REPORT 15

SURVEY OF CALIFORNIA AND OTHER INDIAN LANGUAGES

Structure and Contact in Languages of the Americas

John Sylak-Glassman and Justin Spence, Editors
Andrew Garrett and Leanne Hinton, Series Editors
# Table of Contents

John Sylak-Glassman and Justin Spence  
*Introduction*  
iv

---

Natalia Chousou-Polydouri and Vivian Wauters  
*Subgrouping in the Tupi-Guarani Family: A Phylogenetic Approach*  
1

---

Jessica Cleary-Kemp  
*A ‘Perfect’ Evidential: The Functions of -shka in Imbabura Quichua*  
27

---

Clara Cohen  
*Hierarchies, Subjects, and the Lack Thereof in Imbabura Quichua*  
51  
*Subordinate Clauses*

---

Iksoo Kwon  
*One -mi: An Evidential, Epistemic Modal, and Focus Marker in Imbabura Quechua*  
69

---

Ian Maddieson and Caroline L. Smith  
*The Stops of Tlingit*  
87

---

Yoram Meroz  
*The Plank Canoe of Southern California: Not a Polynesian Import, but a Local Innovation*  
103

---

Lindsey Newbold  
*Variable Affix Ordering in Kuna*  
189

---

Daisy Rosenblum  
*Passive Constructions in K’wak’wala*  
229

---

Justin Spence  
*Dialect Contact, Convergence, and Maintenance in Oregon Athabaskan*  
279

---

John Sylak-Glassman  
*Affix Ordering in Imbabura Quichua*  
311
Introduction

JOHN SYLAK-GLASSMAN and JUSTIN SPENCE
University of California, Berkeley

The editors are pleased to present Structure and Contact in Languages of the Americas as Volume 15 of the Survey Reports from the Survey of California and Other Indian Languages at the University of California, Berkeley. Since Volume 8 in 1993, the Survey Reports series has published the proceedings of important conferences on language revitalization, California languages, and languages of the Americas in general. However, the Survey Reports began as a venue for presenting various shorter pieces of descriptive work, and this volume follows in that tradition, featuring work on languages of North, Central, and South America in a variety of linguistic subfields. A number of the papers in this volume are the result of the 2009-2010 linguistic field methods class on Imbabura Quichua at UC Berkeley. These and the rest of the papers in the volume have benefited from feedback from audiences at UC Berkeley and elsewhere. The editors wish to acknowledge here the support of this broader community of Americanists whose insights and ongoing commitment to languages of the region have made the research featured in this volume possible.
Subgrouping in the Tupí-Guaraní Family: A Phylogenetic Approach*

NATALIA CHOUSOU-POLYDOURI and VIVIAN WAUTERS
University of California, Berkeley

1 Introduction

In recent years, interdisciplinary collaboration between linguists and biologists has led to new insights in comparative linguistics. By analyzing linguistic data using phylogenetic methods of species subgrouping, we can quantify statistic probabilities of various subgrouping hypotheses and enrich our understanding of the diachronic relationships between descendant languages in a family. Even as the preliminary results of this work have developed important strands of inquiry (Atkinson and Gray 2003; Rexová et al. 2006; Dunn 2009), many have criticized the method, indicating essentially that “we still do not have evidence that any of these methods [of phylogenetic analysis] is capable of accurate estimation of linguistic phylogenies” (Nichols and Warnow 2008:814; see also Heggarty 2006).

This paper presents a preliminary analysis of the Tupí-Guaraní language family using phylogenetic methods. Through use of parsimony and Bayesian analyses on a set of lexical items divided into cognates, we conclude that phylogenetic methods do produce useful and interesting results that will inform more traditional reconstruction. We also discuss the many potential confounds that are present in the current analysis and their solutions, and point to future research directions that will make further use of this data in conjunction with a “by hand” reconstruction of the Tupí-Guaraní proto-language.

1.1 Language Background

The Tupí-Guaraní language family forms one major subgroup of the Tupian stock, which is one of the largest macro families in South America (Derbyshire 1994). By the time of European invasion, the Tupian languages were spread throughout the Amazon and beyond, from the Andes to the Atlantic Ocean, mostly following the complex river systems of the regions (Lathrap 1970). Following contact, many languages were lost, and many others moved location, either through forced re-location or in the process of fleeing the settlers (for a discussion of these processes on Omagua, see Michael 2010). However, the recordable remnants of this widely spread family are visible in the dispersion of Tupí-Guaraní languages, as seen in Figure 1.

1.2 Previous Study

While South America is one of the linguistically least understood areas of the world (Dixon and Aikhenvald 1999), the Tupí stock, and particularly the Tupí-Guaraní family, has a somewhat es-

---

*Many thanks to Lev Michael, who runs the Tupí-Guaraní Comparative Project, as well as colleagues Keith Bartolomei, Zachary O’Hagan, and Michael Roberts, each of whom is responsible for collecting all the lexical items for a portion of the languages.
Figure 1: Tupí-Guaraní language map
Established tradition of description ascribed to it (Soares and Leite 1991:36). In addition to the description efforts on various daughter languages of the family, Rodrigues (1958, 1984, 2007) and Rodrigues and Cabral (2002) all focus on various aspects of reconstructing proto-Tupí-Guaraní. Broader morphosyntactic reconstruction efforts appear to be based on Lemle (1971). This reconstruction is referenced and utilized in much of Rodrigues’ work, and contributed to the influential set of reconstructions by Cheryl Jensen (1989, 1998). The latter of the two, Jensen (1998), titled “Comparative Tupí-Guaraní Morphosyntax,” is the source of most linguists’ understanding of subgrouping in the Tupí-Guaraní family. Both Jensen (1998) and Rodrigues (1984) posit only a single level of subgroupings of the Tupí-Guaraní languages; Rodrigues’ grouping, copied into Jensen (1998) is shown in Table 1, with the languages considered in this work in bold. While this is not the only subgrouping that has been proposed, the more recent Rodrigues and Cabral (2002), which included more detailed subgroups within the eight established groups, differed only minimally from its antecedent: Kayabí is considered part of group VI and not V. This is also the case in Etnolingüística (2011).

