

## FOREWARD

It is with great pleasure that we present the proceedings of the eleventh Workshop on American Indigenous Languages (WAIL 2008). In continuing a tradition begun with the student discussion group on North American Indigenous Languages (NAIL), the evolving membership wishes to deeply thank 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 and contributions concerning the indigenous languages of the Americas.

Andrea Berez  
Salomé Gutiérrez  
Lea Harper  
Timothy Henry  
Dan Hintz  
Dianne Hintz  
Joye Kiester

Carrie Meeker  
Verónica Muñoz-Ledo  
Carlos M. Nash  
Daisy Rosenblum  
Rebekka Siemens  
Elizabeth Shipley  
Alexander N. Walker

## Irrealis in Blackfoot?

Leora Bar-el and Ryan Denzer-King

University of Montana

**1. Introduction**<sup>1</sup>. The category “irrealis” has been a useful label for numerous languages that show a grammatical contrast between “real” and “unreal” situations. However, the category has been “inconsistently defined” across languages (Kinkade 1998: 234) and the misalignment between how it has been defined and the distribution of irrealis-marked forms across languages has also led to the claim that the realis/irrealis distinction is not cross-linguistically valid (Bybee et al. 1994:237-8). Kinkade (1998) suggests that “it is necessary to distinguish between that which is actually unreal and an irrealis grammatical category”. Thus we might expect all languages to exhibit constructions that describe “logically unreal” contexts, but as Kinkade states “none of these notions must be marked *grammatically* (i.e., morphologically or syntactically) as unreal” (p. 234). The purpose of this paper is to explore whether Blackfoot, an Algonquian language spoken in Montana and Alberta, provides evidence for irrealis as a grammatical category in the language<sup>2</sup>.

Tense, aspect, and even more so, mood, are understudied phenomena in Blackfoot<sup>3</sup>. The language has been described as having an “irrealis” mood; Uhlenbeck (1938) states that one of the “repressional” moods marked by the suffix *-opi* (and its variants) is used to express “a supposition, nearly always an unreal one, and may therefore be called ‘irrealis’” (p. 169). More recently, this suffix has been described as a marker of the “unreal” paradigm, used in “counterfactual and hypothetical subordinate clauses” (Frantz 1991:115), more specifically “in the apodosis of conditional sentences”, expressing “the action or state which would result if the contrary-to-fact statement of the protasis should happen” (Taylor 1969:170). At first, then, it appears that Blackfoot does show evidence for the category irrealis, not only because it has morphology that has been labeled as such, but because it appears in counterfactual contexts (1), which are among the most typical irrealis contexts:

- (1) Nitsítssáyoyi**htopi**, nitáaksoyi ánnohka  
nit-it-say-loyi-**htopi** nit-áak-loyi annohka  
I-then-neg-eat-**unreal** 1-fut-eat now  
'If I hadn't eaten then, I'd eat now' (Frantz 1991:115, ex. x)

However, *-opi* does not appear in all contexts that are associated with “unreality”. Conditionals, for example, are marked by subjunctive morphology (2), not by the “unreal” paradigm:

- (2) Ikkamáyo'kainoainiki, nitáakahkayi.  
 ikkam-á-yo'kaa-**inoainiki** nit-yáak-wa:hkayi  
 if-dur-sleep(AI)-**2p(subjunctive)** 1-fut-go^home  
 'If you2p are sleeping, I'll go home' (Frantz 1991:113, ex. l)

Yes/no questions are marked by nonaffirmative suffixes (3), and not “unreal” morphology:

- (3) Kitsikákomimmoki**hpa**?  
 kit-íkakomimm-o:k-i-**hpa**  
 2-love-inv-1-**nonaffirm**  
 'Do you2s love me?' (Frantz 1991:133, ex. d)

We propose that irrealis is not a relevant grammatical category in Blackfoot because a variety of “logically unreal” contexts are encoded by different morphology. While we may not expect that in a given language irrealis morphology would mark *all* constructions describing logically unreal contexts, we might expect that if irrealis were a relevant category in Blackfoot that it would at least be used in more than one unreal context.

The paper is organized as follows: we first present a brief overview of the category irrealis across languages, viz., how it is described and the instability of the category (§2). We then look at an analysis of irrealis in another Algonquian language: the Moose Cree “preterit” morphemes *-pan* and *-htay* (James 1991), which are most likely historically related to Blackfoot *-opi* (§3). We turn next to our argument that irrealis is not a relevant grammatical category in Blackfoot, by showing that various “logically unreal” contexts in Blackfoot have completely different morphology, as well as commenting on the possible origins of the Blackfoot “unreal” (Frantz 1991) morpheme *-(ht)opi* (§4). We conclude with a brief discussion of the implications of our proposals and a summary of issues for further research (§5).

**2. The Status of Irrealis as a Grammatical Category.** In this section we give an overview of the motivation for positing an irrealis category in some languages. We also address why this does not necessarily lend credence to the idea of irrealis as a universal grammatical category in all languages.

**2.1. Irrealis Cross-Linguistically.** Chung and Timberlake (1985) define the realis/irrealis split as a distinction between actual and non-actual events. Realis morphology attaches to clauses, verbs, or arguments that refer to some aspect of the world as it is (utterances that make some statement about the state of the world, whether they have a positive or negative logical truth value). Irrealis morphology attaches to clauses, verbs, or arguments that refer to a world other than the one that exists at the time

of the utterance (including statements which refer to the future world, a conditional world, or a counterfactual world that is in some way the opposite of reality at the time of utterance<sup>4</sup>). This is echoed by Mithun (1999) who further suggests that actual events are those which have occurred or are actually occurring, and which are “knowable through direct perception”, while “irrealis portrays situations as purely within the realm of thought, knowable only through imagination” (p. 173). Cross-linguistically, irrealis is used in a variety of “unreal” contexts, such as conditionals, counterfactuals, imperatives, futures, questions, negatives, obligations, potentials, warnings, etc. Although the origin of the term “irrealis” to describe these contexts is not known (Bybee et al. 1994<sup>5</sup>), its use arises as a result of the observation that different constructions are marked in the same way and that the shared characteristic among the constructions is “unreality” or “nonactuality”. Caddo (Caddoan; Oklahoma), for example, encodes the realis/irrealis distinction in pronominal prefixes in the verbal domain (Chafe 1995:354, as cited in Mithun 1999:178-9):

(4) Negation*kúyt’áybah*

kúy-t’a-yibahw

negative-1agent.irrealis-see

‘I don’t see him’

(5) Yes/no questions*sàyybáwnah*

sah?-yibahw-nah

2agent.irrealis-see-perfect

‘Have you (irrealis) seen him?’

(6) Conditionals*hít’áybah*

hí-t’á-yibahw

conditional-1agent.irrealis-see

‘if I see it’

The Caddo data are meant to show that irrealis is a relevant category in this language given that negation, yes/no questions and conditionals are all marked by the same pronominal prefixes. Thus a variety of logically unreal contexts are morphologically marked in the same way in this language.

**2.2. Problems with Irrealis.** Bybee et al. (1994) suggest that “realis/irrealis is rarely realized in a language as a binary morphological distinction. It appears to be more common to have multiple markers in both domains” (pp. 237-8). The dichotomy between

logically and grammatically unreal, or “that which is actually unreal and an irrealis grammatical category” (Kinkade 1998: 234), is a necessary one because there are few if any languages that encode every possible irrealis context with a unique grammatical morpheme signalling the irrealis mood. However, generalizations about irrealis *can* be drawn: Mithun (1999) suggests that “some constructions, such as conditionals and counterfactuals, are classified as irrealis in all systems” (p. 179)<sup>6</sup>. Thus while languages differ according to which logically unreal contexts are marked with irrealis morphology, we would expect from Mithun’s claim that any language in which a realis/irrealis distinction is useful would mark conditionals and counterfactuals in the same way.

Irrealis has been proposed to be relevant in at least one other Algonquian language, Moose Cree. We now turn to an examination of this paradigm in the following section.

**3. Irrealis in Moose Cree.** James (1991) claims that in Moose Cree, the morphemes *-pan* and *-htay* have come to form a single “preterit paradigm” which has two usages: an imperfective past aspectual use and an irrealis modal usage which indicates “that the proposition is unreal or hypothetical as opposed to real and factual” (p. 285). The data suggest that this paradigm is used in both present counterfactual (7) and past counterfactual (8) contexts:

- (7) kiša:spin      iskwe:wit,      ta-miloma:kosi:**pan**  
 if                    he-be-woman    he-will-be-good-looking-**PRET**  
 ‘If he were a woman, he would be good-looking’  
 (James 1991: 286, ex. 7)

- (8) kiša:spin      ki:-wa:pama:kopane:,  
 if                    ki:-he-see-her-DUB-**PRET**  
 ka-ki:-wi:htama:ko**hta**:naw.  
 ki:-he-will-tell-us-about-it-**PRET**  
 ‘If he had seen her, he would have told us about it’  
 (James 1991: 287, ex. 8)

The preterit paradigm is also used in future conditional contexts (9). As James states, it can be used to describe “something which might take place or be the case in the future, where this is contingent upon some other event taking place” (p. 287):

- (9) kiša:spin    itohte:yin    mo:sonihk,    ka-milowe:lihte:**htay**      anta  
 if            you-go      to-Moosonee    you-will-like-it-**PRET**      there  
 ‘If you went to Moosonee, you would like it there’  
 (James 1991: 287, ex. 9)

In each of these cases, the so-called preterit marker is used in conjunction with future time reference morphology glossed as ‘will’. The data in (7) shows that *-pan* need not contribute any sense of past tense; in this case it contributes the meaning associated with irrealis counterfactual mood. The past counterfactual meaning of (8) is contributed by *-pan* as well as the preverb *ki:-*. The future conditional sense of (9) may be due to the status of the verb as eventive rather than stative (as in (7)); this, however, would require a closer look at predicate classes in Moose Cree and thus is left for future research.

James’s proposal is that the Moose Cree preterit marker *-pan* developed from Proto-Algonquian (henceforth PA) *\*(e)pan* which “originally indicated past events specifically not relevant to the present” (p. 290). Some Algonquian languages retained this original meaning; however, many developed other usages (e.g., past, irrealis, evidentiality, etc.). Some modern reflexes of PA *\*(e)pan* are shown in Table 1.

**Table 1. Reflexes of PA *\*(e)pan***

Language	Morpheme	Usage	Source
Ojibwe	<i>-bany</i>	preterit	Rogers (1978:168)
Moose Cree	<i>-pan</i>	preterit, counterfactual	James (1991)
Montagnais	<i>-shapan</i> <sup>7</sup>	past, indirect evidential	James et al. (2001)
Naskapi	<i>-shapan</i>	past, indirect evidential	James et al. (2001)
East Cree	<i>-shapan</i>	past, indirect evidential	James et al. (2001)
Plains Cree	<i>-pan</i>	preterit	Wolfart (1973:44)

Several Algonquian languages have retained the past tense/aspectual usage that has been proposed for the PA morpheme, but Moose Cree has also developed an irrealis counterfactual usage, and in Montagnais, Naskapi, and East Cree the morpheme is primarily now used as an evidential indicating indirect evidence (James et al. 2001).

James (2001), describing the *-shapan* morpheme in Cree/Montagnais/Naskapi, states that speakers can use the suffix “to indicate that they are distanced in some way from the event of state of affairs described” (p. 240). This distancing effect may be what prompted the development of the irrealis usages in many Algonquian languages. This possible relation between past and irrealis as one of distance or irrelevance to present events has been described in Steele (1975) for Proto-Uto-Aztecan; she shows that the irrealis and past tense morphemes in Proto-Uto-Aztecan apparently came from a single irrealis/preterit morpheme. James (1991) echoes this as an avenue of explanation for Proto-Algonquian. Indeed, the connection between past and irrealis is not an uncommon one cross-linguistically: in Itzaj Maya, “the modal-semantic category ‘irrealis’ interacts with the aspectual distinction, ‘perfect,’ and with tense” (Hofling 1998:214). The fact that so many Algonquian languages have developed some kind of irrealis usage from what is

typically understood to be a preterit morpheme in PA leaves open the question as to what the actual meaning of PA *\*(e)pan* was.

Given this brief overview of irrealis in Moose Cree and other Algonquian languages, we now turn to an exploration of irrealis in Blackfoot with the aim of understanding whether irrealis is indeed a relevant category in this language.

**4. “Irrealis” in Blackfoot.** Blackfoot *-opi*, and its variants *-htopi*, *-ohtopi*, and *-wahtopi*, is labeled “unreal” by Frantz (1991) as well as Taylor (1969), who suggests that it is “found in the apodosis of conditional sentences” (Taylor 1969:170), and generally conveys a counterfactual. James (1991, citing Proulx, p.c.) suggests that, like Moose Cree *-pan*, Blackfoot *-opi* probably also developed from PA *\*(e)pan*.

