Lingüístico de Verano A.C.
- Rojas, Basilio. 1958. Miahuatlán: Un Pueblo de México. Monografía del distrito de Miahuatlán, estado de Oaxaca.
- Sicoli, Mark. 1998. "Glottalization and Categories of Possession in Lachixío Zapotec." Paper presented at the annual meeting of the Society for the Study of the Indigenous Languages of the Americas, New York.
- Smith Stark, Thomas C. 2003. "Algunas isoglosas zapotecas." Trabajo preparado para las actas del III Coloquio Internacional de Lingüística "Mauricio Swadesh." México, D.F.: UNAM.
- Suárez, Jorge A. 1973. On proto-Zapotec phonology. *IJAL* 39: 236-49.
- Swadesh, Morris. 1947. The Phonemic Structure of Proto-Zapotec. *International Journal of American Linguistics* 13: 220-30.

## Theoretical Issues in the Representation of the Glottal Stop in Blackfoot\*

Tyler Peterson

University of British Columbia

### 1.0 Introduction

This working paper examines the phonemic and phonetic status and distribution of the glottal stop in Blackfoot.<sup>1</sup> Observe the phonemic distribution (1), four phonological operations (2), and three surface realizations (3) of the glottal stop:

- (1) a. *otsi* 'to swim'                      c. *otsiʔ* 'to splash with water'  
b. *oʔtsi* 'to take'

- (2) a. Metathesis:  $V\underline{?}VC \rightarrow VV\underline{?}C$

*nitáóʔmaiʔtakiwa*  
nit-áʔ-omaíʔtakiwa  
1-INCHOAT-believe  
'now I believe'

- b. Metathesis → Deletion:  $V\underline{?}V:C \rightarrow VV\underline{?}C \rightarrow VV:C$

*kátaookaawaatsi*  
*kátaʔ-ookaa-waatsi*  
INTERROG-sponsor.sundance-3s.NONAFFIRM  
'Did she sponsor a sundance?' (Frantz, 1997: 154)

- c. Metathesis → Degemination:  $V\underline{?}V\underline{?}C \rightarrow VV\underline{?}C \rightarrow VV\underline{?}C$

*áóʔtooyiniki*  
*áʔ-oʔtoo-yiniki*  
INCHOAT-arrive(AI)-1s/2s  
'when I/you arrive'

- d. Metathesis → Deletion → V-lengthening:  $V\underline{?}VC \rightarrow VV\underline{?}C \rightarrow VV:C$

*kátaoottakiwaatsi*  
*kátaʔ-ottaki-waatsi*  
INTERROG-bartender-3s.NONAFFIRM  
'Is he a bartender?' (Frantz, 1997)

- (3) Surface realizations:  $[V\underline{?}C] \sim [VV\underline{?}C] \sim [VV:C]$

*aikaiʔni* 'He dies'

a. [aɪkaɪʔni]

b. [aɪkaɪmi]

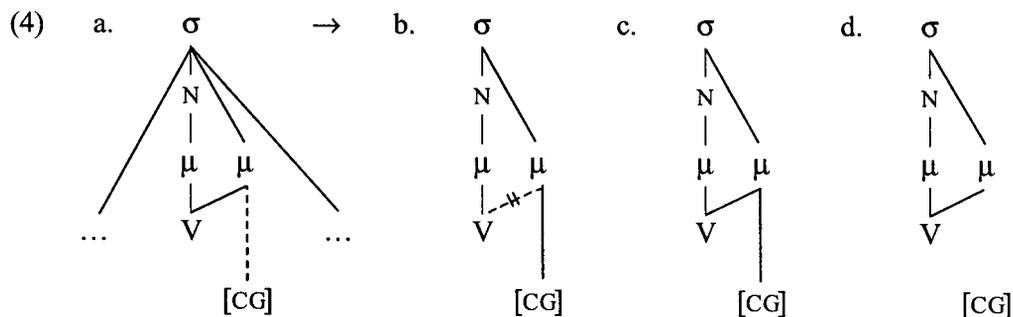
c. [aɪkajni]

The glottal stop appears to have a unique status within the Blackfoot consonant inventory. The examples in (1) (and §2.1 below) suggest that it appears as a fully contrastive phoneme in the language. However, the glottal stop is put through variety of phonological processes (metathesis, syncope and degemination) that no other consonant in the language is subject to. Also, it has a variety of surface realizations in the form of glottalization on an adjacent vowel (cf. (3)), and/or compensatory lengthening of that vowel (cf. (2)c.) – effects that no other consonant have. These observations of the distributional and surface properties of the glottal stop in Blackfoot motivate the following investigation: what exactly are the properties of the glottal stop that set it apart from other consonants in the inventory? Once those properties are identified, how can they be represented in the phonology of Blackfoot? At this point, these questions make a priori assumptions about the glottal stop as a full consonantal segment, and this paper will challenge that assumption by suggesting that the glottal stop in Blackfoot is not in fact a full phonemic segment of the inventory, but actually the phonetic realization of a glottalization on an underlyingly long vowel, the surface realizations the result of parsing and licensing strategies at work in the syllable with adjacent vowels.

The goal of this paper is to develop an adequate phonological representation that can unify the various phonological (cf. (2)) and phonetic patterns (cf. (3)) of the glottal stop in Blackfoot under one account. The distribution of the glottal stop and its surface realizations has offered a frequent challenge to the phonological accounts of various languages. This is partly due to its interaction with (or manifestation as) different phonation types of vowels (i.e. creaky voice), tone (Picanço 2002; Yip 1995), compensatory lengthening alternations (Kavitskaya 2001), it's ability to surface as either a full segment or glottalization in the same language (Zoll 1998; Brown 2004), as well as it's versatility to class with either resonants or stops (Shaw, p.c.). In addition this the glottal stop in Blackfoot is subject to restrictions that other segments are immune to, for example gemination, or the ability to function as an onset – all phenomena that must be considered in any adequate phonology of the glottal stop.