<table>
<thead>
<tr>
<th>GROUP</th>
<th>LANGUAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Old Guaraní; Mbyá Guaraní [Mbyá]; Xetá; Šandeva; Kaiwá [Kaiowá]; Paraguayan Guaraní; Guayakí; Tapieté; Chiriguano; Izoceño</td>
</tr>
<tr>
<td>II</td>
<td>Guarayú; Sirionó; Hora; [Yuki]</td>
</tr>
<tr>
<td>III</td>
<td>Tupinambá; Língua Geral Paulista; Língua Geral Amazônica (Nheengatu); Cocama [Kokama]. Cocamilla [Kokamilla]; Omagua</td>
</tr>
<tr>
<td>IV</td>
<td>Tocantins (or Trocará) Assurini [Asurini do Tocantins]; Tapirapé; Ava (Canoeiro); Tocantins Suruí (Akewere); Parakaná; Guajajara; Tembé [Tenehetara]</td>
</tr>
<tr>
<td>V</td>
<td>Kayabí; Xingu Assurini [Asurini do Xingu]; Araweté</td>
</tr>
<tr>
<td>VI</td>
<td>Parintintún; Tupí-Kwahib; Apiaká</td>
</tr>
<tr>
<td>VII</td>
<td>Kamaiurá</td>
</tr>
<tr>
<td>VIII</td>
<td>Takunyapé; Emerillon; Urubú-Kapor; Wayampi; Amanayé; Anambé; Turiwara; Guajá</td>
</tr>
</tbody>
</table>

Table 1: Jensen (1998) Tupí-Guaraní language subgrouping

1 Names in square brackets represent the names of the languages as used in this study, when different from those used in the source material. Yuki is not included in this subgrouping, but is classified in Etnolingüística (2011) as a member of Group II. Also, our analysis treats Guajajara and Tembé as dialects that can be subsumed under the title Tembé-Tenetehara, often shortened to Tembé. Similarly, Kokama and Kokamilla are dialects that will be referred to under the single name Kokama.
Despite these published reconstruction efforts, none of the previously mentioned sources have actually reconstructed proto-Tupí-Guaraní using rigorous historical linguistics methods like the Comparative Method. What seems more accurate is that Lemle (1971) reconstructed a phoneme system of proto-Tupí-Guaraní by deduction from segments in the daughter languages, and used that phoneme set to establish proto-forms of some 200 words. While this is only one in the list of reconstructions, it further appears that most subsequent work on the proto-language has used the Lemle (1971) data as its basis. While these extant reconstructions do seem (from observation) to be generally correct, the lack of precision inherent in the methods used makes current scholarship on proto-Tupí-Guaraní ultimately uninformative.

In contrast to these “reconstructions,” there have been two published works on the diachronic relationship of languages in the Tupí-Guaraní family that make conclusions based on data rather than intuition. The first of these is “More Evidence for an Internal Classification of Tupí-Guaraní Languages” by Wolf Dietrich (1990). Unlike the other reconstructions, Dietrich does not attempt to make a tree structure for the diversification from proto-Tupí-Guaraní, but rather, using 29 languages, he examines phonological and morphological characteristics of the languages, and, using a summary of pairwise combinations between languages, posits a gradient analysis of each Tupí-Guaraní language from “conservative” to “innovative.” Under this analysis, Dietrich does not propose any proto-forms but groups the languages based on their level of conservativeness into a number of low-level subgroups. These results differ from the other reconstructions in some important ways. Most notably, Kokama (Cocama) does not form a subgroup with Tupinambá. Also, the two Group 6 languages of previous reconstructions that were included in the sample do not pattern together, but rather Kayabí (Group 6) appears to be much closer to Tapirapé (a Group 4 language) and Kamaiurá (Group 7) than to Asurini do Xingu (Group 6) (Dietrich 1990:116). Notably, Dietrich freely admits that while his study is able to indicate similarities between language pairs and small groups, it does not constitute a fully developed description of the subgrouping in Tupí-Guaraní, which he considered to be “far from clear” (Dietrich 1990:116).

The second of these reconstructions is “A Historical Study of the Tupí-Guaraní Language Family” (Estudo histórico da família linguística tupí-guaraní; Mello 2000). This study actually reconstructs a number of cognate sets based on regular sound correspondences, and forms subgrouping hypotheses based on sound changes as well as cognate isoglosses; the grouping is shown in Table 2.

---

2 There is another study that reconstructs proto-Tupí-Guaraní forms based on what appear to be more principled methodologies, which is Schleicher (1998). This reconstruction does not, however, propose any subgrouping hypothesis, instead citing Dietrich (1990) as the best attempt.