The Blackfoot “unreal” paradigm is used in past counterfactual (10) and present counterfactual (11) conditions:

- (10) Nitsítssáyoyi**htopi**, nitáaksoyi ánnohka  
 nit-it-say-Ioyi-**htopi** nit-áak-Ioyi annohka  
 I-then-neg-eat-**unreal** 1-fut-eat now  
 ‘If I hadn’t eaten then, I’d eat now’ (Frantz 1991:115, ex. x)
- (11) kátá’yo’kaa**wahtopi**yaawa, áaksstaayaaw mááhksoyssaawa  
 kátá’yo’kaa-**wahtopi**-yi-aawa yáak-sstaa-yi-aawa m-ááhk-Ioyi-hsi-aawa  
 neg-sleep-**unreal**-3p-PRO fut-want-3p-PRO 3-might-eat-conj-pro  
 ‘If they weren’t asleep, they’d want to eat’ (Frantz 1991:115, ex. y)

If irrealis were a grammatical category in Blackfoot, we would expect that multiple “logically unreal” contexts would be marked with similar morphology (following Mithun 1999, minimally counterfactuals and conditionals). This prediction is not borne out, since counterfactuals are marked by the “unreal”, while conditionals are marked by the subjunctive:

- (12) Ikkamín**mmiinnaaniki**, nitáaksowatoo’pinnaana  
 ikkam-Ini-**mmiinnaaniki** nit-yáak-Iowatoo-‘p-innaan-wa  
 if-see(TI)-**1p(subjunctive)** 1-fut-eat(TI)-theme-1p-in.s  
 ‘If we see it, we’ll eat it’ (Frantz 1991:113, ex. m)

Moreover, other logically unreal contexts are marked in even different ways. Imperatives are marked by the suffixes *-t* or *-k*, and not by the “unreal” paradigm:

- (13) Ooyít!                      Ooyík!  
 ooyi-**t**                      ooyi-**k**  
 eat(AI)-**2s(imp)**        eat(AI)-**2p(imp)**  
 ‘Eat!’ (Frantz 1991: 114, ex. r)

Future tense is marked by the prefix *yáak-*, and not by the “unreal” paradigm<sup>8</sup>:

- (14) Nitáakitsiniki  
 nit-**yáak**-itsiniki  
 1-**fut**-relate  
 ‘I will tell a story’ (Frantz 1991: 31, ex. b)

Interestingly, the only morpheme which, to our knowledge, appears in more than one logically unreal context is *-hpa*, which encodes what Frantz (1991) terms “nonaffirmative”. It appears in both yes/no questions (15) and negative statements in the independent verb paradigm (16):

- (15) Yes/no questions  
 Kikáta’yáaka’po’taki**hpa**?  
 k-Ikáta’-yáak-a’p-o’taki-**hpa**  
 2-interrog-fut-PREF-work-**nonaffirm**  
 ‘Will you work?’ (Frantz 1991:133, ex. J)

- (16) Negative statements in the independent verb paradigm  
 Nimáátáóoyi**hpa**.  
 n-Imáát-á-ooyi-**hpa**  
 1-neg-dur-eat-**nonaffirm**  
 ‘I’m not eating.’ (Frantz 1991:85)

Table 2 summarizes the ways in which logically unreal contexts in Blackfoot are marked morphologically. With the exception of future and imperative, these morphemes are taken as representative of their respective paradigms, within which morpheme shape is dependent on person marking and verb order:

**Table 2. Morphology used in Blackfoot logically unreal contexts.**

Counter-factual	Conditional	Future	Imperative	Negative	Yes/No Questions
<i>-opi</i>	<i>-iniki</i>	<i>-yáak</i>	<i>-t, -k</i>	<i>maat-...-hpa</i>	<i>-hpa</i>

Of the six logically unreal contexts given in Table 2, *-opi* is used in only one of those contexts, the counterfactual. Although *-opi* probably did develop from the PA *\*(e)pan* preterit marker and now has an irrealis usage parallel to Moose Cree *-pan*, its modern use in Blackfoot is too restricted to constitute a grammatical category irrealis.

To our knowledge, no origin has been proposed for the *-(ht)-* portion of the Blackfoot morpheme, but we cannot help noticing the similarity to Moose Cree *-htay*, which James argues had an original irrealis meaning. To account for why these two morphemes *-pan* and *-htay* have come to form a single paradigm in Moose Cree, James appeals to Steele's (1975) claim that past and irrealis "have in common the semantic primitive DISSOCIATIVE.... Past time is dissociated from present time. Irrealis is dissociated from reality" (p. 216-7). A similar explanation might be appropriate for Blackfoot where reflexes of these two morphemes may have merged to form a single morpheme. Unlike Moose Cree where, in addition to having a past tense morpheme *ki:-*, the preterit paradigm has both an irrealis usage and an imperfective past usage, Blackfoot *-opi* does not seem to have distinct imperfective past usage; furthermore, Blackfoot lacks an overt past tense morpheme (see Frantz 1991 and Armoskaite to appear). Thus, while the Moose Cree preterit paradigms seems to have retained some of the meaning of *both* the original PA morphemes, Blackfoot has not.

**5. Conclusion.** In this section we briefly summarize our proposals and discuss two issues for further research: (i) the patterning of yes/no questions and negation and (ii) the relation between the past and irrealis.

**5.1. Summary.** Unlike in Moose Cree, Blackfoot counterfactuals and future conditionals are not marked in the same way. The only two logically unreal contexts which are marked similarly are negation and yes/no questions. If irrealis were a grammatical category in Blackfoot, we would expect to find similar morphology across at least some logically unreal contexts (minimally counterfactuals and conditionals). Since the only logically unreal contexts which pattern together in Blackfoot are negation and yes/no questions, neither of which use the "unreal" paradigm, we conclude that irrealis is not a grammatical category in Blackfoot.

Blackfoot *-opi* probably did develop from PA *\*(e)pan*. Counterfactual usage of the modern reflexes of the PA preterit morpheme is well-documented in other Algonquian languages. Thus both the observed context of use and the phonetics of Blackfoot *-opi* point to this as a reflex (James 1991:291, citing Proulx, p.c.). While the morphological and semantic relation between Moose Cree *-pan/-htay* and Blackfoot *-(ht)opi* is clear, we suggest that the restriction of Blackfoot *-opi* to a single unreal context is not enough to justify irrealis as a grammatical category in the language.

Bybee (1998b) suggests that "the term 'irrealis' is simply too general to be useful, except as a pointer to a very broad domain" (p. 269; see also Bybee et al. 1994 for a

claim against a universal category irrealis). There are languages where reference to the category irrealis seems to be of use (e.g., Caddo; see also Comrie (1985), Mithun (1991) and Bybee (1998a) for further examples). Note that the lack of irrealis as a grammatical category in Blackfoot does not necessarily have implications for the status of irrealis in other languages. Claims about languages that apparently lack the categories tense, number or gender have not all concluded that these categories are irrelevant cross-linguistically, only that they are not relevant in those languages. Likewise, we have shown that irrealis is not a relevant category in Blackfoot, but may still be a relevant category in other languages. Our proposal supports Kinkade's (1998) important observation that logical irrealis does not predict grammatical irrealis.

**5.2. Issues for Further Research.** It remains to be explained why yes/no questions and negation pattern together (marked by the nonaffirmative *-hpa*), but not with content questions, counterfactuals, or conditionals (or other “logically unreal” contexts). Examining non-affirmative endings in Blackfoot, Louie (2008) shows that they have the same distribution as negative polarity items (NPIs) cross-linguistically: they appear in questions and negative constructions, but not in corresponding positive constructions. This analysis may account for why it might be the case that among the logically unreal contexts in Blackfoot, yes/no questions and negative constructions are the only two that pattern together morphologically. Her analysis of non-affirmative in Blackfoot as NPIs seems to be restricted to non-speech act participants and thus a remaining question is whether it extends to speech act participants as well.

As negative statements are marked with the nonaffirmative suffix in addition to the negative prefix, Blackfoot apparently treats positive assertions and negative assertions differently (perhaps negative statements are not treated as assertions at all), an understanding of which we leave for future research. Furthermore, it remains to be seen whether the Blackfoot “unreal” presupposes that the counterfactual condition is false. It has been shown that in English, this is not necessarily the case; although the typical interpretation of (17) below is that John did not come, as Palmer (1986: 191) suggests the sentence “could be used where the speaker does not know whether John came or not”:

(17) If John had come, Mary would have left.

Matthewson, Rullman and Davis (2005:7) argue that the irrealis usage of *ka-* in St'át'imcets (Interior Salish) “requires that the proposition it operates on is false...the falsity cannot be cancelled”. We leave the presupposition facts about counterfactuals in Blackfoot as another issue for further research.

The final remaining question concerns what exactly the relationship is between past and irrealis in Blackfoot. Of particular interest is whether there is a parallel between the development of a PA preterit morpheme into a Blackfoot counterfactual marker and

the lack of an overt past tense morpheme in Blackfoot (see Frantz 1991 and Armoskaite to appear; but also Reis Silva and Matthewson 2007 for the claim that Blackfoot has a phonologically null past tense and Ritter and Wiltschko 2004 for the claim that Blackfoot lacks a T node). This contrasts with Moose Cree, for example, where the preterit paradigm marks imperfective aspect in the past in addition to irrealis. Given the different usages of \*-(e)pan reflexes in modern Algonquian languages, the morpheme most likely had a more complex usage than simple past (possibly related to irrealis). The inconsistency of the category irrealis across Algonquian languages points to the need for an in depth study of irrealis (and more broadly, mood) across the Algonquian language family.

### Notes

<sup>1</sup>Our thanks go to Joshua Birchall, Lisa Matthewson, Mizuki Miyashita, Meredith Ward, Becky Wood and audience members at WAIL 11 for valuable discussion.

<sup>2</sup>The Blackfoot data in this paper is drawn from Frantz (1991).

<sup>3</sup>See, however, recent work by Ritter and Wiltschko (2004, 2005), Reis Silva and Matthewson (2007), Reis Silva and Gougie (2007), Durham (2007), Louie (2008) and Armoskaite (to appear) for analyses of Blackfoot tense and aspect.

<sup>4</sup>Given that irrealis terms refer to a world other than the current one, it might be natural to include “past” or “preterit” in this paradigm. This is not the case in the morphology of languages in which irrealis is a distinct grammatical category. Interestingly, though, there does seem to be a relationship between past and irrealis. Steele (1975) suggests that past and irrealis both signify remoteness from present reality (see §4).

<sup>5</sup>According to their study, however, Uhlenbeck’s (1938) description of irrealis in Blackfoot is the earliest recording of the term.

<sup>6</sup>We follow Chung and Timberlake (1985) and assume that past and present conditions reduce to counterfactual conditions while future conditions are neither actual nor counterfactual, but potential.

<sup>7</sup>In Montagnais, Naskapi, and East Cree, this morpheme is directly descended from Proto-Algonquian \*-(e)sapan, which is in turn purported to be a combination of \*-(e)pan and \*-(e)san (James et al. 2001:246).

<sup>8</sup>See also Reis Silva and Gougie (2007) for analysis of two future tense morphemes in Blackfoot.

## References

- Armoskaite, Solveiga. To appear. 'Blackfoot stem initial eventive predicates: shape shifting perfectives.' *Proceedings of the 39<sup>th</sup> Algonquian Conference*. Toronto: York University.
- Bybee, Joan (ed.) 1998a. *Anthropological Linguistics* 40.
- Bybee, Joan. 1998. "'Irrealis" as a grammatical category.' *Anthropological Linguistics* 40:257-271.
- Bybee, Joan, Revere Perkins and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in Languages of the World*. Chicago/London: University of Chicago Press.
- Chung, Sandra and Alan Timberlake. 1985. 'Tense, aspect and mood.' *Language Typology and Syntactic Description* 3, ed. by T. Shopen, 202-258. Cambridge: Cambridge University Press.
- Durham, Joel. 2007. 'The "durative" in Blackfoot.' *Proceedings of SULA 4: Semantics of Under-Represented Languages in the Americas*, ed. by Amy Rose Deal, 49-64. (University of Montana Occasional Publications in Linguistics 35.) Missoula: University of Montana Press.
- Ellis, C. Douglas. 1961. 'The So-Called Interrogative Order in Cree.' *International Journal of American Linguistics* 27:119-124.
- Frantz, Donald G. 1991. *Blackfoot Grammar*. Toronto: University of Toronto Press.
- Hofling, Charles Andrew. 1998. 'Irrealis and Perfect in Itzaj Maya.' *Anthropological Linguistics* 40:214-227.
- James, Deborah, Sandra Clarke and Marguerite MacKenzie. 2001. 'The Encoding of Information Source in Algonquian.' *International Journal of American Linguistics* 67:229-263.
- James, Deborah. 1991. 'Preterit Forms in Moose Cree as Markers of Tense, Aspect, and Modality.' *International Journal of American Linguistics* 57:281-297.
- Kinkade, M. Dale. 1998. 'Is Irrealis a Grammatical Category in Upper Chehalis?' *Anthropological Linguistics* 40:234-244.
- Louie, Meagan. 2008. Atemporal anchoring of individuals, events and sub-events in Blackfoot: Consequences for the syntax-semantics interface. MA thesis, University of Toronto.
- Matthewson, Lisa, Hotze Rullman and Henry Davis. 2005. 'Modality in Stl'atl'imcets.' *Papers for the International Conference on Salishan and Neighbouring Languages* 40, ed. by J. C. Brown et al., 166-183. (University of British Columbia Working Papers in Linguistics 16.) Vancouver: University of British Columbia Press.

- Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge: Cambridge University Press.
- Reis Silva, Amélia and Jennifer Glougie. 2007. 'The form and function of two futures in Blackfoot and St'at'imcets.' *Proceedings of SULA 4: Semantics of Under-Represented Languages in the Americas*, ed. Amy Rose Deal, 183-196. (University of Montana Occasional Publications in Linguistics 35.) Missoula: University of Montana Press.
- Reis Silva, Amélia and Lisa Matthewson. 2007. 'An instantaneous present tense in Blackfoot.' *Proceedings of SULA 4: Semantics of Under-Represented Languages in the Americas*, ed. Amy Rose Deal, 197-213. University of Montana Occasional Publications in Linguistics 35. Missoula: University of Montana Press.
- Ritter, Elizabeth and Martina Wiltschko. 2005. 'Anchoring Events to Utterances without Tense.' *Proceedings of the 24<sup>th</sup> West Coast Conference on Formal Linguistics*, ed. John Alderete et al., 343-351. Somerville: Cascadilla Proceedings Project.
- Ritter, Elizabeth and Martina Wiltschko. 2004. 'The lack of tense as a syntactic category: evidence from Blackfoot and Halkomelem.' *Papers for the 39<sup>th</sup> International Conference on Salish and Neighbouring Languages*, 341-370. (University of British Columbia Working Papers in Linguistics 14.) Vancouver: University of British Columbia Press.
- Rogers, Jean H. 1978. 'Differential Focusing in Ojibwa Conjunct Verbs.' *International Journal of American Linguistics* 44:167-179.
- Steele, Susan. 1975. 'Past and irrealis: just what does it all mean?' *International Journal of American Linguistics* 41:200-217.
- Taylor, Allan Ross. 1969. *A Grammar of Blackfoot*. Ph.D. dissertation, University of California, Berkeley.
- Uhlenbeck, C. C. 1938. *A Concise Blackfoot Grammar*. Amsterdam: North Holland Publishing Company.
- Wolfart, H. Christoph. 1973. 'Plains Cree: A Grammatical Study.' *Transactions of the American Philosophical Society*. 63:1-90.

# Determining Gender Markedness in Wari'

Joshua Birchall

University of Montana

## 1. Introduction

Grammatical gender is rare across Amazonian languages. The language families that possess grammatical gender properties are Arawak, Chapakuran and Arawá. Wari', a Chapakuran (Txapakura) language of Brazil, is unusual among these languages in that it possesses three gender categories: what we can refer to as masculine, feminine and neuter. While efforts have been made to posit the functionally unmarked gender in Arawak and Arawá languages, no such analysis has been given for any Chapakuran language. In this paper I propose that neuter is the functionally unmarked grammatical gender of Wari'.

Aikhenvald (1999:84) states that masculine is the functionally unmarked gender for non-Caribbean Arawak languages, while Dixon (1999:298) claims that feminine is the unmarked gender in Arawá languages. Functionally unmarked gender refers to the gender that is obligatorily assigned when its designation is otherwise opaque. With attempts, as in Greenberg (1987), to claim a genetic relationship between Chapakuran languages and the Arawak and Arawá families, the presence of gender is a major grammatical similarity and possible motivating factor for such a hypothesis and thus should be examined.

In order to investigate this claim, I first present an outline of the Wari' gender distinctions and then proceed to describe how gender is manifested within the clause. A case is then presented for neuter as the functionally unmarked gender. This claim is further examined for its historical implications, especially regarding genetic affiliation. All of the data<sup>1</sup> used in this paper come from Everett and Kern (1997).

## 2. Gender Distinctions

The three gender distinctions we find in Wari' are largely determined by the semantic domain of the noun. Even though gender assignment is somewhat arbitrary in the language, Everett and Kern (1997) use two features as general domains for distinguishing the genders: human and female.

The feminine gender is by far the most semantically restricted of the gender classes, containing only human females, children, and human collective nouns that are of mixed sexes. Some examples feminine nouns are *narima* "woman", *panxi* "child", and '*oro wari*' "people". Semantically, feminine nouns can be described as [+feminine], [+human].

The masculine gender has a less restricted semantic domain than feminine, containing nouns that are human males, certain animals, plants, and insects, along with "culturally significant objects" (Everett and Kern 1997:296). Cultural significance is described as "familiar to the Wari' before contact" and "played an important role in the pre-contact worldview". Some example masculine nouns are *tarama*' "man", *hwam*

“fish” and *panawo* ‘moon’. Given the semantic features mentioned above, masculine nouns can be described as [-feminine].

The neuter gender has the least restricted semantic domain for assignment. Non-human entities can be assigned to the neuter class, including inanimate objects, most plants, some animals, fish, and insects, as well as loanwords, body parts and subordinate clauses. Some example neuter nouns are *xe* “fire”, *tononoin* “grass”, *came* “capybara”, and *tocoxi* ‘eye’. Neuter nouns can be described as having the semantic feature of [-human].

### 3. Gender Agreement

Gender is not overtly marked on the head noun, such as in the –o/-a distinction in certain Romance languages. Instead, gender is realized through agreement within the clause. Gender agreement features are manifested through verbal inflectional clitics<sup>2</sup> (VICs), possessive constructions, demonstratives and prepositions. Gender is only marked in the third person and only human nouns trigger number agreement.

Verbal inflectional clitics appear at the right edge of the verb and indicate subject and object agreement information along with tense and mood. Agreement between the verb and its arguments is realized on the VIC. A verbal inflectional clitic can be composed of one or two morphemes, depending on whether the clause is transitive. Intransitive VICs, as seen in (1), contain tense and mood information and agree with the subject in number and person. As example (2) illustrates, transitive VICs also contain a morpheme that agrees with the object in person, number and gender.

- (1) mao **na**.  
go **3s:rp/p**  
“He went” (Everett and Kern 1997:121)

- (2) cao’ **‘ina-on** hwam.  
eat **1s:rp/p-3sm** fish  
“I ate the fish” (Everett and Kern 1997:152)

Depending on the construction, some VICs can be tenseless, carrying only argument agreement properties. Intransitive and transitive tenseless VICs can be seen in (3) and (4) respectively.

- (3) ‘om ca mao **ca**.  
not:exist infl:nrp/p go **3sm**  
“He did not go” (Everett and Kern 1997:121)

- (4) ma-in ca to’ ‘ac **ca-em?**  
that:prox:hearer-n infl:nrp/p hit travel **3sm-2s**  
“Where (on your body) did he hit you?” (Everett and Kern 1997:51)

A second way that gender agreement is realized within a clause is through demonstrative constructions. Demonstratives show a gender agreement relationship with the noun that they modify. In addition to marking for proximity to the speaker or hearer, they agree with the gender of the head noun. The forms only distinguish between neuter and non-neuter nouns, as illustrated in the following examples:

- |     |   |     |   |
|-----|---|-----|---|
| (5) | hwam <b>cwa'</b><br>fish <b>this:m/f</b><br>“this fish”                   | (6) | xirim <b>ca'</b><br>house <b>this:n</b><br>“this house”                             |
| (7) | pana <b>cain</b><br>tree <b>that:n:distal</b><br>“that tree (over there)” | (8) | carawa <b>ma'</b><br>animal <b>that:prox:hearer</b><br>“that animal (close to you)” |

A third way that gender agreement is realized within a clause is through possessive constructions. Possession relationships can be marked in two ways<sup>3</sup>: with a nominal inflectional clitic (NIC) following the head noun (9) or a possessive suffix (10). Both NICs and the possessive suffixes occur at the right edge of possessed nouns and contain information on the person, number and gender of the possessor.

- |     |  |      |  |
|-----|--|------|--|
| (9) | xirim <b>nucun</b> Mirin<br>house poss: <b>3sm</b> name:m<br>“Mirin’s house” | (10) | xiri- <b>con</b><br>house- <b>3sm</b><br>“his house” |
|-----|--|------|--|

Prepositional constructions are the final major way that agreement is realized within a clause. There is a single inflected prepositional paradigm used to mark oblique objects on the verb. Prepositions agree with their complements in person, number and gender. Selection of a verbal complement is dependant on the semantic hierarchy in (11) below.

- (11) Goal > Circumstance > Theme > Benefactive > Comitative > Location > Time  
(Everett and Kern 1997:4)

All other arguments present in the clause are marked as oblique objects through the use of a preposition. In example (12) below, the Patient *hwam* (m) “fish” is treated as an oblique object through agreement with the preposition. *Humaxicam* (f) “her children” is acting as the Goal and is therefore marked on the VIC.

- (12) hoc mi' **nanam** **con** hwam humaxicam  
cook give **3s:rp/p-3pf** **prep:3sm** fish children-3sf  
“She cooked fish for her children” (Everett and Kern 1997:125)

With an understanding of how gender is realized within a clause, it is now possible to examine the case for neuter as the functionally unmarked gender.

#### 4. Case for Unmarked Neuter

In this section I argue that neuter is the functionally unmarked gender in Wari' based on the following evidence: the grammatical incorporation of loanwords, gender assignment to mixed groups, interrogatives and subordinate clause agreement. This claim is of significance because Everett and Kern (1997) do not analyze for gender markedness.

The most convincing piece of evidence for an unmarked neuter is that when new words enter the lexicon through borrowing they are “assigned to the neuter gender class” (Everett and Kern 1997:301). The gender of borrowed lexical items in the donor language does not influence gender assignment, e.g., *sal* “salt” is masculine in Portuguese yet is assigned neuter in Wari'.

#### (13) Example Loanwords from Portuguese

<i>sal</i> “salt”	<i>canoa</i> “canoe”
<i>semana</i> “week”	<i>quilometro</i> “kilometer”
<i>motor</i> “motor”	<i>dinheiro</i> “money, price”
<i>mesa</i> “table”	<i>segunda-feira</i> “Monday”

Although noting that multiple factors like phonological shape and semantic domain can influence the gender assignment of loanwords, Poplack et al. (1982:21) state “that loanwords tend to adopt the unmarked gender of the language into which they are borrowed.” While Wari' does not condition gender assignment on phonological factors, semantic domain plays a central role. It must be noted that these instances of loanwords from Portuguese, which are all incorporated as neuter nouns, may not fall into the semantic domain of assignment into any gender besides neuter. However, examples in (13) are only the instances in the available materials. I expect that more data will demonstrate that all loanwords in Wari' belong to the neuter gender class.

A second piece of evidence for an unmarked neuter is that in instances where gender is assigned to groups composed of entities that are both masculine and neuter, the group is always treated as neuter<sup>4</sup>. This process is known as gender resolution (see Corbett 1991: Ch.9). In example (14) below, we see that a group composed of two masculine entities *me* “bird” and *jowin* “monkey” and a single neuter entity *carawa jimao* “different animals” triggers neuter object agreement on the VIC.

- (14) Om      ca            taraju xo'            camain'    **cain**    'oro me  
 not:exist infl:nrp/p hear    correctly at:all    3sm-3n coll. bird(m)
- 'oro jowin                    cwa' 'oro carawa    jimao.  
 coll monkey:spec.(m) this coll. animal(n) different

“He doesn't hear correctly at all, the birds, the *jowin* monkeys, all different animals.” (Everett and Kern 1997:494)

Corbett (1991) notes that markedness is an inadequate explanation for all gender resolution rules since trinary gender systems must also account for groups that fall outside of the domain of the unmarked gender. For example, in Wari' groups of mixed gender humans are treated as feminine because the entities fall outside of the [-human] semantic domain of the neuter gender class. Since semantic factors are primary in the assignment of gender in Wari', this should be of no surprise. However, the fact that neuter overrides masculine even in instances where there is no difference in semantic domain within the group, as in (14), strengthens the claim that neuter acts as a default when gender assignment is unclear.

Opacity in gender assignment is most evident in interrogative structures where it is uncertain which gender the object being questioned belongs to. Wari' primarily employs two interrogative markers, *cain'* "that:n:distal" and derivatives of *ma'* "that:prox:hearer". Both forms are morphologically derived from the demonstrative paradigm. Derivatives of *ma'*, which can be inflected for gender, are used in questions denoting "who?", "whom?" or "which?". Except for the neuter form *main*, which can also mean "how?" or "where?", the derivatives of *ma'* are used in instances where the object being questioned is semantically [+human].

The other interrogative marker *cain'* is used to mark questions denoting "what?", "when?", "how?", "how many?". The objects being questioned typically fall outside of the semantic domain of [+human]. We can see through agreement with the neuter inflectional particle *ca* "infl:nrp/p" that *cain'* still carries its neuter properties.

- (15) **Cain'**        **ca**        tomi'    cama?  
           that:n:distal infl:nrp/p speak 3sf  
           'What did she say?' (Everett and Kern 1997:51)

Since it is unclear what gender class the object being questioned belongs to, a neuter form is employed. The use of the unmarked gender in interrogative structures where the gender of the referent is unavailable is not uncommon<sup>5</sup>. Dixon (1995) uses similar criteria from interrogatives in Jarawara (Arawá) to posit feminine as its functionally unmarked gender.

The final piece of evidence that points towards neuter as the functionally unmarked gender in Wari' is agreement with a subordinate clause. All subordinate clauses are introduced by a preposition. Treated as a prepositional object, subordinate clauses trigger agreement as if they were a neuter entity.

- (16) Querec te        'inem        **pain**    ca        hwet    mapa'  
           See     father:1s 1s:rp/p-2s prep:3n infl:nrp/p appear 2s-1s  
           'I am happy that you came to me' (Everett and Kern 1997:297)

Since a clause cannot typically possess a grammatical gender, Wari' defaults to neuter because agreement is necessary in a prepositional construction.