### 1.1 Proposals

The specific proposal that will be sketched out is that what surfaces both phonemically and phonetically as the glottal stop in Blackfoot is the result of the realization of a morpheme-level, floating feature [CONstricted GLOTTIS] ([CG]) feature, which is prosodically licensed by a 'non-nuclear' mora. This is schematized (4)a.:



Given certain assumptions about syllable structure and syllabification, and prosodic features such as the nucleus and mora (both will be explained in more detail below), the syllable template in (4)a. coupled with the proposal that the floating [CG] feature is prosodically licensed by this non-nuclear mora, makes at least three predictions regarding the interaction between vowel length (as represented by the mora) and the realization of this feature, all of which are borne out in Blackfoot. When [CG] is licensed by a non-nuclear mora (as opposed to nuclear one), a de-linking of the vowel features occurs and a full glottal stop surfaces (V?) as in (4)b. If there is no de-linking of the vowel, as in (4)c., laryngealization (creaky voice) on that vowel surfaces ( $V_1V_1$ ). (4)d. demonstrates that if floating [CG] is left floating (i.e. unlicensed), a bimoraic long vowel surfaces (V:). This effectively derives all of the surface realizations of the glottal stop.

One crucial prediction that follows from this proposal is that all morphemes with previously assumed V? sequences are actually underlyingly bimoraic long vowels (V:) (cf. (4)a.), and the present analysis will take quite literally the phonological representation of a long vowel as two versus one mora. Both impressionistic and preliminary phonetic analysis suggests that when the glottal stop surfaces (as a result of parsing the [CG] feature to the non-nuclear mora, cf. (4)b.) it shortens the long vowel *or* creaky voice appears on the last half of the long vowel (cf. (4)c.) The non-nuclear or 'second' mora provides a natural phonological representation for the locus of processes like 'compensatory shortening' of a vowel (or a glottal stop surfacing) and creaky voice on the second half of a long vowel. This approach, coupled with the proposed non-nuclear mora licensing condition on [CG], will interact in a principled way in capturing the metathesis and deletion (degemination) patterns. Within a moraic-prosodic framework these processes will actually follow instead from universal conditions on syllable weight restrictions and a ban on moraic onsets (Shaw p.c.) and not directly from conditions on the glottal stop itself. For example, when [CG] is licensed by the non-nuclear mora and the vowel is de-linked, the resulting full glottal stop is then treated as a legitimate consonantal onset; however, its moraic status bans it from being an onset. The result is a diphthongization of the two vowels under one nuclear mora and the non-nuclear [CG] glottal stop mora is left in a suitable coda position (cf. (4)a.); thus capturing the underlying generalization that the glottal stop can never occur as an onset (cf. (4)-0). Under this analysis, deletion is viewed as an unparsed [CG] feature – the underlying bimoraic, long vowel is simply left intact (cf. (4)d.).

The paper is organized as follows: in §2 the established distributional properties and previous approaches to its distribution of the glottal stop in Blackfoot will be re-examined. This will take the form of reviewing various data as well as incorporating data that has been collected from a speaker. In section 3, the above proposals will be discussed in greater detail and implemented.

## 2.0 Background and Observations

Consonant length in Blackfoot is contrastive. There are no voiced non-sonorant consonants underlyingly in Blackfoot, nor are there any liquids. Phonemically, there are only three vowels /i/, /o/, and /a/, with various surface (lax) realizations. Vowel length is also contrastive, however only tense vowels can be long. Vowels may also be voiceless in certain environments (i.e. word-finally). Frantz (1997) observes three diphthong in Blackfoot: [ai], [ao] and [oi].

(5) Consonant Inventory:

p	t	k	ʔ (')
	s	ts	x (h)
ps	ts	ks	
m	n		
w		j (y)	

Vowel Inventory:

	i	u	
'i'	ɪ	ʊ	
	e	ɔ	'o'
'a'	ɛ	ɔ	
	æ	a	

### 2.1 The phonemic status of the glottal stop

The glottal stop appears to have phonemic status as a full segment, as it contrasts in the following minimal pairs (All examples in this subsection are from Frantz (1995)):

(6) a.	<i>ini</i>	'to see'	a'	<i>iʔni</i>	'to die'
b.	<i>sik</i>	'black or dark'	b'	<i>siʔk</i>	'cover/blanket'
c.	<i>itsi</i>	'high quality'	c'	<i>iʔtsi</i>	'to feel emotion'
d.	<i>itsk</i>	'past; by'	d'	<i>iʔtsk</i>	'bare'
e.	<i>oto</i>	'spring'	e'	<i>oʔto</i>	'to pick up/take'
f.	<i>moksis</i>	'awl'	f'	<i>moʔksis</i>	'armpit'
g.	<i>ipot</i>	'to beat a drum'	g'	<i>ipoʔt</i>	'his eyes are closed'
h.	<i>opam</i>	'across'	h'	<i>opaʔm</i>	'lull to sleep or relaxation'
i.	<i>á</i>	DURATIVE ASP.	i'	<i>áʔ</i>	INCHOATIVE PREFIX
j.	<i>ika</i>	'foot'	j'	<i>ikaʔ</i>	'for now; presently'
k.	<i>ini</i>	'to see'	k'	<i>iʔni</i>	'to die'

In some examples it can contrast in three different positions (7), or with another stop (8):

(7) a.	<i>otsi</i>	'to swim'	c.	<i>otsiʔ</i>	'to splash with water'
b.	<i>oʔtsi</i>	'to take'			
(8) a.	<i>moʔtsís</i>	'arm/hand'	a'	<i>mottsís</i>	'gut/intestine'
b.	<i>otaʔs</i>	'on/upon'	b'	<i>otats</i>	'horse'

From this, a basic generalization emerges: in non-derived forms, the glottal stop never appears as a syllable onset.<sup>2</sup> This generalization will also hold in derived environments: where the vowel occurs at the right edge of a morpheme, upon concatenation with a

vowel initial morpheme, it will metathesize to the closest post-vocalic, pre-consonantal position; or, as will be demonstrated in §3.0, a post-nuclear, pre-coda position.