3 Languages included in our current sample appear in bold and languages that will eventually be included appear in italics. For ease of comparison, the language names have been regularized from Mello’s labels to those used throughout the rest of this study.
### Subgrouping in the Tupí-Guaraní Family

<table>
<thead>
<tr>
<th>GROUP</th>
<th>SUBGROUP</th>
<th>LANGUAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>A</td>
<td><em>Mbyá, Paraguayan Guarani, Old Guarani</em></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>Chiriguano, Chané, Izoceño</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>Guayaki</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>Xetá</td>
</tr>
<tr>
<td>II</td>
<td></td>
<td>Sirionó</td>
</tr>
<tr>
<td>III</td>
<td></td>
<td>Guarayú</td>
</tr>
<tr>
<td>IV</td>
<td>A</td>
<td>Sirionó, Amundava, Urueuwauwau</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>Tenharín, Karipúna</td>
</tr>
<tr>
<td>V</td>
<td></td>
<td>Kayabí, Kamaiurá, Apiaká</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>Asuriní do Tocantins, Parakaná, Suruí</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>Tembé</td>
</tr>
<tr>
<td>VI</td>
<td>C</td>
<td>Tapirapé</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>Asuriní do Xingu</td>
</tr>
<tr>
<td>VII</td>
<td></td>
<td>Guajá, Araweté, Anambé, Aurê e Aura</td>
</tr>
<tr>
<td>VIII</td>
<td></td>
<td>Wayampí do Amapari/Jarí, Emerillon, Urubu-Kaapor</td>
</tr>
<tr>
<td>IX</td>
<td></td>
<td>Tupinambá, Kokama, Lígua Geral Amazônica</td>
</tr>
</tbody>
</table>

Table 2: Mello (2000) Tupí-Guaraní language subgrouping

While Mello (2000) does not include a number of the languages in our current study in his subgrouping analysis, there are a number of significant and interesting differences between his analysis and those of the other subgroupings. Most notably, Mello adds a Group 9, in which he puts Tupinambá and Kokama, which corresponds to Group 3 in the other analyses. He also subdivides the traditional Group 2 into two groups, one with Sirionó and the other with Guarayú. Furthermore, his Group 6 corresponds to an amalgamation of the other analyses’ Groups 4 and 6, with the notable exception of Parintintin. Finally, he clusters Kayabí and Kamaiurá together unlike previous analyses. These significant differences, particularly in light of the potentially more reliable data collection and analysis methods used in Mello (2000) make it particularly important as a basis for comparison in our current study.

---

4 We assume that Omagua would cluster with Kokama in Group IX, but Tapieté, Yuki, and Ava Canoeiro are also missing, with few clues as to their distribution.
2 Methodology

2.1 Data Collection and Diversity

Linguists traditionally use the Comparative Method, both for reconstruction of proto-segments and words, as well as to determine subgrouping within language families. Languages are grouped according to shared innovations, developments from the proto-language that appear in a subset of languages. Subgrouping using the Comparative Method is based on the assumption that it is more parsimonious to assume that languages that exhibit shared innovations inherited those innovations from an intermediate proto-language, rather than proposing that each language developed the innovation independently. Of course, this is complicated by possible contact effects and diffusion of features, but traditional reconstruction assumes that such effects are negligible compared to the main effects of shared innovations. In either case, shared retention is not indicative of any particular subgrouping.

While the Comparative Method tests theories about language relationships, phylogenetics assumes relatedness, and provides subgroupings using an optimality criterion. Thus, a requirement of using phylogenetic methods in linguistics is that one considers the languages to be related at some level. In the case of the Tupí-Guaraní family this is fairly unproblematic. Published work on the family is, for the most part, in agreement about which languages form part of the family. There are two notable exceptions. First, some authors, such as Schleicher (1998) place Sirionó outside of the Tupí-Guaraní family. More significantly, both Omagua and Kokama, two languages often treated as dialectal variants of one another, have previously been analyzed as sufficiently mixed in heritage to not qualify as strictly Tupí-Guaraní (Cabral 1995; Schleicher 1998; see also Dietrich 1990).

For the phylogenetic analysis, we used a dataset of 20 languages, 18 of which are Tupí-Guaraní and two of which are Tupian. Based on the extant subgrouping hypotheses, we made sure to include a language from each of the eight groups, with some groups represented by more than a single language. Within each group, representative languages were chosen to be included in the analysis based on the availability and quality of primary sources. Not all Tupí-Guaraní languages have sufficient description to gather the necessary data for our purposes, and we have not completed data collection for all languages for which it is possible. The two Tupian languages were chosen to be outgroup languages because they are conventionally ascribed to sit just outside of the Tupí-Guaraní cluster in the larger Tupian family. Aweti is sister to Tupí-Guaraní and Satere-Mawé is sister to the Aweti–Tupí-Guaraní subgroup.

The reason for including data outside the Tupí-Guaraní family is so that we are able to root the tree. As discussed clearly in Dunn et al. (2008), rooting a tree consists of picking up an unrooted tree indicating relationships between languages and “suspend[ing] it at different points” (Dunn et al. 2008:723). By including two outgroup languages, we know from which point to suspend our tree: from the point at which the more distant outgroup language connects to the unrooted tree. The reason to include at least 2 outgroup languages is in order to test the monophyly of Tupí-Guaraní (i.e., if the Tupí-Guaraní languages form a subgroup, leaving the second outgroup

5 This will naturally lead us to re-test the data using similar methods once our sample is complete with all 27 languages that do have sufficient data represented.
language outside).

For each of the 20 languages included in the study, we collected data using primary sources such as wordlists, dictionaries, grammars, and personal fieldnotes. We chose which lexical items to include in our list by taking a 200-item Swadesh list which was expanded with words that are culturally important in the area (animal names, food items, tools, weapons, other artifacts, and verbs) and words that had reconstructed protoforms in the previous reconstruction work. Words in the Swadesh list that were irrelevant for climatic, cultural, or linguistic reasons were removed. We then searched the sources in each language for words bearing the meaning of our chosen lexical list items as well as words with closely related meanings. Our final dataset for this study included 572 word meanings (5 numbers, 52 body part terms, 25 food terms, 52 animal terms, 42 kinship terms, 7 color terms, 18 time-related terms, 41 natural environment terms, 49 artifacts, 63 adjectives, 13 non-adjectival descriptive words, and 205 verbs). The full list of meanings and the corresponding words found for each language are available upon request.