The four pieces of evidence presented above all demonstrate that when gender agreement is necessary but opaque, the language uses neuter as the functionally

unmarked gender class for assignment. While semantic properties are the main driving force behind the assignment of gender, when these properties alone are insufficient to determine proper agreement, markedness helps to motivate the designation of grammatical gender. This claim not only contributes to the typological data available for Amazonian languages, but also has some historical implications.

### 5. Historical Implications

Greenberg (1987) posits a genetic relationship between the Chapakuran language family and the Arawak and Arawá families. He includes Chapakuran languages in an “Arawakan” subgrouping with these two other families. Given the broad scope of his research, Greenberg presents little evidence to motivate this claim. The only evidence he cites is that there are traces of *n*- 1<sup>st</sup> person singular and *m*- 2<sup>nd</sup> person singular in the pronouns (p.51). However, this evidence is unconvincing considering the high occurrence of these phonemes in pronominal sets of unrelated languages all across the Americas, e.g. the Algonquian and Salishan families. Acknowledging the paucity of data, he states that his analysis is based off of earlier works by Paul Rivet and others.

In Crequi-Montfort and Rivet (1913), the possibility of a genetic relationship between Chapakuran and Arawak languages was entertained. However, after a comparative analysis between the families, they conclude regardless of notable resemblances that:

"Despite the considerable number of these matches, we cannot believe that we can, for the moment at least, conclude a parent, even distant. The presence of these foreign words in our vocabulary may indeed be explained by borrowing, since all the peoples of our group are in contact on all sides with Arawak tribes. It is indeed remarkable that these borrowed words never coincide with those that we have established the parents of from different languages of the Chapakuran group" (Crequi-Montfort and Rivet 1913:146, translated by author)

It seems that even Rivet was in doubt of a genetic relationship between Arawak and Chapakuran languages. It is thus unclear what led Greenberg to propose a closer genetic relationship between Arawak and Chapakuran languages than other languages in his Equatorial grouping.

Everett and Kern (1997:3) suggest that another factor may have contributed to Greenberg’s analysis: the presence of grammatical gender. However, they dismiss the claim of a genetic relationship between Arawak, Arawá and Chapakuran languages by stating that “they share no obvious cognates and the presence of gender is hardly a sufficiently reliable criterion alone to posit any genetic affiliation.” Upon close inspection of these families’ gender systems, we find that they are considerably different.

Arawak languages typically have a binary gender class distinction. Aikhenvald (1999:84) claims that masculine is the functionally unmarked gender in the non-Caribbean Arawak languages. Only Ignaciano (Southern) and Palikur (Eastern) show a

neuter distinction in their pronominal systems. At least in the case of Ignaciano, the presence of a third gender category may be a result of contact.<sup>6</sup>

Arawá languages also only possess a masculine/feminine gender distinction. Dixon (1999:298) analyzes feminine as the functionally unmarked gender category. Although its genetic relationship to Arawak languages has been proposed since Ehrenreich (1897), recent scholarship has cast doubt on such a claim. Aikhenvald and Dixon (1999:14) state that “it must be concluded that there is no evidence whatsoever that... the Arawá and Arawak language families are genetically related”.

Without a complete reconstruction of Proto-Chapakuran phonology and grammatical forms<sup>7</sup>, a comparative approach to resolving the question of its relationship to Arawak languages is untenable. However, we can rely on grammatical properties such as gender systems to provide insight into such questions.

The available data on Chapakuran languages all point towards an unmarked neuter across the family. The Moré lexicon in Angenot de Lima (2000) and a Torá wordlist from Nimuendajú (1925) show that these languages use similar semantic criteria for the assignment of gender as Wari'. The only available research on 'Oro Win in Popky (1999) shows a very similar gender system to Wari'. Without an extensive analysis we can provisionally assume that neuter is the functionally unmarked gender across Chapakuran languages.

If one were to posit a genetic relationship between Arawak, Arawá and Chapakuran languages, an account of the differences in their gender systems is in order, especially since we see a high degree of regularity in these systems within their respective families. An account for the innovation of a third grammatical gender class in Chapakuran languages, as well as a shift in (un)markedness to this novel category, would be critical in demonstrating a genetic relationship among these groups.

Since Greenberg's claims in *Language in the Americas* were published, they have been widely adopted by other branches of social sciences and the public at large. While grammatical properties can help us investigate such claims, more work involving comparative reconstruction is necessary before any definitive claim on genetic affiliation can be reasonably accepted. Therefore, due to a lack of evidence as well as considerable methodological concerns (see Campbell 1991), the claim that Chapakuran, Arawak and Arawá languages share a genetic relationship must be rejected.

## 6. Conclusions

In this paper, I have demonstrated that neuter is the functionally unmarked grammatical gender in Wari'. This claim is supported by four main facts. First, loanwords are always assigned to the neuter gender class. Second, gender resolution shows that groups of mixed genders trigger neuter agreement when semantically permitted. Third, the neuter gender is used in interrogative structures where the gender of the object being questioned is unclear. Finally, subordinate clauses trigger neuter agreement with the preposition that introduces the clause. This analysis is of significance because no work on gender markedness in Wari' has been attempted.

This claim of gender markedness can be provisionally extended to all Chapakuran languages. In regards to proposals of a genetic relationship between Chapakuran,

Arawak and Arawá languages, there is not significant evidence to support this conclusion and the claim must be rejected. In general, the terms “Arawakan” or “Macro-Arawakan” should be avoided since they acknowledge the existence of this relationship without any basis in an empirically driven approach of comparative reconstruction. Gender markedness in Wari’ helps to shed light on the dubiousness of such a proposal by highlighting the stark differences between the gender systems of these languages. As more data are available and the analyses of these languages further develop, we will be able to get a clearer perspective on the historical relationships in Western Amazonia.

### Notes

\* Thanks to Leora Bar-el, Ryan Denzer-King, Mizuki Miyashita, Anthony Mattina and the audience at WAIL 2008 for their help and comments. Additional thanks goes to Barbara Kern for help with some of the data. All errors are my own.

<sup>1</sup> Because the traditional Wari’ orthography employs an apostrophe [‘] to represent a glottal stop, a double quotation mark will be used to contain the gloss accompanying the data.

<sup>2</sup> These bundles of agreement information have been analyzed as clitics rather than affixes principally on a phonological basis. For a more detailed discussion of this analysis see Everett and Kern (1997:332-3).

<sup>3</sup> Inalienably possessed nouns (e.g. kinship terms and body parts) may only take the possessive suffix.

<sup>4</sup> Groups consisting of entities of both masculine and feminine gender are treated as feminine. No data is available for groups that contain feminine and neuter entities.

<sup>5</sup> Thanks to Brook Lillehaugen for pointing out that the use of the gender with the broadest semantic domain in questions may result from the need for sufficient vagueness in the proposition.

<sup>6</sup> d’Orbigny (1839) notes a high degree of contact between the Chapakura and the Moxo (Ignaciano) people upriver.

<sup>7</sup> An attempt at a phonological reconstruction of Proto-Chapakura is presented in Angenot de Lima and Angenot (2000).

### Abbreviations

rp/p = realis past/present

m = masculine

f = feminine

n = neuter

s = singular

p = plural

coll = collectivizer

infl = inflectional particle

prep = preposition

prox = proximal

poss = possessor

## References

- Aikhenvald, A. and R.M.W. Dixon (eds.). 1999. *The Amazonian Languages*. Cambridge: University Press.
- Aikhenvald, Alexandra. 1999. 'Arawak'. In A. Aikhenvald and R.M.W. Dixon (eds.), *The Amazonian Languages*. 65-106. Cambridge: University Press.
- Angenot de Lima, G. 2000. *Description Phonologique, Grammaticale et Lexicale du Moré, Langue Amazonienne de Bolivie et du Brésil*. Porto Velho: EDUFRO
- Angenot de Lima, G. and J.P. Angenot. 2000. 'A reconstrução do Protochapakura.' In Hein van der Voort and Simon van de Kerke (eds.), *Indigenous Languages of Lowland South America*. 149-164. Leiden: ILLA, CNWS Publications.
- Campbell, Lyle. 1991. 'On so-called Pan-Americanisms (in Notes and Reviews).' *International Journal of American Linguistics*, Vol. 57, No. 3: 394-399.
- Corbett, Greville. 1991. *Gender*. Cambridge: University Press.
- Crequi-Montfort, G. and P. Rivet. 1913. 'Linguistique bolivienne: La famille linguistique çapacura.' *Journal de la Société des Américanistes de Paris*, 10:119-171.
- Dixon, R.M.W. 1995. 'Fusional Development of Gender Marking in Jarawara Possessed Nouns.' *International Journal of American Linguistics*, Vol. 61, No. 3:263-294.
- Dixon, R.M.W. 1999. 'Arawá'. In A. Aikhenvald and R.M.W. Dixon (eds.), *The Amazonian Languages*. 292-306. Cambridge: University Press.
- d'Orbigny, A.D. 1839. 'Nación Chapacura.' *L'homme américain de l'Amérique méridionale, considéré sous ses rapports physiologiques et moraux*. 340-3. Paris.
- Ehrenreich, P. 1897. 'Materialen zur Sprachenkunde Brasiliens: Vokabuläre von Purus-Stämmen.' *Zeitschrift für Ethnologie*, 24:59-71.
- Everett, Daniel and Barbara Kern. 1997. *Wari': The Pakaas Novos Language of Western Brazil*. London: Routledge.
- Greenberg, Joseph. 1987. *Language in the Americas*. Stanford: Stanford University Press.
- Nimuendajú, Curt. 1925. 'As tribus do Alto Madeira.' *Journal de la Société des Américanistes de Paris*, 17:137-172.

Popky, Donna. 1999. *Oro Win: A Descriptive and Comparative Look at an Endangered Language*. University of Pittsburg. M.A. Thesis.

Poplack, S., A. Pousada and D. Sankoff. 1982. 'Competing Influences on Gender Assignment.' *Lingua*, 57:1-28.

## Neologisms in Indigenous Languages of North America

Ryan Denzer-King

University of Montana

**1. Introduction<sup>1</sup>.** North American languages have been exposed to a wide range of new ideas, animals, and objects since the arrival of Europeans four hundred years ago. Many words original to indigenous languages are not morphologically simple, for example, Blackfoot *apahsípoko*, from *ipahs*, ‘mushy, curdled, thickened’ + *ipoko*, ‘taste’. These descriptive words and phrases may have acted as a model for the coining of morphologically complex words to describe new items. This paper compares a group of words (selected so as not to be words that would have existed before European contact) both cross-linguistically and within large families, using languages from several different language groups (including Algonquian, Salish, and Na-Dene). While this paper shows that language-internal word coining is in fact a widespread feature of Native American languages (Leonard 2008), statistical tests show that the strategies employed by language communities differ according to what type of object is being named. In addition, even languages that use a descriptive method for word coining often describe different aspects of the same object. These different views of the same object may be taken to represent different ways of interacting with the world, though a cultural investigation of this sort is beyond the scope of this paper.

I first discuss the methodology used in gathering and analyzing the data (2), including overviews of the language families and individual languages surveyed and the sources used for each language. Section 3 discusses the different types of lexical innovation employed by North American languages. Section 4 presents an analysis of the data collected, discussing the similarities and differences among the languages surveyed. The final section presents the conclusions of this research (5), including a summary of the arguments put forth in the paper, the implications for language revitalization and preservation, and the difficulties inherent in such a survey.

**2. Methodology.** During this project I consulted dictionaries as well as, when possible, speakers. Native speakers are especially invaluable because they may have critical insight into the transparency, usage, and sometimes history of the various terms. My goal in selecting words to compare was to create an inventory of words that would represent a balanced sample, that is, a sample including representatives from all categories I felt might be relevant, as opposed to a completely random sample. This meant including terms that existed at the time of European contact (e.g., ‘gun’), when indigenous language use was widespread, terms that have evolved or developed since the

decline of indigenous language use (e.g., ‘radio’), as well as a mixture of terms relating to the natural world (e.g., ‘cat’) and man-made objects (e.g., ‘screwdriver’). I also selected some objects which were similar to items present in pre-Columbian indigenous societies (e.g., ‘road’) in order to ensure at least a few tokens of semantic extension.

**2.1. Language Overviews<sup>2</sup>.** Because this paper compares strategies within and across linguistic families, it is necessary to be aware of forms which are similar because of genetic inheritance or areal diffusion, rather than because of using similar strategies for lexical innovation. Thus it is important to understand, at least generally, which languages surveyed are in contact, as well as which are related and how closely. Table 1 gives the major genetic family and smaller subbranch for each languages, as well as where the languages is or was spoken. All references to genetic classification and location are from Gordon (2005). (AB = Alberta, AK = Alaska, AZ = Arizona, BC = British Columbia, CA = California, IA = Iowa, ID = Idaho, IL = Illinois, IN = Indiana, OK = Oklahoma, MT = Montana, NM = New Mexico, OR = Oregon, WA = Washington.)