## 2.2 Phonological processes

As can be seen in the examples in (2), three phonological processes can be observed operating on the distribution of the glottal stop: metathesis, deletion (including degemination) and assimilation. Frantz (1997: 154) observed that the triggering of these three processes is motivated by a constraint banning the glottal stop from surfacing intervocalically. This distributional property can be seen in the roots in §2.1, but also applies as a condition in a derived environment. Metathesis is thus a reflex of this condition moving the glottal stop to the nearest post-vocalic, pre-consonantal position. Frantz's rule in (9) can be seen operating on (11) and (12):

- |   |  |
|---|--|
| (9) <i>Glottal Metathesis</i><br>$?V \rightarrow V? / V\_C$                                   | (10) <i>Glottal Syncope/Degemination</i><br>$? \rightarrow \emptyset / \left\{ \begin{array}{l} V_1 V_2 \_ C \\ \_ ? \end{array} \right\}$ |
| (11) <i>nitáóʔmaiʔtaki</i><br>nit-áʔ-omaíʔtaki<br>1-INCHOAT.-convinced<br>'now I'm convinced' | (12) <i>kátaookaawaatsi</i><br>kátaʔ-ookaa-waatsi<br>INTEROG-sponsor.sundance-3s.NONAFFIRM<br>'Did she sponsor a sundance?'                |

Metathesis feeds other phonological processes, such as deletion. A similar intervocalic environment surfaces in (12)-(13) and metathesis is triggered; however, once the glottal stop is moved to this position it is deleted. This is reflected in the formulation of (10): When two glottal stops are adjacent degemination occurs (presumably after the application of metathesis: *áʔ-oʔtoo-jiniki* → *áóʔʔtoojiniki* → *áóʔtoojiniki*):

- (13) *áóʔtoojiniki*  
*áʔ-oʔtoo-jiniki*  
INCHOAT-arrive-1s/2s  
'when you arrive'

Finally, metathesis and deletion occur in (15) when the glottal stop precedes a long consonant, accompanied by compensatory lengthening of the preceding vowel.

- (14) *Glottal Assimilation*  
 $V? \rightarrow V: / \_ C: \text{ (where } C \text{ [s])}$
- (15) *kátaoottakiwaatsi*  
kátaʔ-ottaki-waatsi  
INTEROG-bartender-3s.NONAFFIRM  
'Is he a bartender?'

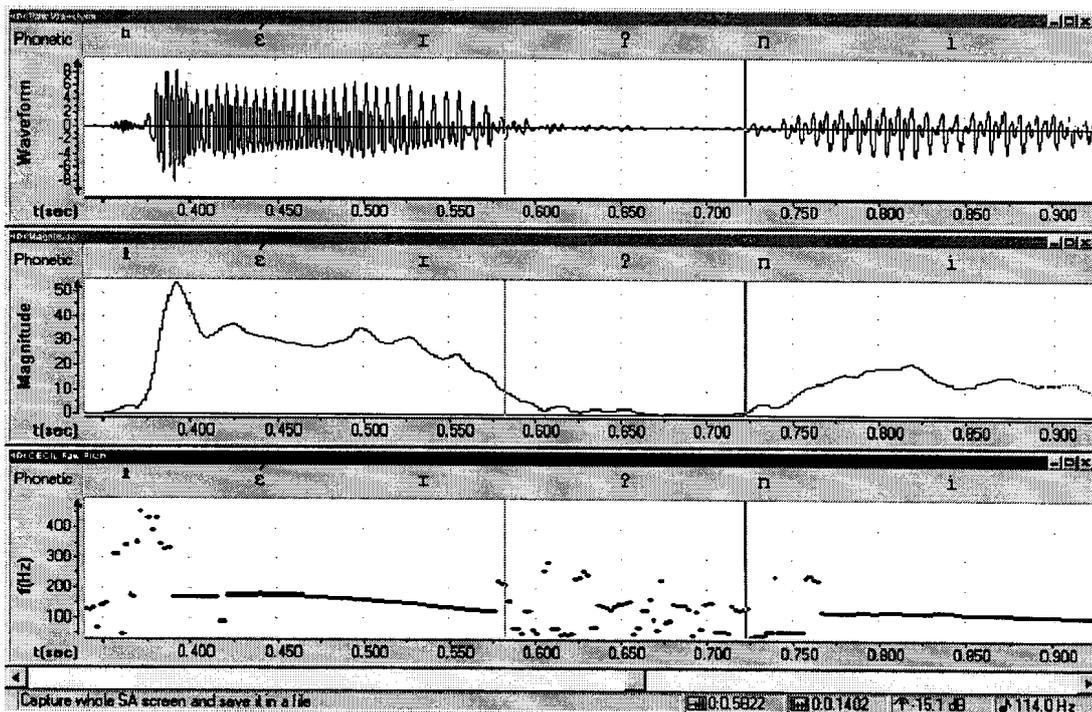
### 2.3 Speaker Data and Surface Variations: ? ~ ʔ ~ V

McLennan and Bortolin (1995) suggest a diachronic change between what speakers call ‘Old Blackfoot’ and ‘New Blackfoot’. This is characterized by several phonological changes, among which the glottal stop is gradually being replaced by creaky voice and/or long segments; or in other words, the right edge of a vowel is gradually assimilating the glottal stop, sometimes surfacing as creaky voice or lengthening. There is evidence that this is also a synchronic phenomenon, where the surface variations in (16) are in free alternation:

(16) [...VʔC...] ~ [...Vʔ̤C...] ~ [...V:C...]