2.2 Character Coding

Characters used for computational phylogenetic methods need to follow a number of assumptions in order for the methods to be trustworthy. First of all, researchers need to be able to compare characters across languages, i.e., to formulate hypotheses of homology. It must be noted here that character coding is a hypothesis of homology for many characters to be tested with an optimality criterion (parsimony, maximum likelihood). In the absence of evidence to the contrary, the researcher accepts common ancestry as the simplest explanation for the observed similarity. Characters must also be “heritable,” which is not a trivial concept in linguistics. At a microscale, most linguists accept that language is not “inherited,” but acquired in a more or less unique way by an individual based on his or her experience within a linguistic environment. At a more macroscopic level, though, the language of a speech community exhibits “heritability” as its characteristics are retained, modified, or lost. The second main assumption for characters used in a phylogenetic analysis is that they are independent, i.e., that they evolve separately from each other. Thus, their congruence is evidence for shared history.

There are two ways that lexical data can be coded for phylogenetic analyses: treating each etymon as a character with states present or absent in each language, or treating each meaning as a

---


7 In biology, homologous morphological structures, genes, or nucleotides can be defined as being manifestations of the same morphological structure, gene, or nucleotide present in the common ancestor of the group studied (e.g., the limbs of amphibians, reptiles, birds, and mammals are homologous and they are modifications of the limbs of the common ancestor of all tetrapods).
character with states corresponding to cognate sets within this semantic slot.

The first coding method is based on etyma, which are unambiguously “inherited,” but which might not be independent. The independence of etyma which are close to each other in semantic space is essentially a question on the prevalence or not of semantic shift: When a new etymon enters a semantic slot, will the old etymon disappear or change meaning? The answer is that both outcomes can happen. If semantic shift happens very rarely, then etyma are clearly non-independent and should not be used as the basis for characters. However, if semantic shift is more or less common, then etyma are independent to a certain extent and there is subgrouping information that can be lost if it is not taken into account. This coding scheme results in binary characters with states (present/absent) that are comparable across many characters, making it easy to apply a common evolutionary model and use likelihood-based and Bayesian methods.

The second coding method is based on the homology of meanings, which are not as intuitively “inherited” as etyma, but they are independent. In fact, “meanings” in the form of body parts, animals, natural phenomena, and others exist outside language and language applies a word to them. The assumption here is that these “meanings” are comparable across languages. This coding scheme results in multistate characters (each state corresponds to a cognate set) without the implication of a unified mechanism for state changes. This makes the characterization of models, necessary for likelihood-based and Bayesian methods, more complicated and could result in overparameterization. Nevertheless, multistate characters of this type are ideal for parsimony methods.

Most previous works (Greenhill and Gray 2005; Gray et al. 2009; Atkinson and Gray 2003) have used essentially the second type of coding (but see Rexova et al. 2006), followed by a step of binary recoding, where each cognate set within a semantic slot is considered a character and is coded as present or absent. This step introduces a great number of non-independent characters in the data, which artificially inflates support for certain groupings and could make ancestral state reconstructions difficult to interpret. We therefore decided not to use binary recoding in our analysis.

As our dataset had many more word meanings than a typical Swadesh list, we were better able to detect semantic shifts. So, we decided to use the etymon-based approach (the first approach) as a first step. All lexical items were put in cognate sets by the Tupi-Guarani Comparative Project team. Cognate words that had undergone semantic shift were included in their respective cognate

---

8 To our knowledge, there is not much more than speculation on the frequency of semantic shift versus loss.
9 This is not generally true as differences in the semantic boundaries among languages are common. But, many meanings are indeed comparable and stable across languages in general or at least across languages from a certain family due to shared environment, culture, etc.
10 Typically the more parameters a model has, the better it fits the data. Nevertheless, too many parameters can create an overparameterization problem where the data are not enough to estimate with confidence all the parameters, which results in low accuracy.
11 This last step makes the resulting matrix very similar to the first type of coding with one crucial difference: the binary recoding ignores presences of an etymon in a language if it has undergone a semantic shift.
12 We are also planning to code the same dataset in a meaning-based multistate approach and compare the results.
13 Team members include: Keith Bartholomei, Zachary OHagan, Michael Roberts, and Vivian Wauters, all under the guidance of Professor Lev Michael.
Compound words were temporarily placed only in one cognate set due to time constraints. Each cognate set was then converted to a binary character (1: etymon present in the language, 0: etymon absent). When a language had no entry for a semantic slot, it was coded as unknown for all cognate sets associated with this meaning.

This procedure resulted in a matrix of 3,361 characters with 31.15% missing data. We detected 548 cases of semantic shift out of 10,600 entries or 5.17%. Semantic shift is probably overestimated in the dataset for two reasons. First, there were cases where presumably the same historical semantic shift was counted multiple times (once per daughter language that shows the shift). Second, some of the sources had more generic or vague meanings than others, thus creating “extra” cases of semantic shift. The influence of missing data and semantic shifts on our analysis are further discussed in §4.