**Table 1. Language Classifications and Locations.**

Language	Family	Subbranch	Location
Tlingit	Na-Dene <sup>3</sup>	Tlingit	AK panhandle, parts of BC
Koyukon	Athapaskan	Ingalik-Koyukon	AK, Koyukon and Koyukuk rivers
Gwich'in	Athapaskan	Canadian	NE AK, N BC
Navajo	Athapaskan	Apachean	NM, AZ
Western Apache	Athapaskan	Apachean	NM, AZ
Squamish	Salish	Central	SW BC, Vancouver area
Lushootseed	Salish	Central	W WA, around Puget Sound
Thompson	Salish	Interior	S BC, Fraser and Thompson rivers
Okaganan	Salish	Interior	N WA, Colville reservation
Umatilla	Sahaptian	Sahaptin	N OR, Columbia river
Blackfoot	Algonquian	Plains	AB, MT
Miami	Algonquian	Central	IN, NE OK, IL, IA
Kootenai	isolate	isolate	SE BC, N ID, MT
Luiseno	Uto-Aztecan	Takic	S CA

**2.2. Sources.** Since for the most part I was unable to use linguistic consultants for this research, most forms come from dictionaries. The main disadvantage to using dictionaries rather than consulting members of language communities is that few dictionaries of indigenous languages give usage notes (generational, cultural, gender based, etc.) of a given word or phrase, and where multiple entries exist for the same word (e.g., ‘car’ in Blackfoot), it is usually not marked whether these forms are in some kind of

free variation, or whether there are specific differences between dialects or certain speakers. Thus it is often difficult to determine what method of coining new words is most common in a given language, since some entries may include examples of multiple methods without noting which is in more common or more recent usage. Many dictionaries also do not list the full etymologies of certain words, even those which are fully semantically transparent, both diachronically and synchronically. That being said, the publication of such dictionaries makes it possible to do broad surveys such as this one. Such a study of North American languages certainly would not have been possible even thirty years ago.

In this paper, unless otherwise noted, Blackfoot data comes from Frantz (1995), Gwich'in from Gwich'in Language Center (2005), Koyukon from Jetté & Jones (2000), Kootenai from Kootenai Culture Committee (1999), Luiseño from Bright (1968), Lushootseed from Bates (1994), Miami from Costa (1991), Navajo from Young & Morgan (1980), Okanagan from Mattina (1987), Thompson from Thompson & Thompson (1996), and Western Apache from Bray (1998). All Umatilla data is from Thomas Morningowl (p.c.). Tlingit data comes from Roby Littlefield (p.c.) unless otherwise noted. Squamish data comes from Leora Bar-el (p.c.) unless otherwise noted. Any morphemic analysis also comes from these sources unless otherwise noted.

**3. Types of Lexical Innovation.** In order to discuss lexical innovation, it is first necessary to describe the main types of word coining that are used. The three<sup>4</sup> types of lexical innovation this paper focuses on are neologisms (language-internal coinings), borrowings, and semantic extensions. Neologisms are descriptive words or phrases used to describe a hitherto unknown item<sup>5</sup>. Often they utilize productive morphemes in the language to generate a new word for a new item, e.g., Umatilla Sahaptin *pluuswit'awas*, 'computer,' composed of *pluus*, 'brain' + *wit*, abstractive suffix + *awas*, instrumentative suffix. When analyzing neologisms, it is often useful to discuss what aspect of the item in question the language chose to describe. I will use the term feature selection to refer to picking the most perceptually salient attribute to describe when a language is creating a neologism. Borrowings occur when one word is taken directly from another language, usually with concomitant phonological adaptation (e.g., Koyukon *kelaandas*, 'pencil,' from Russian *karandásh*<sup>6</sup>). Semantic extension is what happens when a new term or concept is similar enough to an existing object that the semantic scope of the original word is expanded to include the new item. An example of this is the introduction of the domestic cat to North America. Since felines were already well known, some languages expanded the usage of an existing term to include this new animal (e.g., Miami *pinšiwá*, 'cat, lynx'; Teetl'it Gwich'in *niinjii zhuu*, 'cat,' lit. 'young lynx'). An example from

Blackfoot is the word *stamik*, which used to mean ‘male bison,’ but in modern times it has come to mean ‘steer,’ or ‘bull,’ referring to cattle instead of bison.

**4. Comparison of Language Strategies.** The limited number of possibilities available for coining new words creates some uniformity across languages. However, the application of similar methods cross-linguistically does not necessarily result in lexical items which are semantically similar. In a group of languages that all use descriptive neologisms, for instance, languages may differ widely in feature selection. The words for ‘car’ are presented in table 2<sup>7</sup>. (Ath = Athapaskan, S = Salish, Alg = Algonquian, iso = language isolate, N-D = Na-Dene. In the strategy column, N = neologism, B = borrowing, E = semantic extension.)

**Table 2. Words for ‘car’.**

	<b>Language</b>	<b>Word</b>	<b>Meaning</b>	<b>Strategy</b>
a)	Koyukon (Ath)	kk’o’eelbaalee	that on which one rolls around	N
b)	Gwich’in (Ath)	iitsii khał	metal wagon	N
c)	Navajo (Ath)	chidí	onomatopoetic	N
d)	Lushootseed (S)	tuł <sup>w</sup> us	clapping or pounding sound in front	N
e)	Thompson (S)	ka(h)	borrowing from English	B
f)	Okanagan (S)	nqýqýx <sup>w</sup> ups	stinking end	N
g)	Blackfoot (Alg)	áíksisstoomatokska’si	starts to run without apparent cause	N
h)	Kootenai (iso)	kqaqana’łkqac	travels on its own	N

A comparison of these languages shows about as much variety as is possible, both in strategies for creating new words and feature selection (in neologisms). Koyukon (2a), Gwich’in (2b), Okanagan (2f), Blackfoot (2g), and Kootenai (2h) all use descriptive phrases for the word ‘car’, but each selected a different feature to describe. In Koyukon it is the rolling aspect of a car that is most salient (cf. *baatl*, ‘roll, revolve’), in Gwich’in it is the material makeup of the car, in Okanagan it is the smell produced by the car, and in Blackfoot and Kootenai it is the fact that cars move essentially on their own, with nothing pulling them (compare former ‘horseless carriage’ in English). The similarity between Blackfoot and Kootenai feature selection may be because of their geographical proximity: they border each other in Montana and Alberta. Navajo (2c) and Lushootseed (2d) both refer to the sound of the car, in the case of Navajo by literally imitating the

sound (*chid*, the sound a car makes + *i*, a type of nominalizer meaning, ‘the one which’ (Young & Morgan 1980)), while Lushootseed describes the sound. Thompson (2e) borrowed the word ‘car’ from English (dropping the foreign rhotic).

In these examples it is evident that languages do not always opt for the same method when creating a new lexical item. Koyukon, Gwich’in, and Navajo are all Athapaskan languages, yet Koyukon and Gwich’in use descriptive phrases, while Navajo made up an entirely new word based on onomatopoeia. Areal influences seem to have more of an effect, since Koyukon and Gwich’in use the same strategy (though not the same feature selection), as do Blackfoot and Kootenai, which are unrelated.

The words for ‘telephone’ in table 3 display less variety in feature selection than seen in table 2, though a thorough investigation into why this should be the case is beyond the scope of this paper.

**Table 3. Words for ‘telephone’.**

	Language	Word	Meaning	Strategy
a)	Western Apache (Ath)	bésh bití’yá’ití’í	what one talks with	N
b)	Navajo (Ath)	béésh bee hane’é	lit. ‘tool with there is conversation’	N
c)	Gwich’in (Ath)	t’yah vizhit tr’igijikhii	lit. ‘line in it talking’	N
d)	Tlingit (N-D)	atóodei yóo xh’aduwatángi át	that thing you talk into	N
e)	Lushootseed (S)	səx <sup>w</sup> ǰ <sup>w</sup> udǰ <sup>w</sup> ud	device for conversing	N
f)	Thompson (S)	cənxíc	to ring, to call s.o. by telephone	E
g)	Okanagan (S)	tq <sup>w</sup> lq <sup>w</sup> ltiws	from <i>q<sup>w</sup>l</i> , ‘talk’	N
h)	Blackfoot (Alg)	iihtáipii’poyo’p	what one talks afar with	N
i)	Kootenai (iso)	kuqunał cxanam	? (possibly related to <i>qunał</i> , ‘go over to do’)	N

In this case it is evident that most languages have converged on one aspect of the new concept: talking. All the Na-Dene languages (3a-d), as well as at least half of the other languages listed, make the verb for ‘talk’ the root from which the neologism is derived. In the Navajo (3b) example, the word *béésh* means ‘tool’<sup>8</sup>, the word *bee* is something of

an instrumental marker often translated ‘with it,’ and *hane’é* means ‘talking takes place’ or ‘there is conversation.’ While the meaning of the Kootenai example (3h) is not apparent, it is clearly a descriptive phrase, since it is multiple words and begins with the nominalizer *k*. Thus out of examples from nine languages, eight are neologisms, with Thompson (3e) being the only outlier by using a semantic extension of the verb ‘to ring.’ This shows a marked difference from the terms for the word for ‘cat,’ shown in table 4.

**Table 4. Words for ‘cat’.**

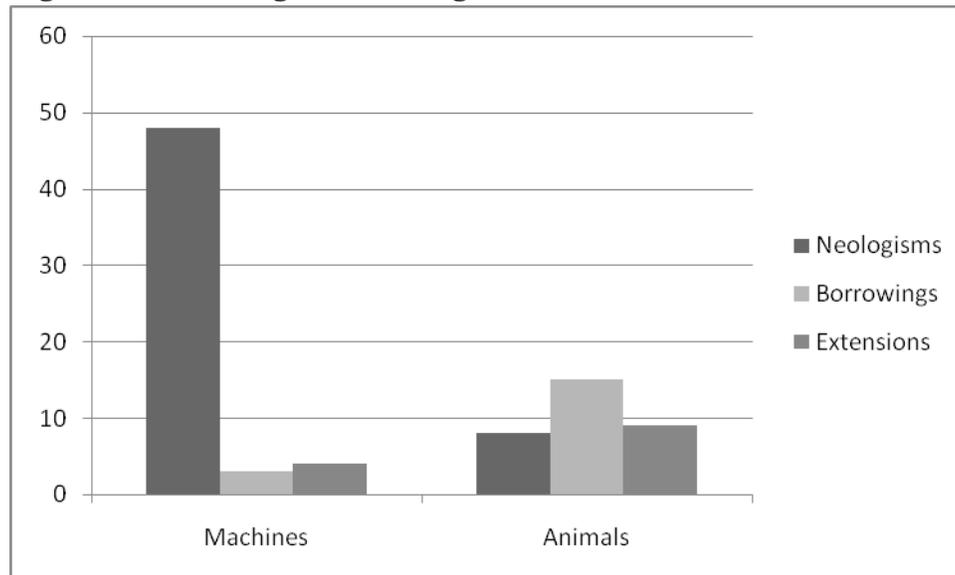
	<b>Language</b>	<b>Word</b>	<b>Meaning</b>	<b>Strategy</b>
a)	Tlingit (N-D)	dóosh	loan from Chinook Jargon	B
b)	Western Apache (Ath)	gídí	loan from English	B
c)	Navajo (Ath)	mósí, másí	loan from Spanish	B
d)	Gwich’in (Ath)	niinjii zhuu	young lynx	N
e)	Koyukon (Ath)	k’oots’eege	the one that whines at s.t.	N
f)	Lushootseed (S)	píšpiš	loan from English (?)	B
g)	Thompson (S)	pós(i), pús	loan from English	B
h)	Okanagan (S)	pus	loan from English	B
i)	Squamish (S)	push	loan from English	B
j)	Blackfoot (Alg)	ohpoos, poos	loan from English	B
k)	Kootenai (iso)	pus	loan from English	B

Several things need to be noted about the forms in table 4. The Tlingit word *dóosh* (4a) is indeed from Chinook Jargon, but since the borrowed word was *púsh* or *púshpúsh* (also *púspus* in Southern Oregon), this is clearly from the same English loan as many of the other languages surveyed. It should also be noted that Tlingit borrowings often have opaque phonological changes created by the lack of labials in Tlingit. A comparison of borrowings in Tlingit with the original source word shows that labials are often mapped onto labialized velars, but this can be unpredictable. See Crippen (2007) for more discussion of phonological mapping in Tlingit loanwords. The Gwich’in example given (4d) is from Teetl’it Gwich’in, one of two dialects which have significant lexical differences. The form given in (4d) is the Teetl’it version. In the Gwichyah dialect, the term is similar (*niinjii zheu*), but *niinjii* has been extended to mean ‘cat,’ so that the term for cat actually means ‘young cat’ instead of ‘young lynx.’ This can probably be traced through an original word or phrase which meant ‘young lynx,’ after

which the simple term *niinjii* came to mean ‘cat’ by itself. Finally, the high level of convergence among the Salish and Algonquian languages (4f-j), as well as Kootenai (4k), Chinook Jargon, and Tlingit (4a), needs more investigation (though see Denzer-King (2008)). They may all date back to a loan from Dutch *pooschees* (Mithun 1999), but in this paper I will follow the widespread assumption that the loan is directly from English ‘puss’ or ‘pussy.’