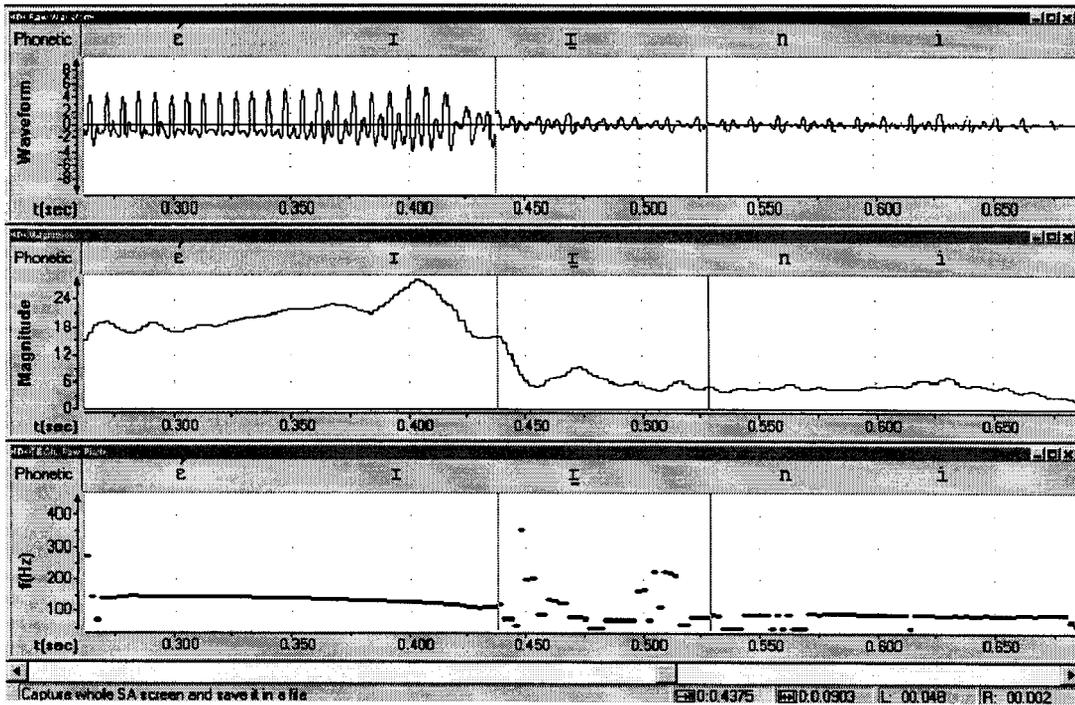
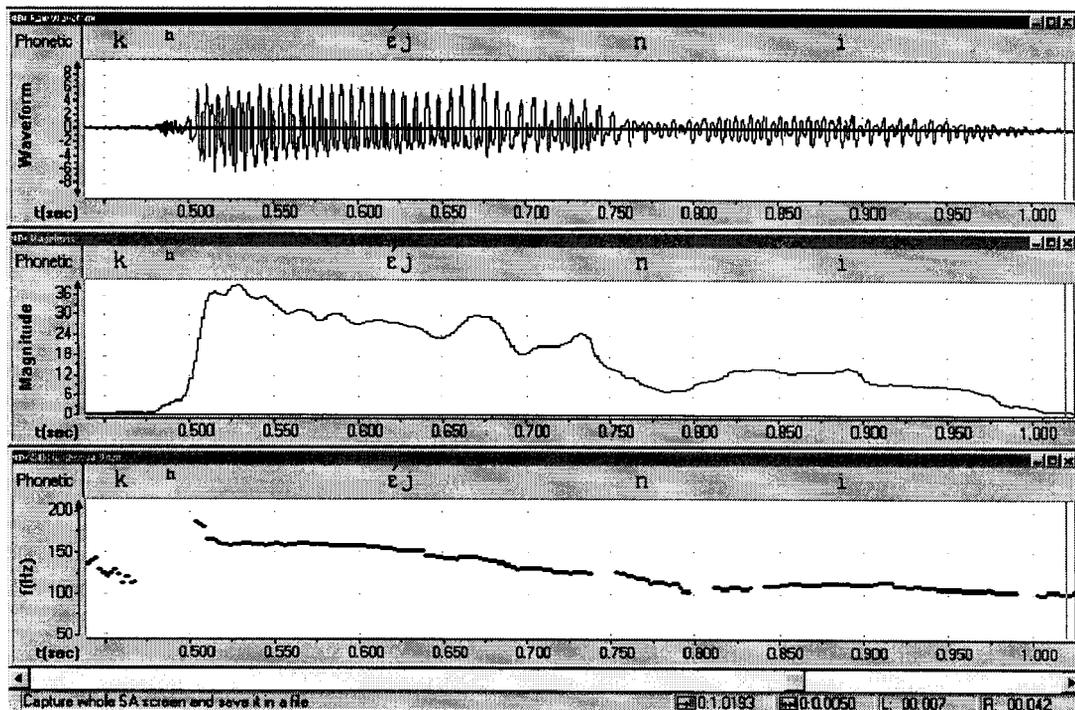
This has been observed in BPC’s speech, a speaker of New Blackfoot: in *aikaiʔni* ‘He dies’, preliminary phonetic analysis reveals that the glottal stop can surface as a full stop (17), what appears to be creaky voice (18), or almost completely reduced or assimilated to the preceding vowel (19).<sup>3</sup>

(17) [aɪkaɪʔni] *full glottal stop*



While there is clearly a post-vocalic complete stop (17) [aɪkaɪʔni], there is a continued low magnitude, periodic waveform after the vowel in [aɪkaɪʔni], suggesting possible creaky voice. In (19) there is very little reduction in magnitude and no negligible stop before the [n] segment:

(18) [aɪkaɪmi] possible creaky voice

(19) [aɪk<sup>h</sup>ejni] almost completely reduced

However, in some cases the surface realization of the glottal stop was quite stable, as can be observed in the following minimal a/b pairs of (20)-(22):

- (20) a. /ini/ ‘to see’ → a’ *nitáinip* [nitéinip] ‘I see it’  
 b. /iʔni/ ‘to die’ → b’ *aikaiʔni* [aikéiʔni] ‘He dies’
- (21) a. /sik/ ‘black or dark’ → a’ *siksináttsi* [siksínáttsi] ‘it’s black’  
 b. /siʔk/ ‘cover/blanket’ → b’ *siʔkaan(i)* [siʔká:n(i)] ‘blanket’
- (22) a. /otsi/ ‘to swim’ → a’ *tóótsit* [dú:tsit] ‘swim!’  
 b. /oʔtsi/ ‘to take’ → b’ *oʔtsit* [uʔtsit] ‘take it!’