2.3 Parsimony Analysis

Parsimony is an intuitive optimality criterion used extensively in both biology and linguistics even before computational methods were developed. It is based on the notion that the simplest explanation (the one that involves the least changes or steps) is the preferable one. One of its main advantages, apart from its simplicity and intuitiveness, is that it does not need an explicit model of evolution and thus can be used on a variety of character types. Parsimony has been shown to be statistically inconsistent under certain conditions (Felsenstein 1978). These conditions, usually described as “long branch attraction,” involve long branches (such as language isolates without close relatives in the analysis) and quickly evolving characters that have a limited number of possible states (Schulmeister 2004).

We decided to do a parsimony analysis of our dataset mainly for comparative purposes, as plain parsimony is not the best method to analyze our characters. The reason for that has nothing to do with the “long branch attraction” problem, as we have no language isolates, and although our characters are binary, we do not expect them to evolve quickly (i.e., we do not expect any given etymon to switch many times from present to absent and vice versa). The main problem is that we expect an asymmetry in the rate that cognates are lost versus gained and straight parsimony assigns the same “penalty” to both directions. One could argue that we could weigh the gains more than the losses, but this would involve a more or less arbitrary choice of a number. So, we decided to do a plain parsimony analysis and compare the results with those of the more realistic Bayesian analysis (see below).

As a support measure for our parsimony analysis, we used bootstrapping, a resampling method. For each iteration of the bootstrap algorithm (called a pseudoreplicate), a surrogate dataset equal in size with the original is produced through resampling the characters of the original dataset with replacement. This results in certain characters being omitted and others overrepresented. Then, the same analysis is performed on this surrogate dataset as in the original dataset and the resulting

---

14 All the resulting cognate sets are available on request.
15 The section of kinship terms was not included when calculating the amount of semantic shift because some of the sources were not accurate in their description of the kin system and the amount of semantic shift was overestimated.
16 A method is statistically consistent if it is guaranteed to approach the correct solution if given enough (technically infinite) data that are generated according to a specific set of rules (the model).
shorter tree(s) are recorded. This procedure is repeated a large number of times and the results are summarized in a 50% majority tree (where 50% means that a clade/subgroup with 50% bootstrap value was present in 50% of the collection of trees at the end of the bootstrap analysis). A level of 50% is usually chosen as a cut-off point because clades present in more than 50% of the trees are guaranteed to be compatible with each other. There is considerable disagreement and debate on what bootstrap values actually mean in a phylogenetic context and how to interpret them (for more information see §4).

The parsimony analysis was performed in PAUP* 4.0b10 (Swofford 2003). Characters were equally weighted. We performed 40 heuristic searches starting each time with a tree built by stepwise addition of languages and the addition sequence was random. We also performed bootstrapping to calculate support values for our nodes (Felsenstein 1985). We did 1000 pseudoreplicates on our whole dataset and 5000 pseudoreplicates on only the parsimony informative characters, as too many parsimony uninformative characters might artificially decrease support (Soltis and Soltis 2003).

2.4 Bayesian Analysis

The essential outcome of a Bayesian analysis is the posterior probability of the parameters of our model given the data we have. The stereotypical coin-flip equivalent is, how possible is it that the coin is fair (the parameter of our model) given the fact that we just got 90 heads in 100 flips (our data). For a Bayesian analysis, we start with a model of evolution that has some parameters, with our prior beliefs for the probability distribution of these parameters, and with our data, which we assume were generated according to this model. The algorithm then gives us the posterior probability distribution of our parameters and a typical way to summarize the results is to take the 95% confidence interval for each parameter.

For our analysis we chose the restriction site model implemented in MrBayes 3.2.0 (Huelsenbeck et al. 2001; Ronquist and Huelsenbeck 2003). This model allows for two different rates of cognate gain and loss. We believe that this is a more realistic model for our data than a symmetrical model, as we expect that cognate loss is easier than cognate gain. However, it should be noted that no rates were indicated a priori, rather the prior for these two rates is equal and they are able to diverge from there in the expected or the opposite direction. The data were corrected for non-observable all absent sites. We also allowed for variation in the rates across characters, to simulate the situation of quickly evolving etyma (high probability for loss and gain, e.g., through borrowing) and slowly evolving ones. The rate variation across etyma was set to follow a gamma

---

17 Of course, if the data are really generated according to our model, then the method is guaranteed to find the correct values of the parameters if given enough data. In reality, though, any evolutionary model is less complex than the real evolutionary process that generated the data. Simulations with generated data under our most sophisticated models and analyzed with our crudest models still give correct results in most cases, showing that phylogenetic methods are robust when the model assumptions are violated (Sullivan and Swofford 2001).

18 The 95% confidence interval encompasses the most probable values of each parameter given our data (the remaining 5% is the tails of the posterior probability distribution.

19 This is because we cannot observe etyma that do not exist now in any language because they were gained and then lost from every language. This correction allows the model to be estimated properly by adding dummy characters to compensate for this bias.
shape distribution and the $\alpha$ parameter had a uniform prior from 1 to 200.$^{20}$

Bayesian analysis was conducted with MrBayes 3.2.0. We performed two independent Metropolis-Coupled Markov Chain Monte-Carlo (MCMC-MCMC) runs of 4 chains each (1 cold and 3 heated).$^{21}$ The analysis ran for 5 million generations and was sampled every 1,000 generations.

3 Results

3.1 Parsimony

There were 1,162 parsimony informative characters. The heuristic searches resulted in 2 equally parsimonious trees of 4,958 steps. The two trees differed only in the position of Guarayú (sister to Xetá and Tapieté or sister to Paraguayan Guaraní). The strict consensus of the 2 trees is shown in Figure 2. Bootstrap values are visible for nodes with more than 50% support. The bootstrap values for a few branches increased a bit when using only parsimony informative characters, but overall the two bootstrap analyses gave very similar results and no value changed from lower than 50% to higher or vice versa. The numbers shown are from the bootstrap analysis on the parsimony informative characters only.