Compared to the words for ‘telephone’ in table 3, it is obvious that the strategies employed for ‘cat’ are completely different. Out of eleven languages, nine borrowed the word and only two used neologisms (for discussion on whether ‘young lynx’ should be considered a neologism or extension, see 5.1 below). In addition, Koyukon does have a borrowed form of cat (*gusge*, from Russian *kóshka*), though it is not the first entry listed in Jetté’s dictionary. This preference for neologisms describing machines or mechanical inventions/tools and borrowings or extensions for animals and other natural items seems to hold for almost the entire sample surveyed for this paper. Figure 1 shows the number of word coinings present in the entire cross-linguistic sample for neologisms, borrowings, and semantic extensions.

**Figure 1. Cross-linguistic Strategies for Mechanical Inventions vs. Animals.**



Clearly neologisms are favored for mechanical inventions, with forty-five instances of neologisms for such terms, compared to only four instances of semantic extension and three of borrowing. Items included in this category include telephone, airplane, car, screwdriver, gun, radio, computer, television, etc. In the animals category, the strategies were more evenly applied, but borrowings seem to be favored (15,

compared to 9 semantic extensions and 8 neologisms). While the preference for borrowings in the animals category is not statistically significant because of the small sample size, a chi-square test shows that the preference for neologisms among mechanical inventions/tools is ( $p=0.0000001$ ), as is the difference of preferences between the two categories ( $p=0.00000003$ ).

**5. Conclusions.** The major purpose of this study was not merely to look at individual forms in individual languages, but to analyze possible cross-linguistic trends or tendencies. By using a statistically valid sample size, it was possible to demonstrate that there is a real difference between strategies employed for naming man-made objects versus natural objects.

**5.1. Summary.** Languages use three main strategies for coining new terms: neologisms (descriptive words or phrases), borrowings (usually with phonological adaptation), and semantic extension (in which an older word is assigned to a newer concept, sometimes still retaining the older meaning, sometimes replacing it). Calques are not covered in this paper because they are rare and generally fall under the purview of neologisms. This paper shows that languages use different strategies for coining new words, even for the same concepts. Certain categories of lexical innovation show statistically significant differences in choice of strategies, viz., languages appear to distinguish between man-made objects and natural objects when coining new words. A summary of which strategies are used most often by each language is shown in table 5.

**Table 5. Strategies Presented by Language.**

Language	Neologisms	Borrowings	Extensions	Total
Squamish <sup>9</sup>	2	3	0	5
Lushootseed	6	2	2	10
Okanagan	17	5	4	26
Thompson	14	6	10	30
Tlingit	7	26 <sup>10</sup>	3	36
Gwich'in	15	0	3	18
Koyukon	8	1	2	11
Navajo	22	2	4	28
Western Apache	14	2	5	21
Blackfoot	41	2	4	47
Kootenai	15	2	2	19
Luiŕeño	8	2	3	13
Umatilla	8	0	3	11
<b>Total</b>	<b>171</b>	<b>51</b>	<b>43</b>	<b>267</b>

Table 5 shows that indigenous languages more often opt for creating new descriptive phrases rather than borrowing from another language or using an older word and ascribing new properties to it ( $p=5 \times 10^{-12}$ ). This is not to say that languages *only* use this strategy; almost every language surveyed used all three strategies even in the small samples inspected here. However, it does show that, possibly because of the model set by earlier multimorphemic noun phrases, indigenous languages of North America utilize productive grammatical processes to coin new words rather than seeking language-external sources. The general trend seems to be to prefer neologisms to both extensions and borrowings ( $p=.0008$ ). Tlingit must be left out of any statistical analysis because the sample I used was so heavily skewed towards borrowings. In the future, I will hopefully be able to use a more balanced sample of Tlingit data. A project of this type obviously benefits from massive amounts of data, so another direction for future research will be to include more languages from more and varied families in order to determine whether the trends analyzed here hold throughout the entire North American continent. It would also be a worthwhile pursuit to look into indigenous languages in other areas of the world to see if the same patterns exist elsewhere.

**5.2. Implications.** The use of neologisms rather than borrowings has important implications for native language revitalization. Borrowings cannot be said to degrade language in any way, but when borrowings come from a culturally dominant language (like English in the United States), younger speakers may feel pressure to simply use the dominant language instead of just borrowing words from it. The continued use of neologisms is thus a vital part of language preservation, because if people are going to continue to use a language, they must be able to say what they want to say. A language with no word for “cell phone” or “computer” is less likely to be used by younger generations than one which innovates. As Grenoble & Whaley (2006) put it, “creating a language revitalization program frequently involves updating the lexicon of a local language to meet the demands of the domains in which it will be used” (p. 181). Neologisms are also an important part of continued cultural traditions, since modern neologisms, especially those which are created by a language advisory board or committee, are specifically constructed to reflect certain social or cultural views on the object being named, e.g., Umatilla *shapa’ayayshwit’awas*, ‘television,’ literally, ‘thing that makes one stupid’ (see Hinton & Hale 2001 for more). In addition, “an active awareness and open dialogue about lexical innovation proposals and guiding ideologies facilitates community cohesion in language reclamation efforts” (Leonard 2008). The construction of language-internal neologisms can reverse the devaluation that may occur with extensive borrowing.

**5.3. Difficulties.** To compare strategies for lexical innovation it is of course necessary to first lump individual tokens into broader categories, i.e., to look at a given

word and label it as either a neologism, a borrowing, or an instance of semantic extension. In many cases this is straightforward, as with Thompson *ka(h)*, ‘car.’ This is clearly a borrowing, and in no way resembles either of the other two categories. Likewise, Blackfoot *innóísttoani*, ‘American,’ literally ‘have long knife,’ is clearly a neologism; it is a descriptive polymorphemic word which is not a borrowing or extension. However, not all cases are so clear-cut. An example of this is the common reference to newly encountered items as “white man’s \_\_\_\_\_,” filling in the blank with whatever familiar item is most similar. This is a trend among Thompson fruit names, e.g., *q<sup>w</sup>ʔéps e sémeʔ*, ‘apple,’ literally ‘white man’s crabapple;’ *skəz’kəz’s e sémeʔ*, ‘banana,’ literally ‘white man’s prickly pear fruit.’ Should these ‘white man’s’ items be counted as neologisms or semantic extensions? Clearly the original word is being “extended” in some way, because it is being applied to a novel object. However, for the purposes of trend analysis in this paper they were counted as neologisms, with semantic extensions being reserved for those words which are simply used as they are to describe new concepts, without being modified in any way to signify the extension of meaning.

In addition, the use of dictionary entries can be problematic because they do not always reflect community-wide usage and connotation. For instance, the word *pájaro* means ‘parrot’ in Spain, but in South America has become a vulgar slang term for male genitalia. While most extensive Spanish dictionaries will note this, indigenous languages frequently do not have extensive dictionaries, and many have only word lists. Because of this, using dictionaries of indigenous languages can be problematic because the words included therein can be reflective of a single person’s idiolect. Though it is worth noting, this problem is not necessarily a serious one for a project such as this, since neologisms often originate with a single person, and afterwards catch on in the community. Thomas Morningowl (p.c.) noted that as a translator for religious ceremonies (from English into Umatilla Sahaptin), he often makes up words as he is translating, and in many cases this is how neologisms catch on in the community.

## Appendix

Due to space limitations, preclude a complete set a data analyzed in this study cannot be included. Below are sample data sets from Koyukon and Thompson.

Koyukon, from Jetté & Jones (2000)

book	neelts'aanonolkkuge	lit. 'that which falls apart repeatedly'
car	kk'o'eelbaalee	lit. 'that which rolls around'
cat	k'oots'eege	lit. 'the one that whines at sth.'
gun	deltudle	lit. 'the thing which makes a loud noise'
necktie	denaalaan hedeeloye	lit. 'that which rests against the neck'
paper	nedenledege	orig. 'strike with arrow'
pencil	kelaandas	loan from Russian <i>karandásh</i> , a brand of pencil
radio	beyekk'e kk'ena k'etelaaye	lit. 'that through which words are sent'
road	tene	'trail, path, road', from <i>ten</i> , 'trail'
screwdriver	baahaa noch'enedetugge	lit. 'that with which sth. is unscrewed'
table	betleekk'e k'edone	lit. 'on which people eat'

Thompson, from Thompson & Thompson (1996)

airplane	nxʔiyxtn	‘it rises of its own accord’
apple	ʔépls	borrowing from English
banana	skəz’ kəz’s e sémeʔ	lit. ‘white person’s prickly pear’
book	sc’óq <sup>w</sup>	from c’óq <sup>w</sup> , ‘mark/draw/write’
car	ka(h)	loan from English
cat	pós(i), pús	English loan
chicken	cíkn	borrowing from English
compass	ncúlmn	from cúł, ‘indicate, point out’
English	səmeʔcín	from sémeʔ, ‘white person’
gun	ck <sup>w</sup> ín’ek	from ck <sup>w</sup> , ‘pull’
horse	nc’esqáxaʔ	from qáxaʔ, ‘dog, horse’
hotel	n <sup>w</sup> oy’téłx <sup>w</sup>	from <sup>w</sup> oy’t, ‘sleep’
paper	sc’óq <sup>w</sup>	from c’óq <sup>w</sup> , ‘mark/draw/write’
potato	pəték	borrowing from English
radio	nk’éy’ntn	from k’éy’nm, ‘listen’
road	xwél, xəwél	‘trail, road’, from xw, ‘trail’
ship	sc’əqʔéwł	‘boat, canoe, ship’
silver	snuyehéy’st	lit. ‘beaver stone’
table	típəł ~ típəl	loan from English
telephone	çənxíc	from çəñ, ‘ring’
Thursday	smusésq’t	lit. ‘fourth day’
tiger	ʔesə’səp tək spzúʔ	lit. ‘animal with vertical stripes’

## Notes

<sup>1</sup>For specific language data and depth of knowledge about how neologisms are formed I am indebted to James Crippen (Tlingit), Roby Littlefield (Tlingit), and Thomas Morningowl (Umatilla Sahaptin). This paper also benefitted greatly from useful comments and discussion from Leora Bar-el, Josh Birchall, Mizuki Miyashita, Meredith Ward, and Becky Wood, as well as the audience at WAIL 2008, though this does not necessarily mean that they agree with all the claims presented here. Any errors, in data or analysis, are of course my own.

<sup>2</sup> North American languages, like most languages, often have exonyms at variance with what speakers call their language. The language names I use in this paper are taken from the dictionaries for each language.

<sup>3</sup> Na-Dene is often used as a broader term to include not only Tlingit, Eyak, and the Athapaskan languages, but Haida as well (Gordon 2005). This is a controversial claim and one which is not assumed here; I use Na-Dene for Tlingit-Eyak-Athapaskan.

<sup>4</sup>Baldwin (1994) adds a fourth category, calques (loan translations), but since these play a minor role in the lexical innovation discussed here, I do not comment on this category.

<sup>5</sup>There is also the case of nonsense words which are framed only by a language's phonotactic rules as opposed to any semantic reference (e.g., 'blurb'). Since I have come across no such tokens in the indigenous languages of North America, this type of lexical innovation is not discussed here.

<sup>6</sup>Carmen Jany (p.c.) notes that *karandásh* is not actually the basic word for pencil in Russian, but refers to a specific brand, similar to the use of "Kleenex" in American English for any brand of facial tissue.

<sup>7</sup>All words are presented in the orthography of the dictionary from which they were taken, with the exception of Kootenai, in which two rare symbols have been substituted with more Unicode-friendly counterparts: the Kootenai double-barred 'l' has been replaced with the more common single-barred 'l', and the Kootenai barred 'c' has been replaced with a simple 'c'. Orthographies which differ from standard Americanist orthography are as follows: ' after a vowel represents a glottal stop, vowel doubling indicates double length, Koyukon 'kk' is Americanist 'q', Gwich'in 'kh', Tlingit 'xh', and Kootenai 'x' are 'ǰ', Gwich'in 'zh' is 'ž', Navajo 'ch' is 'č', Blackfoot 'h' is 'x', and Tlingit and Squamish 'sh' are 'š'.

<sup>8</sup>This word has a fairly high functional load in Apachean languages, especially in neologisms, and can also mean 'metal/flint,' but when used in neologisms it is often the 'tool' meaning which is used in translation. The same Apachean reflex is seen in Western Apache *bésh*.

<sup>9</sup>While Thompson, Tlingit, and Blackfoot were the only languages with a large number a words investigated (n>30), most other languages had enough data to draw some tentative conclusions. The Squamish sample size, however, was so small that no conclusions should be drawn about this language.

<sup>10</sup>A significant portion of the Tlingit data comes from a list provided by James Crippen (p.c.) of Tlingit borrowings from Chinook Jargon. Thus, while Tlingit does seem to be

more prone to borrowing than several of the other languages surveyed, the data is heavily skewed towards borrowings, and is not a balanced sample.