In longer phrases, the same variability can be found; and in at least one case (23), the glottal was deleted without lengthening:

- |  |  |
|--|--|
| <p>(23) [kitayoʔkaawatsiksa]<br/> <i>kát<u>a</u>ʔjájóʔkááwaiksaawa</i><br/> <i>kátaʔ-já-joʔkaa-waiksaawa</i><br/>       INTERROG-DUR-sleep-3p:NONAFFIRM<br/>       ‘are they sleeping?’</p>            | <p>(24) [kitaiʔnóókawatsiksa]<br/> <i>kikát<u>ai</u>ʔnóókwaiksaawa</i><br/> <i>k-Ikátaʔ-ino-ok-waiksaawa</i><br/>       2-INTER-see(TA)-INV-3p:NONAFF<br/>       ‘did they see you?’</p> |
| <p>(25) [kitaooykskaʔsiwaatsiksa]<br/> <i>kát<u>ao</u>ʔkskaʔsiwaatsiki</i><br/> <i>kátaʔ-okskaʔsi-waatsiki</i><br/>       INTERROG-run-3s:NONAFFIRM<br/>       ‘did he run?’</p>                       | <p>(26) [kitoʌʔtaki]<br/> <i>kitaoʔtaki</i><br/> <i>kitaʔ-otaki</i><br/>       ‘did he take any?’</p>  |
| <p>(27) [kitaoʔkaawaatsi]<br/> <i>kát<u>ao</u>okaawaatsi</i><br/> <i>kátaʔ-ookaa-waatsi</i><br/>       INTERROG-sponsor.sundance-<br/>       3s:NONAFFIRM<br/>       ‘did she sponsor a sundance?’</p> | <p>(28) [nitóʌʔmaiʔtakʲi]<br/> <i>nit<u>á</u>ʌʔmaiʔtaki</i><br/> <i>nit-áʌʔ-omaiʔtaki</i><br/>       1-INCHOAT-convinced<br/>       ‘now I’m convinced’</p>                              |
| <p>(29) [eoʔtú:niki]<br/> <i><u>á</u>ʌʔtooyiniki</i><br/> <i>áʌʔ-oʌʔtoo-yiniki</i><br/>       INCHOAT-arrive(AI)-1s/2s<br/>       ‘when I/you arrive’</p>  | <p>(30) [áʌʔmaiʔtaki]<br/> <i><u>á</u>ʌʔmaiʔtaki(wa)</i><br/> <i>áʌʔ-omaiʔtaki(wa)</i><br/>       INCHOAT-convinced<br/>       ‘now he believes’</p>                                     |

All of the observed  $\text{ʔ} \sim \text{ʌ} \sim \text{V} (\sim \emptyset)$  variations in BPC’s speech can be accounted for by considering the glottal stop not to be a actual segment in the language, but rather the realization of a laryngeal feature of vowels. The (in)stability of this feature is a reflection of parsing strategies within the prosodic domain – a hypothesis that will be explored in further detail below.

### 3.0 Phonological Representations and the Glottal Stop

Although the above re-write style rules can adequately account for the distribution of the glottal stop in derived environments (as well as reflecting its post-vocalic, pre-consonantal distribution in roots), the phenomena characterized by these rules seem to suggest something more is at work here. The central question is exactly why are these restrictions placed on the glottal stop's distribution? This 'conspiracy' effect can be perhaps characterized as an attempt to converge on a well-formed output, and notions like 'coda-like' position point to an account that includes the syllable as the natural and relevant domain for these alternations. Taking this as a direction, in producing an account for the Blackfoot syllable it is possible to first begin with a set of assumptions that can be inferred from Frantz's rules. This will provide the observational groundwork and insight necessary for capturing the essence of glottal stop distribution within a prosodic framework. Informally, these rules can be interpreted in the following way:

(31) Glottal Metathesis: The glottal stop cannot occur as an onset.

Once metathesis occurs, it feeds further processes. The derivation of [kátaʔ-o:ka:wa:tsi] could yield: [káta.ʔo:ka:wa:tsi]. Metathesis applies as the constraint on a glottal stop onset surfacing: \*[ká.tao:ʔ.ka:wa:tsi]. Although the glottal stop has been moved to an adequate post-vocalic position, it is still deleted: *kátaookaawaatsi*. The glottal stop is always deleted after a surface sequence of three vowels, and this suggests a restriction on syllable weight: the second syllable in [ká.tao:ʔ.kaa.waa.tsi] would be 'too heavy' and the glottal stop is marked for deletion in order to respect syllable weight. When gemination occurs (as fed by metathesis) in [áoʔʔ.to:yi.ni.ki] → *áoʔtooyiniki*, the glottal stop is deleted. Assuming this geminate occurs within the same syllable, this deletion rule would reflect the universal constraint on tautosyllabic geminates:

(32) Glottal Syncope/De-Gemination: Syllable weight must be respected; No tautosyllabic geminates.

So far, what this line of inquiry actually suggests is that there is an alternation between a full vowel and the glottal stop somewhere in the nucleus of a syllable. However, treating the glottal stop as just another consonant in the inventory implies that it is available for syllabification not only as a moraic coda (hence compensatory lengthening alternations), but that it can also function as an onset. A somewhat simplistic but effective suggestion would be to class the glottal stop with other segments that cannot function as onsets: vowels. Taking this at face value, the glottal stop in this language can be viewed a form of glottalized vowel, or in other words, the glottal stop itself is actually a vowel with some added laryngeal feature. This compromises the glottal stop's position in the segment inventory of Blackfoot: instead of being a true member of the inventory, the glottal stop is actually the result of a floating [CG] feature realized at the edge of a nucleus. How this feature is aligned with the nucleus edge will be addressed in §3.2.