The parsimony analysis produced a small number of strong groupings supported by a large number of characters (although this support is partly artificial, see §4). All of them except for one are at the tips of the tree and group only 2 or 3 languages together. All the higher level groupings have much less supporting characters (as the lower than 50% bootstrap values show).

$^{20}$The gamma distribution can take a variety of shapes depending on the value of the parameter $\alpha$.

$^{21}$The samples of the posterior probability distributions are taken only through the cold chain. The 3 heated chains are used to explore tree space more effectively (they can cross areas of low likelihood easier than the cold chain and so they do not get stuck on local optima) and they can swap places with the cold chain if they find an area of higher likelihood.
3.2 Bayesian Analysis

The majority consensus tree resulting from the Bayesian analysis is shown in Figure 3. The names of the languages are color-coded according to the proposed 8 groups. The values on the branches are the posterior probability of the clade supported by each branch. The branch lengths represent the amount of change that has occurred along the branch and in this value, time and rate are confounded (i.e., a long branch could be due to longer time, fast rate of change, or both). The rates of etymon gain and loss estimated through the analysis were very asymmetrical in the expected direction with the loss rate being 9 times the gain rate.

Convergence between the 2 runs was determined by the average standard deviation of split frequencies (which should approach 0 as convergence is reached, and was in fact 0.0072 at the end of the 5 million generations). The Potential Scale Reduction Factor (PSRF), another convergence
metric, was between 0.999 and 1.001 for all parameters, which shows that the sample for the estimation of each parameter was adequate and the 2 runs had converged. 25% of the initial samples were discarded as burn-in, to ensure that each run had reached stationarity before starting to sample it.

![Diagram of language relationships](image)

Figure 3: Majority-rule consensus tree with posterior probability values. Languages are color-coded according to previously proposed groupings.

4 Discussion

4.1 Characteristics of Dataset

4.1.1 Advantages

Our dataset is the most complete to date for the Tupí-Guaraní family. It contains approximately half of the languages in the Tupí-Guaraní family and it has good coverage of the suggested groups within the family. It also includes 572 word meanings, which is almost 3 times the size of a typical Swadesh list used in other studies (Gray et al. 2009; Atkinson and Gray 2003; Nakhleh et al. 2005). We believe that by including as much information as possible, we have a better chance to recover
subgrouping information for the languages in question. The Swadesh list can be a starting point, but there is no a priori reason to restrict the lexical items collected to the ones included in the list, as the terms that are stable for a language group is to a certain degree idiosyncratic. On the other hand, choosing a few “good” cognate sets to analyze without explicit criteria does not test all the available evidence. As far as the coding method is concerned, we are not aware of another study that used etymon-based coding including reflexes that have undergone semantic shift. Overall, our dataset is both extensive and provides good sampling of the Tupí-Guaraní family and is promising for phylogenetic analyses.

4.1.2 Limitations

There are a number of qualities of our current dataset that limit its usefulness. In order to further develop this analysis in the future, it is important to recognize these limiting factors, and how they contribute to deceptive trends in the data, in order to mitigate deleterious effects in the future. Our two main limitations are: 1) non-independence of the data, and 2) low-quality sources providing misleading data.

Non-independence of our characters means that the states of one character affect the states of other characters. In our dataset, this is caused by languages losing an etymon when they gain another etymon for the same meaning, instead of the erstwhile item shifting to a new meaning. The tendency toward loss instead of shift leads to a high probability that the presence of a reflex in one cognate set generally corresponds to a lack of a reflex in the other cognate sets for that particular item, as seen in Example (1). While we hoped that a high level of shift (i.e., non-loss of words even when a new cognate entered a semantic domain) would mitigate the level of non-independence caused by cognate loss, the level of shift, at 5.17% was not high enough to rule out non-independence as a confounding factor. Additionally, it has been shown for Indo-European that not controlling for non-independence can be deleterious to anything but the smaller subgroupings within the family (Rexová et al. 2003). Nevertheless, the low overall percentage of semantic shift does not capture its distribution over different lexical items. In most lexical items, there was no detected shift at all or just sporadic cases of semantic shift, but in a few cases there was extensive semantic shift that essentially made the etyma independent of one another. Some examples of these semantic “complexes” are: face-head-eye-cheek, nipple-breast-chest, heart-liver-intestines-belly, sand-beach-dust, house-village-shelter, plant-bury-hide, want-love-like. We will explore the possibility of including etymon-based coding only for such items (see §4.4).

(1) ‘want’

(1) Paraguayan Guaraní yvoty; Tapieté mba’e poti; Guarayú i’potrí; Yuki bajúti; Tupinambá potyra; Tapirapé ywátyr, patyr, hywatyr; Parakanã potyr; potiša; Tembé putir, potí, potyra, iboti, i-putir; Kayabí ywotyr; Parintintin yvatyrí, yvytyr; Kamaiurá potyt; Guajá mítir, mítí; Awetí potir; Mawé pohit

(2) Omagua ssa; Kokama tsetsa

---

22This percentage of course refers to our dataset, which included only a fraction of the etyma present in the languages.
One effect of non-independence is that it exaggerates any trends in the data, in the following manner: if two languages exclusively share a cognate set, such as Omagua and Kokama, they are coded the same (as 1) for that set, while all other languages are coded as different (as 0). However, in the following set, where all other languages have a cognate except Omagua and Kokama, they will be coded exactly opposite, thus creating two data points that link Omagua and Kokama as an exclusive pair to the exclusion of the rest of the languages. Some cases of trends like this are shown in Figure 4 as well as Example (1) using the words for “neck” and “want,” respectively. In order
to argue for the independence of these two points, we would be positing that the loss of a reflex of one cognate and the presence of a different cognate for the same lexical item are completely unrelated. Since this is not the case, it may have artificially increased the support for subgroupings that have a high number of shared innovations. In our current dataset, this is most likely the case for Omagua and Kokama because of the high level of non-Tupí-Guaraní lexical items in their language. However, this could also be contributing to the high numbers for the Parakanã, Tapirapé, and Asurini do Tocantins subgrouping.