## References

- Baldwin, Stuart J. 1994. 'Blackfoot Neologisms.' *International Journal of American Linguistics* 60:69-72.
- Bates, Dawn, Thom Hess, and Vi Hilbert. 1994. *Lushootseed Dictionary*, ed. Dawn Bates. Seattle: University of Washington Press.
- Bray, Dorothy, ed. 1998. *Western Apache-English Dictionary*. Tempe: Bilingual Press.
- Bright, William. 1968. *A Luiseño Dictionary*. (University of California Publications in Linguistics 51.) Berkeley: University of California Press.
- Costa, David J. 1991. 'The Historical Phonology of Miami-Illinois Consonants.' *International Journal of American Linguistics* 57:365-393.
- Crippen, James. 2007. "Nontraditional Placenames in Tlingit." MS.
- Denzer-King, Ryan. 2008. 'Diffusion of a Word for "cat" in Indigenous Languages of North America.' Paper to be presented at the 8th High Desert Linguistics Conference, Albuquerque.
- Frantz, Donald G. 1995. *Blackfoot Dictionary*. Toronto: University of Toronto Press.
- Gordon, Raymond G., Jr. (ed.). 2005. *Ethnologue: Languages of the World*, 15<sup>th</sup> ed. Dallas, Tex.: SIL International. Online version: <http://www.ethnologue.com/>.
- Grenoble, Lenore A. and Lindsay J. Whaley. 2006. *Saving Languages*. Cambridge: Cambridge University Press.
- Gwich'in Language Centre. 2005. *Gwich'in Language Dictionary*. Ft. McPherson, NT: Gwich'in Language Centre / Gwich'in Social & Cultural Institute.
- Hinton, Leanne and Ken Hale, ed. 2001. *The Green Book of Language Revitalization in Practice*. San Diego: Academic Press.
- Jetté, Jules and Eliza Jones. 2000. *Koyukon Athabaskan Dictionary*. Fairbanks, AK: Alaska Native Language Center.
- Kootenai Culture Committee. 1999. *Kootenai Dictionary*. Elmo, MT: Kootenai Culture Committee, Confederated Salish and Kootenai Tribes.
- Leonard, Wesley Y. 2008. 'Lexical Innovation in Miami.' Paper presented at the SSILA 2008 Annual Meeting, Chicago.
- Mattina, Anthony. 1987. *Colville Okanagan Dictionary*. (University of Montana Occasional Papers in Linguistics 5.) Missoula, MT: University of Montana Press.
- Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge: University of Cambridge Press.

- Thompson, Laurence C. and M. Terry Thompson. 1996. *Thompson River Salish Dictionary: neʔkepmxcín*. (University of Montana Occasional Papers in Linguistics 12.) Missoula, MT: University of Montana Press.
- Young, Robert W. and William Morgan. 1980. *The Navajo Language*. Albuquerque: University of New Mexico Press.

## Relativization versus Nominalization Strategies in Chimariko

Carmen Jany

California State University, San Bernardino

### 1. Introduction

Are relative clauses formally indistinct from clausal nominalization in certain languages? This has been argued for a number of languages, such as Mojave, Diegueño, Luiseño, Wappo, and Quechua (Comrie and Thompsen 1985). This paper examines relative clauses in Chimariko, an extinct language of Northern California, and demonstrates that Chimariko shows no formal distinction between relative clauses and clausal nominalizations, the same as Mojave, Diegueño, Luiseño, Wappo, and Quechua.

Nominalization refers to 'turning something into a noun' (Comrie & Thompson 1995). It is a derivational process that creates nouns from lexical verbs and adjectives. The resulting nouns become the head nouns in a noun phrase. Clausal nominalization is a process 'by which a prototypical verbal clause [...] is converted into a noun phrase' (Givón 1990:498). According to Givón (1990:498) 'a verbal clause is nominalized most commonly when it occupies a prototypical nominal position (or 'function) [...] within another clause'. Generally, this process is accompanied by structural adjustments, such as the absence of tense, aspect, and modal markers and modifications in the case-marking, among others. This work compares relativization strategies to clausal nominalizations. Clausal nominalizations are not prototypical nominalizations in that they do not involve the derivation of a noun from a verb given that the nominalized constituent represents an entire clause. Furthermore, the verb in the nominalized clause often maintains some of its verbal properties. For example, arguments and adjuncts have the same properties as in a non-nominalized clause.

In contrast to clausal nominalization, relative clauses are clauses which restrict the meaning of a noun. They are subordinate clauses embedded inside noun phrases functioning as noun modifiers (Givón 1990:645). Keenan (1985) identifies four characteristics included in relative clause constructions: (1) they are sentence-like, (2) they consist of a head noun which is present or which can be inferred and a relative clause, (3) they have a total of two predicates, and (4) they describe or delimit an argument. Relative clauses in Chimariko include these four characteristics (see section 3). However, the verb in these constructions occurs with a special suffix *-rop/-rot/-lop/-lot* marking dependency. This suffix could also be interpreted as a clausal nominalizer (see section 4).

Relative clauses can be grouped together following three typological parameters: (1) the position of the relative clause with respect to the head noun, (2) the mode of expression of the relativized noun phrase, and (3) which grammatical element can be relativized (Payne 1997). With regard to (1), i.e. the position of the relative clause with respect to the head noun, four possibilities have been identified: relative clauses are either pronominal (the relative clause occurs before the head), postnominal (the relative clause occurs after head), internally headed (the head occurs within the

relative clause, or headless (the head is inferred). Only the latter two possibilities occur in Chimariko (see section 4). The second typological parameter refers to identifying the role of the referent of the head noun within the relative clause. It has also been called the ‘case recoverability strategy’ (Payne 1997). The third typological parameter refers to the relativization hierarchy: subject > direct object > indirect object > oblique (Keenan and Comrie 1977). If a position on the hierarchy can be relativized, then all positions to the left can also be relativized. Given the particular case of Chimariko argument structure which is based on agents and patients and on a person hierarchy, the latter two typological parameters (2) and (3) are not examined here. Neither the agent-patient distinction nor the person hierarchy is reflected in third persons in Chimariko. There is only one marker for third persons, a prefixed or suffixed *h* (see section 3).

Section 2 provides general information about Chimariko and explains the nature of the data. Section 3 treats relativization strategies in Chimariko, while section 4 summarizes nominalization in Chimariko. Section 5 examines relativization in Diegueño, Quechua, and Wappo comparing it to Chimariko.

## 2. The Language and Data

Chimariko is a now extinct Northern California language. It was once spoken in a few small villages along the Trinity River and parts of the New River and South Fork River. The last speaker probably died in the 1940s.

Published and unpublished materials on the Chimariko language and culture are limited to a brief grammatical sketch (Dixon 1910), a few articles, and handwritten notes from data collection sessions by Stephen Powers (1875), Edward Sapir (1927), C. Hart Merriam (1920-1921) and John Peabody Harrington (1921, 1927, 1928), among others. The main source of data for this work comes from 3500 pages of handwritten field notes collected by John Peabody Harrington in the 1920s and the notes of George Grekoff. Harrington collected elicited sentences, vocabulary, and oral narratives from several consultants. A sample page is included in the appendix. Grekoff examined Harrington’s extensive corpus leaving numerous notes and some analyses which have proven useful.

Typologically, Chimariko is a head-marking language. Core arguments are obligatorily marked on the verb and possession is marked on the possessed. Case-marking occurs only with instruments and companions while other nominal syntactic relations are unmarked. Argument structure is based on agents and patients and on a person hierarchy whereby only one argument or overtly expressed on the predicate. Chimariko is a synthetic to polysynthetic language with mainly suffixes. However, personal pronouns and possessors are either prefixed or suffixed. With regard to word order, Chimariko seems to be verb final, though the limited amount and kind of data does not yield a clear picture. No apparent preference or restrictions have been identified for the order of nominal elements within a noun phrase.

## 3. Relativization Strategies in Chimariko

There are two relativization strategies in Chimariko: (1) internally headed relative clauses and (2) headless relative clauses. There are no relativizers or subordinating conjunctions linking a relative clause to its head. Heads occur all within the relative clause as in examples 1-5, and the predicate within the relative clause occurs with a special suffix *-rop* /*-rot* /*-lop* /*-lot* marking dependency. In the examples below, relative clauses are enclosed in brackets, heads are boldfaced, and the special verb form *-rop* /*-rot* /*-lop* /*-lot* is underlined.

1. 'Hopping Game' (Grekoﬀ 004.008)  
*himantamorop map'un, hi'amta*  
 [h-iman-tamo-rop **map'un**] h-i'am-ta  
 3-fall-DIR-DEP that.one 3-beat-DER  
 'Those fellows that went down got beaten.'
  
2. Harrington 20-1097  
*map'un hokote'rot yeči' ?imi'nan*  
 [map'un h-oko-te'-rot] y-eči' ?i-mi'n-an  
 that.one 3-tattoo-DER-DEP 1SG.A-buy 1SG-want-ASP  
 'I want to buy that engraved one.'
  
3. Harrington 20-1103  
*mo'a p<sup>h</sup>uncar huwatkurop p<sup>h</sup>a'yinip*  
 [mo'a p<sup>h</sup>uncar h-uwa-tku-rop] p<sup>h</sup>a'yi-nip  
 yesterday woman 3-go-DIR-DEP thus.say-PST  
 'That woman who came yesterday told me.'
  
4. Grekoﬀ 020-009  
*nač<sup>h</sup>ot yak'orop p<sup>h</sup>a'asu, hik'ot*  
 [nač<sup>h</sup>ot ya-k'o-rop p<sup>h</sup>a'asu] h-ik'o-t  
 1PL 1PL-talk-DEP that.kind 3-talk-ASP  
 'What we talk, she talked.'
  
5. Grekoﬀ 012.014  
*č<sup>h</sup>e'new yewurop hačmukč<sup>h</sup>a č<sup>h</sup>awun*  
 [č<sup>h</sup>e'new y-ewu-rop] hačmukč<sup>h</sup>a č<sup>h</sup>-awu-n  
 bread 1SG.A-give-DEP axe 1SG.P-give-ASP  
 'For the bread I gave him, he gave me an axe.'

As can be seen in 1-5, there are no relativizers and all heads are within the relative clause, either as a head noun as in 3 and 5, or as a relative pronoun *map'un* 'that one' or *p<sup>h</sup>a'asu* 'that kind' as in 1, 2, and 4. In example 1, *map'un* 'that one' is considered to be part of the relative clause rather than the main clause as it occurs within the same

intonation unit. Intonation units are signalled in the data by commas. In 1-5, the relative clauses always precede the main clause. This is not surprising as Chimariko has predominantly verb-final word order. With regard to the order within the relative clause, the heads either precede the dependent predicate, as in 2, 3, and 5, or they follow it, as in 1 and 4. In addition to the special suffix marking dependency, the dependent predicate occurs with pronominal affixes, but it lack tense, aspect, or modal suffixes which are obligatory in independent clauses. Potential restrictions on what can be relativized in Chimariko is unclear. In the available data there are examples of relativized arguments that serve as actors or undergoers in the relative clause.

In general, only one argument is marked on the predicate following a hierarchy whereby speech act participants are favored over third persons. In addition, first persons show a distinction for agents and patients and first and second persons distinguish number. Third person markers are always *h*, regardless of number or semantic role.

The second relativization strategy used in Chimariko are headless relative clauses. As with internally headed relative clauses, the verb form in the headless relative clause occurs with a suffix *-rop* /*-rot* /*-lop* /*-lot* marking dependency. This is illustrated in examples 6-8.

6. Grekoff 012.014  
*yewuxan* *ʔahatew hexačilop šičelaʔi*  
*y-ewu-xan* *ʔahatew [h-exači-lop šičela-ʔi]*  
 1SG.A-give-FUT money 3-steal-DEP dog-POSS  
 ‘I’ll give you money for the stealing by my dog.’  
 (Literally: ‘I give you money for what my dog stole’)
7. Grekoff 020.009  
*hikʔomutarop hitxahta*  
*[h-ikʔo-muta-rop] h-itxah-ta*  
 3-talk-?-DEP 3-stop-ASP  
 ‘He stopped talking’ (Literally: ‘What he was uttering, he stopped it’)
8. Grekoff 020.009  
*šitoita hikʔorop hekʔoʔnačaxat*  
*[šito-ita h-ikʔo-rop] h-ekʔoʔna-čaxa-t*  
 mother-POSS 3-tell-DEP 3-say-APPL-COMP-ASP  
 ‘She told her mother everything’ (‘What she told her mother, she told her all’)

In examples 6-8, there are no relativizers and there are no heads. The heads are inferred. In examples 7 and 8, the relative clauses precede the main clause. The only example where a relative clause follows a main clause is 6. As with internally headed clauses, the dependent predicate occurs with pronominal affixes, but it lack tense, aspect, and modal suffixes.

Overall, internally headed and headless relative clauses show the same structure with the only distinction of having an explicit head or not. Could these constructions be interpreted as clausal nominalizations? The next section examines this question.

#### 4. Relativization versus Nominalization in Chimariko

The predicates in Chimariko relative clauses show properties of both, nouns and verbs. They are noun-like in that (a) they lack any tense, aspect, or modal affixes and (b) they cannot form clauses by themselves. In Chimariko verbs can form clauses by themselves. They are verb-like in that (a) they occur with pronominal marking and (b) they can take arguments. In clausal nominalization the verb retains some of its verbal properties. The Chimariko verb retains some of its verbal properties in relative clauses, such as pronominal marking and the possibility of taking arguments. Furthermore, in clausal nominalization there are structural adjustments in the process. The lack of tense, aspect, and modal suffixes is often such a structural adjustment. Tense, aspect, and modal suffixes are absent in Chimariko relative clauses. Another piece of evidence for clausal nominalization is the position of the nominalized clause. A nominalized clause occupies a prototypical nominal position or function within another clause. Chimariko is predominantly verb-final and the relative or nominalized clauses occur before the main predicate with the exception of example 6, i.e. in the prototypical nominal position. What about nominal function? In clausal nominalization a verb phrase is turned into a noun phrase: VP → NP. There are no nouns or independent pronouns in some of the constructions found in Chimariko, as in example 7, and the dependent predicate shows properties of both nouns and verbs making it difficult to determine whether they are nouns or verbs. A relative clause *Srel* restricts the meaning of a noun phrase. The construction can be summarized as follows: [[*Srel*]NP] + VP. In examples 1-7 the relative clauses restrict the meaning of a head, present or inferred. This is less clear in example 8. Overall, relative clauses in Chimariko could structurally be interpreted as clausal nominalizations paralleling constructions found in Diegueño and in other languages (see 5), but functionally they are best viewed as relative clauses given that they restrict the meaning of a head.