### 3.1 Morpheme-level Subsegmental Features: [CONSTRUCTED GLOTTIS]<sup>4</sup>

Zoll (1998) proposes a single underlying distinction between full segments and subsegmental elements, or *floating* features (7). A floating laryngeal feature such as [CG] can move around to find a place to dock and attach to existing vowel's root node. An alternative is that the [CG] feature can be directly licensed by a prosodic node, without an intermediary root node. Brown (2004) and Gittlen and Marlett (1989; cited in Macaulay and Salmons 1995) have proposed that glottalization, is a feature of syllables in Mixtec, and this can extend in principle to glottalization in Blackfoot.

Blackfoot has contrastive length, therefore length is encoded underlyingly. The hypothesis that will be forwarded is that [CG] is licensed by the second mora of an underlyingly bimoraic vowel – the properties of which are discussed in the following section.

- (33) a. /á/ DURATIVE (cf. *áyoʔkaawa* 'he is sleeping')  
 b. /áʔ/ INCHOATIVE (cf. *áʔyoʔkaawa* 'now he is sleeping')

- (34) a.  $\left[ \begin{array}{c} \mu \\ | \\ \acute{a} \end{array} \right]$  DURATIVE ASPECT      b.  $\left[ \begin{array}{c} \mu \quad \mu \\ \diagdown \quad / \\ \acute{a} \\ [CG] \end{array} \right]$  INCHOATIVE PREFIX

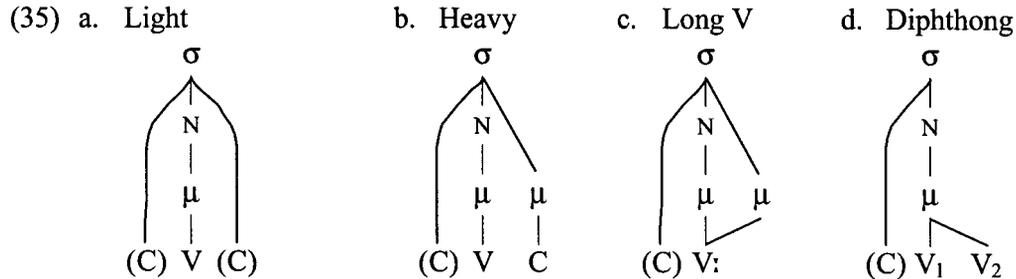
The morpheme level specification of the floating [CG] feature can be highlighted in the minimal pairs DURATIVE/INCHOATIVE minimal pairs, now represented in the lexicon as [á]/[áʔ]<sub>[CG]</sub>. There are advantages and disadvantages to root node vs. prosodic licensing, and it is beyond the scope of this paper to fully compare in detail this distinction (see Zoll 1998 for detailed discussion), but by linking the [CG] feature directly to the prosody we gain the power of predictability in the surface distribution of the creaky voice and the glottal stop. If [CG] could only licensed by a (vocalic) root node, we would need to appeal to alignment or other measures in order to derive the surface position of glottalization.

Given the assumption that prosody licenses glottalization, so far there is no way to distinguish between the two moras in a long vowel, as it would seem the first mora could also license the [CG] feature. From here, we are now ready take this preliminary theory of the floating feature [CG] as a lexically-specified, morpheme-level feature, prosodically licensed by a mora, and observe how it interacts with the next level: the syllable. In the next section, I will elaborate on this inherent properties and asymmetries of the first and second mora in a long vowel syllable and how they determine the distribution of glottalization.

### 3.2 A Syllable Template: the Nuclear Moraic Model

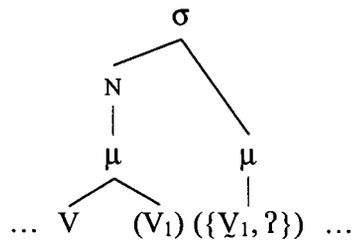
Syllable weight is standardly defined in terms of the mora (Hyman 1985, Hayes 1989, Zec 1988, etc.). This provides a fundamental dichotomy of syllable types into light monomoraic vs. heavy bimoraic syllables. Although this claim has empirical support

from the behaviour of certain stress systems, compensatory lengthening processes, contrastive length etc., it is proposed that there is also a distinction between the two moraic constituents of a bimoraic syllable: the second mora of a bimoraic syllable is characterized as actually ‘weak’ in relation to the first mora. This asymmetry is captured by assigning a ‘Nucleus’ status to the first mora – an intervening head between the mora and the syllable node – vs. a non-nuclear mora which is headed by the syllable node directly (Shaw, 1992):



(35)a. represents a light, monomoraic syllable, while (35)b. and d. represent a closed, bimoraic syllable and one with a long vowel respectively. (35)d. is again a light syllable diphthong. Evidence for these representations (an adaptation of Shaw, 1992) follow from a closer examination of the function of the second mora and how it is actually distinct from the first. When viewed in a certain way, a bimoraic syllable can be seen as a temporal metric that measures across the most sonorous part of a syllable. A nuclear mora will always head a vocalic (or at least sonorant) segment, while the second mora typically a less sonorant segment such as a consonant, representing a decreasing sonority sequencing temporally across the two moras. The second mora can encode vowel length (underlyingly), but it is dependent on the root node features of the vowel that makes up the actual nucleus (‘N’), the only obligatory feature of a syllable. What is relevant for the present analysis is the fact that the second mora can effectively represent the locus of the ‘weakening’ of the right edge through creaky voice or shortening of the underlyingly long vowel. In addition to this, the first two vowels of a three-vowel nucleus in Blackfoot are consistently a diphthong, and that a maximal sequence of three vowels is always  $[V_1 V_2]$  ( $*V_1 V_2 V_3$ ;  $*V_1 V_2$ ), and glottalization is always post-vocalic, or in this approach, post-nuclear. This yields the following nucleus template for Blackfoot (Peterson, to appear; Shaw 1992, 1993; and see Yip ??? for a similar treatment in Choayang):