Another effect of non-independence is realized in the marking of unknown character states. As described above, for each lexical item that we did not find an entry for in a particular language, we coded each character (etymon) associated with this semantic slot as unknown, with a question mark (‘?’). In this way, the analyses treated these unknown cases as either present or absent, in order to fit the data. The number of unknown states for a given character was entirely dependent on how many known cognates for that character there were in the other languages. For example, our data on Xetá lacks both a word for “flute” and “bead.” However, there are 12 different cognate sets for “flute,” and only one for “bead,” which result in a total of 13 unknown character states in Xetá. It is still unclear what the final effect of these characters is on our analysis.

Another possible distortion from the data stems from variability in data quality. Each of the sources has a different level of detail in differentiating the semantic domain of each lexical item. In particular, the length of the branches in MrBayes could be distorted by an overrepresentation of single set cognates in a single language for a single lexical item. This is most evident in Parintintin, which is represented with a very long branch length; the source for Parintintin is a dictionary that makes very little subcategorization of terms in the dictionary. For example, the lexical entry “bee” has 14 entries, none of which are cognate, the lexical for “cousin” (gender-unspecified) has 8 non-cognate entries, and the lexical entry “road” has 6 non-cognate entries. This creates an abundance of single-item cognate sets for Parintintin, which leads MrBayes to analyze it as extremely divergent compared to other languages. This is also the case for Tembé, which has for instance 8 non-cognate words under the lexical item “beat.” While one would expect that each of the words listed under a single heading actually have some semantic distinctions between them, the lack of description in the literature causes a spurious effect in the data.23

4.2 Comparison of Methods

The parsimony analysis recovered a few low-level subgroups supported by a large number of characters, as well as the Tupí-Guaraní family. All other groupings had low bootstrap values. As mentioned in the methods section (§2), bootstrapping has been used as a measure of support or even accuracy (Hillis and Bull 1993), but there is considerable debate over its interpretation (Soltis and Soltis 2003). The only thing the low bootstrap values tell us for sure about a branch is that there are few characters supporting this branch within our dataset. This does not necessarily mean that the clade is not accurate. In fact, if the branch is short (because the period of shared history for the clade was short), it is very probable that we will have few characters that change along this

23Note, however, that the branch length of Yuki does not appear to be an artifact of under-informative data, but rather the Yuki lexicon is remarkably divergent, having lost many of the Tupí-Guaraní cognates found in many other languages in the family.
branch. In this case, we will never get a high bootstrap value, even if there is nothing else that contradicts this grouping. So, bootstrap values are good indicators of how much support there is in the dataset for a particular grouping and if there are characters that support conflicting groupings.

Both the parsimony and the Bayesian analyses recovered the same highly supported low-level subgroups (Tupinambá and Tembé, Omagua and Kokama, Tapirapé sister to Parakaná and Asurini do Tocantins), as well as the Tupí-Guaraní family node. However, the Bayesian analysis showed two additional well-supported subgroups: the clade including all languages from Groups 1 and 2 (although its internal relationships are not fully resolved) and its sister relationship with the Tupinambá-Tembé clade. All of these languages were grouped together in the parsimony analysis, but with low support, except for Yuki which was grouped with the Omagua-Kokama clade.24

As mentioned in §2, we consider the model used in the Bayesian analysis much more realistic than the symmetrical penalty of the parsimony analysis. The rates of cognate gain and loss estimated through the Bayesian analysis confirmed our expectations, as they were highly asymmetrical. Also, the Bayesian analysis is able to use all the coded characters (while parsimony was using only a third of them, which were parsimony informative). Therefore, we consider the results of the Bayesian analysis to be more trustworthy with this type of data. However, even with the better model, the posterior probability of higher-level subgroupings in the Bayesian analysis is still low, in many cases much lower than the 95% confidence interval. Many higher-level branches are also fairly short. It could be that the Tupí-Guaraní languages diversified fast and there is no adequate evidence to resolve the topology of the tree at that level. Additional data from other lines of evidence (phonology, morphology) could help clarify these relationships (see §4.4). In case the base of the Tupí-Guaraní family is a real polytomy, with many subgroups diversifying more or less at the same time, phylogenetic reconstructions will be sensitive to the addition and removal of characters, as there will be many conflicting characters for different topologies.

4.3 Promising Trends

Despite the possible confounds to our analysis, this preliminary study is demonstrably useful based both on a number of expected subgroupings based on previous reconstructions, as well as particular structures that point to revelatory possibilities for re-analysis of the relationships between specific languages.

4.3.1 Comparison with Previous Reconstructions

Because a full understanding of the accuracy and practicability of using phylogenetic methods on linguistic data has not been conclusively determined, one of the goals of recent phylogenetic studies has been to evaluate the results gathered in comparison to previous research (Dunn et al. 2008; Atkinson and Gray 2003). In order to determine the basic validity of our data and methods, we used two outgroup languages, as discussed above (§2). We rooted the tree using Mawé, since it is considered to be further outside the family than Awetí, and then our analysis of the data as a family was validated by the unequivocal placement of Awetí as outside of the Tupí-Guaraní family.