While there is no construction representing clausal nominalization in Chimariko, other than the relative clauses, lexical nominalizations are common and are formed with the nominalizer *-ew*, as in example 9.

9. Nominalizations with the nominalizer *-ew*
- |                       |    |  |                       |
|-----------------------|----|--|-----------------------|
| <i>ama</i> ‘to eat’   | => | <i>h-am-ew</i><br>POSS-eat-NOM<br>‘food’     | (‘Woman wanders’)     |
| <i>ik’o</i> ‘to talk’ | => | <i>h-ik’-ew</i><br>POSS-talk-NOM<br>‘talker’ | (Harrington 020-1133) |

The prefix *h-* in example 9 is a possessive prefix. Given that verbal pronominal affixes and nominal possessive affixes are almost identical in shape, the pronominal affixes in relative clauses could also be interpreted as possessive affixes. The only difference in shape occurs in the first person plural forms and in some third person forms. Example 10 shows that the affixes in relative clauses are in fact pronominal and not possessive affixes as in the nominalized verbs in 9.

- 10a. (same as 4) Grekoff 020-009  
 [nač<sup>h</sup>ot *ya-k'o-rop*    *p<sup>h</sup>a'asu*]    *h-ik'o-t*  
 1PL    1PL-talk-DEP    that.kind    3-talk-ASP  
 'What we talk, she talked.'
- 10b.                    *č<sup>h</sup>a-sot*                    *h-usot*  
 1PL.POSS-eye                    3.POSS-eye  
 'our eye'                    'his eye'

In example 10a, the pronominal prefix for the first person plural *ya-* is different from the first person plural possessive prefix *č<sup>h</sup>a-* in 10b.

While Chimariko verbs take pronominal, tense, aspect, modal, and derivational affixes, nominal stems can take possessive, privative, locative, definitive, and case affixes for instruments and companions. None of these nominal affixes are found on predicates in relative clause constructions. However, the verbal suffix marking dependency *-rop/ -rot /-lop /-lot* is similar in shape to the nominal suffix marking definiteness, shown in examples 11 and 12.

11.    Definite suffix *-ot* (Harrington 020-1093)  
*šičelot č<sup>h</sup>awin, č<sup>h</sup>utpai, č<sup>h</sup>awin*  
*šičel-ot*    *č<sup>h</sup>-awi-n*                    *č<sup>h</sup>-utpa-i*                    *č<sup>h</sup>-awi-n*  
 dog-DEF 1SG.P-afraid-ASP 1SG.P-bite-MOD 1SG.P-afraid-ASP  
 'I am afraid of **the dog**, he might bite, I am afraid'.
12.    Definite suffix *-op* ('Fugitives at Burnt Ranch')  
*hek'omatta, hak<sup>h</sup>ote' č'imarop, xawiyop hak<sup>h</sup>ote'n*  
*h-ek'o-ma-tta h-ak<sup>h</sup>o-te' č'imar-op*    *xawiy-op*    *h-ak<sup>h</sup>o-te'-n*  
 3-say-?-DER 3-kill-DER person-DEF Indian-DEF 3-kill-DEF-ASP  
 'He (the boy) told (it), they killed the boy, **the people, the Indians** killed him'.

Is this similarity indicative of clausal nominalization rather than of a relative clause construction? Given that these two affixes differ in their functions and that they are only similar in shape, the answer is no. The suffix *-rop/ -rot /-lop /-lot* marking dependency in relative clauses does not mark definiteness.

The next section examines relativization and nominalization in three of the languages where the two constructions have been claimed to be the same and compares them to relative clause constructions in Chimariko.

## 5. Relativization and Nominalization in Diegueño, Quechua, and Wappo

In certain languages clausal nominalization is formally indistinct from relativization (Comrie and Thompson 1985). Three of these languages are examined below: Diegueño, Wappo, and Huallaga Quechua.

### 5.1 Diegueño

Diegueño is a language of Southern California. Gorbet (1976) examines relative clauses in Diegueño and notes that the verb in the relative clause occurs with nominal suffixes marking definiteness and case. This is illustrated in example 13.

- 13a. *[i:pac 'wu:w]-puc-c*                      *ciyaw*  
 man I.saw-DEM-SUBJ                      sing  
 'The man I saw sang'
- 13b. *i:pac 'wu:w*                                      *i:pac-puc-c*                      *ciyaw*  
 man I.saw    man-DEM-SUBJ                      sing  
 'I saw the man'                                      'The man sang'

The relativized noun *i:pac* 'man' in 13a does not change its position or case-marking when compared to 13b. The verb in the relative clause in 13a bears the definiteness and case markers, *-puc* and *-c* respectively, indicated from the function of the relativized noun in the main clause. When compared to a nominalized sentential object clause, as in example 14, the relative clause in 13a shows no structural difference. Hence, clausal nominalization and relative clause constructions are formally indistinct in Diegueño.

14. *'nʷa:c 'i:ca-s*                                      [*puy ta-'nʷ-way]-pu-∅*  
 I-SUBJ I-remember-EMPH                      there PROG-I-be-there-DEM-OBJ  
 'I remember that we were there'

The predicate in Chimariko relative clauses does not occur with any case or definiteness markers. Therefore, the structural similarity between relative clauses and clausal nominalizations is less clear in Chimariko.

### 5.2 Wappo (Li and Thompson 1978)

Wappo is a Northern California language with a rich case system and a verb-final word order. In subordinate clauses, subjects appear in the accusative case. Li and Thompson (1978) identify three relativization strategies in Wappo: (a) internal head constructions, (b) a pre-posing strategy and (c) a post-posing strategy. Only the first one is examined here. The internal head construction is shown in example 15.

15.   ʔah    [ʔi    k'ew-ø    nawta]-ø   hak'se'  
       I       me    man-ACC   saw   ACC   like  
       'I like the man I saw'

The relative clause in 15 occurs in the position in which a simple noun in that function would occur, here the O in SOV. Furthermore, the relative clause in 15 is marked with the appropriate case marker, here - ø for the accusative, and it is clearly subordinate since the subject ʔi 'I' occurs in the accusative. A nominalized clause would show the same structure. In contrast, in Chimariko there is no case marking on the predicates in relative clauses. Therefore, there is less evidence for a structural similarity between relative clauses and nominalizations in Chimariko.

### 5.3 Huallaga Quechua (Weber 1983)

Huallaga Quechua is spoken in Peru. It is an SOV language, but it does not have a rigid SOV word order (Weber 1983). As with Wappo, Huallaga Quechua has a rich case system. According to Weber (1983), there is insufficient evidence for distinguishing nouns and adjectives as distinct lexical categories in the language. Weber (1983) argues that relativizations are formally distinct from nominalizations only to the extent that nouns are distinguished from adjectives. As a result, there is also insufficient evidence that relativized clauses and nominalizations are distinct syntactic classes. Examples 16 and 17 show that internally headed relative clauses and nominalized clauses have the same structure in the language.

16.   [*marka-man*   *chaya-sha:-chaw*]   *hamashkaa*  
       town-GOAL   arrive-SUB.1-LOC   I.rested  
       'I rested in the town to which I arrived'
17.   *qonqashkaa*   *away-shaa-ta*  
       I.forgot       go-SUB.1-ACC  
       'I forgot that I had gone'

The predicate in the internally headed relative clause in 16 is marked with a locative suffix *-chaw*, i.e. a nominal suffix. Both the relative clause in 16 and nominalized clause in 17 are marked with the same substantivizing subordinator (glossed together with the pronominal affix). Hence, there is no structural difference between the two constructions other than the presence of the head noun in 16.

Chimariko is different from Huallaga Quechua in that there are no locative markers on predicates in relative clauses, and in that adjectives are morphosyntactically distinct from nouns.

## 6. Summary and Conclusions

As shown in section 5, in Diegueño, Wappo, and Huallaga Quechua relative clauses and nominalized clauses are structurally the same. There are no relativizers in these constructions. The predicates in both the relative clauses and the nominalized clauses occur with nominal markers, such as case or definiteness affixes, and the clauses occupy a prototypical nominal position in the main clause. The only difference between the two constructions is that relative clauses restrict an internal or inferred head.

Are relative clauses and clausal nominalizations structurally indistinct in Chimariko, the same as in these languages? Like in Diegueño, Wappo, and Huallaga Quechua, there are no relativizers in Chimariko relative clauses, and the relative clauses occupy prototypical nominal positions in the main clause. However, unlike in Diegueño, Wappo, and Huallaga Quechua there are no nominal markers on predicates in relative clauses. While Chimariko has no rich case system, other nominal markers, such as the definite *-op*, are common. The suffix marking dependency in relative clauses, *-rop /-ro t/-lop/-lot* is similar in shape to the nominal definite marker *-op*, but it is clearly different in its function. The suffix marking dependency could also be interpreted as a clausal nominalizer, similar to Huallaga Quechua, but semantically the constructions found in Chimariko represent relative clauses, because they are restrict a head, present or inferred.

To conclude, relative clauses and clausal nominalizations may be formally indistinct in Chimariko. Semantically, however, these constructions represent relative clauses. Given that they are restrictive, i.e. they identify the respective referents, rather than being nominalizations, they are better interpreted as relative clauses. Examples with relativized instruments or companions, i.e. arguments where case marking occurs, or negative relative clauses may offer an additional piece of evidence for the claim that relative clauses and clausal nominalizations are formally indistinct in Chimariko. However, given the nature of the data, no such examples have been identified.

## 7. Bibliography

- Comrie, Bernard. 1998. 'Rethinking the Typology of Relative Clauses'. *Language Design 1*: 59-86.
- Comrie, Bernard and Sandra A. Thompson. 1985. 'Lexical Nominalization'. In: Timothy Shopen ed. *Language Typology and Syntactic Description, Vol. III: Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press. 349-398.
- Givón, T. 1990. *Syntax: A Functional-Typological Introduction*, Vol. II. Amsterdam: Benjamins.
- Gorbet, Larry Paul. 1976. *A Grammar of Diegueño*. New York: Garland.
- Grekoff, George. 1950-1999. Unpublished notes on various topics. Survey of California and other Indian Languages. University of California, Berkeley.
- Harrington, John Peabody. 1921-1928. Fieldnotes. Microfilm reels: Northern California 020-024.

- Jany, Carmen. To appear. *Chimariko Grammar: Areal and Typological Perspective*. UC Publications in Linguistics. University of California Press.
- Keenan, Edward. 1985. 'Relative Clauses'. In: Timothy Shopen ed. *Language Typology and Syntactic Description, Vol. II: Complex Constructions*. Cambridge: Cambridge University Press. 141-170.
- Li, Charles N. and Sandra A Thompson. 1978. 'Relativization Strategies in Wappo'. *Berkeley Linguistics Society* 4: 116-113.
- Payne, Thomas E. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press.
- Weber, David John. 1983. *Relativization and Nominalized Clauses in Huallaga (Huanuco) Quechua*. UC Publication in Linguistics, Vol. 103. University of California Press.

## LIST OF GLOSSES

A	Agent	P	Patient
ASP	Aspect	PTCP	Participle
DEP	Dependent	PST	Past tense
DER	Derivational	PL	Plural
DIR	Directional	POSS	Possessive
FUT	Future	PROG	Progressive
MOD	Modal	Q	Interrogative
NEG	Negative	SG	Singular
OP	Discourse-pragmatic marker	SUB	Substantivizing subordinator
		TERM	Terminative

## Appendix I: Sources for the Narratives Used in the Examples

<u>Narrative</u>	<u>Source</u>
Fugitives at Burnt Ranch	Harrington 021-0007 <sup>1</sup>
Dailey Chased by the Bull	Grekoff 004.008 <sup>2</sup>
Hopping Game	Grekoff 004.008
Crawfish	Grekoff 004.008

<sup>1</sup> These numbers refer to the microfilm reels with Harrington's data. The first three digits indicate the microfilm reel and the following number represents the frame number on the reel. The reels are numbered 020-024 of the Northern California collection. The number of frames on one reel varies.

<sup>2</sup> The Grekoff Collection is housed at the Survey of California and Other Indian Languages at the University of California at Berkeley. The numbers represent their cataloguing of the materials.

## Appendix II: Harrington Sample Page

Textlet on Fugitives at Aunt Ranch.  
 J'war g'tai hetraheskut  
 I three men came as fugitives  
 nwat put, hetraheskut of utamdaye  
 kimot 'u'ia asunda, of lak'o,  
 These are strangers.      of lak'o  
 hetraheskut a sunda,  
 They are running away.  
 H. Casak'ona, we of hup of Casak'ona  
 (said?) don't will them