(36) Blackfoot Syllable Template (Nuclear Moraic Model)



Again, this is drawn from evidence that, in syllables where a glottal stop follows a nucleus, that nuclear sequence is always a diphthong:

$$(37) \quad \begin{array}{c} \mu^N \mu^\sigma \\ \wedge \quad | \\ k i t a o \text{?} t a k i \end{array}$$

Through eliciting other diphthong-long vowel examples of this type, a longer sequence was expected: *kitao<sub>o</sub>kaawaatsi*, but BPC produced several tokens of *kitao<sub>o</sub>?kaawaatsi*. (27) also displays a similar free variation between a glottal stop and lengthened vowel: *kitao<sub>o</sub>kaawaatsi* ~ *kitao<sub>o</sub>?kaawaatsi*. This alternation is also apparent in words with plain long vowels, as in *tápisamikooko* ‘late night’. BPC would produce both V: and V? upon repeated elicitation: [...kooko] ~ [...ko?ko]. It is difficult to tell from just looking at the dictionary if diphthongs inherently a part of this, but BPC’s speech so far has always confirmed it. The glottal stop seems to never appear in any position other than directly following a nucleus, and almost all of the forms produced by BPC contain a diphthong or short vowel followed by a glottal stop. In these cases, once [CG] is licensed, by the ‘weak’ mora, the delinking of the vowel to the weak mora can be said to be optional in BPC’s speech.

The non-nuclear is the locus of glottalization and therefore licenses the [CG] feature. Equipped with this template, and our assumptions regarding the licensing of [CG], we can now observe how these behave in a derived environment, including the various parsing strategies highlighted in (4). Consider the metathesis case:

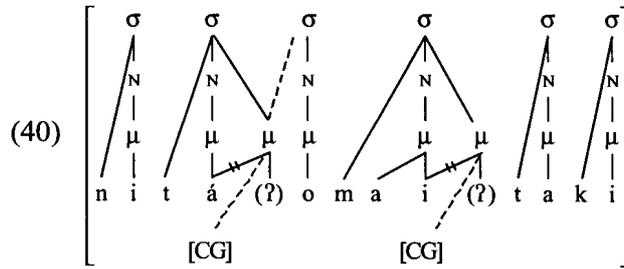
$$(38) \text{ Metathesis: } V?VC \rightarrow VV?C$$

*nitá<sub>o</sub>?mai?takiwa*  
 [nit]-[á:]<sub>[CG]</sub>-[omaɪtakiwa]      (compare UR with (2)a.)  
 I-INCHOAT-believe  
 ‘now I believe’

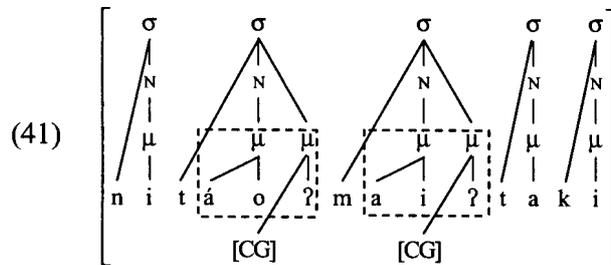
With the revised URs discussed in the previous subsection, all V? sequences are underlyingly V: with lexically specified floating [CG] feature:

$$(39) \quad \left[ \begin{array}{c} \mu \\ | \\ n \quad i \quad t \end{array} \right] \left[ \begin{array}{c} \mu \quad \mu \\ \diagdown \quad / \\ \acute{a} \end{array} \right] \left[ \begin{array}{c} \mu \quad \mu \quad \mu \quad \mu \\ | \quad \diagdown \quad / \quad | \quad | \\ o \quad m \quad a \quad i \quad t \quad a \quad k \quad i \\ [CG] \end{array} \right]$$

The floating [CG] is parsed to the non-nuclear mora triggering the delinking of the vocalic nucleus [a]. The outcome in (40) is the syllabified form \*/ni.tá.?<sub>o</sub>.mai?.ta.ki/, but this is ungrammatical because the moraic glottal stop, surfacing as a consonant, is syllabified as an onset:



It is the assumption that the glottal stop becomes a consonant and then is forced as an onset that invokes the metathesis rule:



The glottal stop is now in a position where it can't be re-syllabified as an onset, but at the expense of deleting a mora. Observing another (universal) constraint on the bi-moraic limit of syllables, this metathesis forces the re-syllabification to correct a potential trimoraic syllable by diphthongization of [...á<sub>μ</sub>o<sub>μ</sub>...] into [...áo<sub>μ</sub>...], where the two segments share a mora. This is confirmed by the template and accent spread (Peterson, to appear). This derivation deals with the metathesis and weight-related deletion. The glottal stop, surfacing as a full, mora-bearing segment cannot be re-syllabified as an onset (Shaw p.c.). It escapes this by moving rightwards to a suitable coda position in the next syllable. Given space considerations, I will defer a detailed analysis of examples of metathesis feeding deletion/degemination (i.e. (2)b. - d.): [káta?-ookaawaatsi] → [kátao?kaawaatsi] → *kátaookaawaatsi*; [á?-o?too-yiniki] → [áo??tooyiniki] → *áo?tooyiniki*; instead suggesting that the above analysis in principle can be extended to these cases as well. A potentially problematic case is the compensatory lengthening case of (2)d.: [káta?-ottakiwaatsi] → [kátao?ttakiwaatsi] → *kátaoottakiwaatsi*. Underlyingly, [káta?] is now viewed as [kátaa], yielding a possible output \**kátaoottakiwaatsi*. This actually follows from the syllable template in (36): the glottal stop would surface as the second [a] and then go through the diphthongization process and ending up as a glottal stop on the non-nuclear [o] mora. This is predicted but does not actually surface \**kátao?ttakiwaatsi*. One could suggest that [CG] is simply left unparsed, leaving the template to rearrange the [...aao...] sequence into the grammatical [...aoo...], but the reason for this is unclear.

#### 4.0 Summary

Synchronically, glottal approximants which result in compensatory lengthening are predicted to be moraic, while glottal stops whose deletion does not trigger compensatory lengthening are predicted to be weightless (Hayes 1989). Kavitskaya (2001) argues that the synchronic variation of this type mirrors the diachronic fate of glottal stops. They fluctuate between glottal approximants and vowels: when they surface as vowels, compensatory lengthening results in monophthongization. The present analysis offered an alternative approach, where the surface realization of either creaky voice or a full glottal stop was the result of the prosodic licensing of a [CG] feature by the non-nuclear mora of a bimoraic vowel. This offered a principled approach to not only predicting the position and distribution of glottalization (i.e. the right edge of a long vowel, or the non-nuclear mora), but also the surface phonetic realization of glottalization (i.e. the parsing strategies outlined in (4)). However, there are several outstanding issues: the two most pressing are the impact of reinterpreting all Vʔ sequences as V: on the lexicon. For example, are there any other contrasts: [á]/[á]<sub>[CG]</sub> or [áá]/[áá]<sub>[CG]</sub>? Secondly, there is typically only one [CG] per root, but potentially more than one bimoraic syllable to license it. How does floating [CG] discern which to parse to? There were also many other facts that were ignored, such as why you get an unparsed [CG] before a geminate consonant (cf. (2)d.). These are questions I will defer to further research. The outcome of this analysis is a way to better capture the distribution and surface realization of glottalization in Blackfoot, through deriving its properties in the prosody and not through sets of (possibly extrinsically) ordered sets of rules that obscure the true function of glottalization.

---

#### Notes

\*Thanks to Pat Shaw, Gessiane Picanço and J.C. Brown for their helpful input. I would also like to thank my Blackfoot consultants Broderick Prairie Chicken, and Beatrice and Melinda Bullshields (examples not cited were elicited by the author). This research was made possible with the financial support of a research grant (SSHRC 410-200-123) awarded to R-M. Déchaine, P.I. and D. Pulleyblank, Co-Inv. All errors are my own.

<sup>1</sup> Blackfoot is an Algonquian language spoken by about 5000 people of the Blood, Peigan, an Siksika tribes in southern Alberta and Northwestern Montana. Its closest sister within the Algonquian family is Cree. In Blackfoot orthography the glottal stop is represented by an apostrophe ('), but for expository purposes in this paper I will use 'ʔ'.

<sup>2</sup> However, the glottal stop can be found intervocalically between two identical vowels. This suggests a purely phonetic realization, where it occurs as the onset to the second member of an 'echo' vowel:

i. <i>saʔai</i>	'duck'	iv. <i>stáʔao</i>	'ghost/spirit'
ii. <i>háʔayaa</i>	'oh oh!'	v. <i>akáistaʔao</i>	Blood clan name
iii. <i>saoʔohkaʔpii</i>	'boring'	vi. <i>iiʔittaki</i>	'to skin'

---

Given this, the general observation above still holds: in non-derived environments, the glottal stop cannot appear as an onset to a syllable unless it surfaces phonetically as part of an echo vowel complex: (V<sub>1</sub>?V<sub>1</sub>).

<sup>3</sup> It should be noted that this is only a preliminary phonetic analysis for expository purposes, and should be followed up by a more detailed and rigorous examination.

<sup>4</sup> After a survey of dictionary entries, it appears that Blackfoot allows only one glottal stop per morpheme.

### References

- Brown, J.C. 2004. "Licensing and Glottalization in Mixtec" in *Papers from the 39th Meeting of the Chicago Linguistics Society*. Chicago: University of Chicago. 319-326
- Frantz, D.G. 1997. *Blackfoot Grammar*. Toronto: University of Toronto Press
- Frantz, D.G. 1995. *Blackfoot Dictionary of Roots, Stems and Affixes*. Toronto: University of Toronto Press
- Hayes, B. 1989. "Compensatory lengthening in moraic phonology". *Linguistic Inquiry* 20, 253-306.
- Hyman, L. 1985. *A theory of phonological weight*. Dordrecht: Foris.
- Kavitskaya, D. 2001. Compensatory lengthening: phonetics phonology, diachrony. Doctoral dissertation, UC Berkeley.
- McLennan and Bortolin 1995. "Blackfoot" Ms. University of Calgary
- Macaulay, M. and J. Salmons. 1995. "The phonology of glottalization in Mixtec". *International Journal of American Linguistics* 61:38-61
- Peterson, T. to appear. "A Template for the Blackfoot Nucleus: Evidence from Pitch Accent and the Distribution of the Glottal Stop" *University of British Columbia Working Papers in Linguistics*, Vol.?
- Picanço, G. 2002. "O sistema tonal de Mundurukú revisitado" [The tonal system of Mundurukú revisited] Proceedings of *O I Encontro Internacional do Grupo de Trabalho sobre linguas Indigenas da ANPOLL: Fonologia, Gramática e História, Vol. 1*. A.S.A. Cabral & A. Rodrigues (eds). Editoria Universitária UFPA: Belém, PA; pp. 243-253
- Shaw, P. 1992. "Templatic Evidence for the Syllable Nucleus" In the *Proceedings of the 23rd annual meeting of the North East Linguistic Society (NELS 23)*. Amherst, Massachusetts: GSLA. 463-477
- Shaw, P. 1993. "The Prosodic Constituency of Minor Syllables" *WCCFL 1993*
- Yip, M. 1997. "Morpheme-level Features: Chaoyang Syllable Structure and Nasalization" in *Proceedings of the 6th North American Conference on Chinese Linguistics: Volume 2*
- Yip, M. 1995. "Tone in East Asian Languages" In *Handbook of Phonological Theory*. J. Goldsmith, (ed), Basil Blackwell, Oxford. pp 476-494.