24Yuki is a highly divergent language and we may need additional and more reliable characters to place it with confidence.
This shows that the family is monophyletic, i.e., that the languages can be reliably determined to be a family to the exclusion of the two outgroup languages.

Having determined that our data did in fact produce an exclusive Tupí-Guaraní clade, it is possible to compare the results of our current analysis with the previous subgrouping hypotheses. As discussed above, the diachronic study of the Tupí-Guaraní family has, for the most part, not been rigorously undertaken, and therefore the 8 conventional groups into which languages have been divided is questionable (of course, Mello 2000 does not follow this convention entirely). Nevertheless we will discuss the points of agreement of our analysis with the 8 groups, and then separately discuss important comparisons with Mello (2000) not covered by the prior hypotheses.

In the parsimony analysis as well as MrBayes, the most consistent groupings reproduced above the pairwise level (i.e., more than two languages put together) were in Groups 1 and 4. For the parsimony analysis, the three Group 1 languages were reproduced in a clade with Guarayú, but the internal topology of the clade was unresolved. This clade is a combined Group 1 and Group 2 clade with the exception of Yuki.

In MrBayes, Guarayú and Yuki were both included with the Group 1 languages. In the case of Group 4, both trees produced an identical subgrouping of three languages: Parakanã with Asurini do Tocantins, followed by an immediately higher node of Tapirapé. Both trees did not include Tembé in this clade, and instead clustered Tembé with Tupinambá, the possible reason for which is discussed above. Also reproduced in these trees was the close relationship between Omagua and Kokama, which, in spite of the problems discussed above, is an encouraging result.

In contrast to the congruence between the current analysis and those of Jensen (1998), the Mello (2000) language groupings are not as well-matched with our analysis, partially due to differences in which languages were included in each sample. Paraguayan Guaraní and Xetá, which form a subgroup in Group 1, do appear as part of a larger clade including Guarayú, which is in Group 3 for Mello (2000). Also, while the Group 6 languages (Group 4 in Jensen 1998) Parakanã, Asurini do Tocantins, and Tapirapé pattern together perfectly, even following the subgrouping in Mello (2000) with Tapirapé appearing one node higher, Asurini do Xingu, which is included in this Group by Mello (2000), is not reproduced in either the parsimony analysis or MrBayes.

While Dietrich (1990) explicitly states that his work is not a subgrouping analysis, the gradient conversativeness rating of each language does result in some low-level groupings, many of which follow Jensen (1998). There are, however, some important differences. Dietrich does not find Tupinambá and Kokama to be similar enough for subgrouping. Furthermore, like Mello (2000), the conservativeness ranking finds Asurini do Tocantins to be an outlier in a Group containing Gua-jajára and Tembé (which we consider dialectal variants of the same language) as well as Kamaiurá, a subgrouping that is not supported by either of our analyses. Further, Dietrich finds Kayabí to be much more closely related to Tapirapé and Kamaiurá than to Asurini do Xingu. This is another subgrouping that is not supported by our analyses, although both the parsimony and Bayesian analyses split Kayabí and Asurini do Xingu into very different trees. This points to considerable confusion about the placement of Kayabí, as it is one of the significant differences between Jensen (1998) and Rodrigues and Cabral (2002) (see §1.2).

In summary, there is a reasonable level of agreement between our analysis and that of Jensen (1998) about some basic groupings in the Tupí-Guaraní family, and slightly less agreement with Mello (2000). However, there were a number of distinctly different subgroupings. Both the parsi-
mony analysis and MrBayes placed Yuki in different positions (with Omagua and Kokama for the parsimony analysis and with Xetá for MrBayes), neither of which matches any previous analysis. As discussed above, we attribute this to the highly divergent lexicon of Yuki. Because non-cognate forms were all coded the same (as 0), the analyses tended to group divergent languages together. However, both parsimony and MrBayes grouped Guajá with Asurini do Xingu, which is not supported by any previous analysis. This particular grouping will be more informative once data on the other languages in our sample have been gathered because currently only one language from each of the previously hypothesized subgroups to which these languages belong has been included in the sample. Therefore, it may be that the two are grouping together based on a lack of more closely related languages, and the inclusion of other languages in Group 6 and 8 will result in a different pattern in our results. If this is not the case, however, it may indicate that the languages in these two groups, like those in Groups 1 and 2, may be better analyzed as all part of a larger clade within the Tupí-Guarani family.

4.3.2 Subgrouping Implications

The original purpose of the Tupí-Guarani Comparative Project was to correctly determine the relationship of Omagua to the Tupí-Guarani family. Based on this research goal, the most promising result from this analysis is the way in which the languages of the erstwhile Groups 1 and 3 combine. The inclusion of Tupinambá and the Omagua-Kokama clade as successive branches of the Group 1 (more or less) clade points to a need for re-analysis regarding the immediate proto-language of Omagua and Kokama, not as the direct descendent of Tupinambá, as previously suggested (Cabral 1995). In previous research, evidence has focused on the relationship of Kokama (and thus Omagua) as Tupí-Guarani or not, always in relation to Tupinambá as the potential predecessor (Cabral 1995). Based on the results of the phylogenetic analysis, however, Omagua and Kokama are very clearly a subgroup, as are Tupinambá and Tembé (although recall that some data issues may be creating a spurious subgrouping in that case), but Tupinambá and the Omagua-Kokama clade are members of a paraphyletic group, in that they do not subgroup to the exclusion of other languages. Instead, the MrBayes analysis suggests that Paraguayan Guarani and and the other languages in Groups 1 and 2, as well as Tupinambá and the Omagua-Kokama subgroup all share a common ancestor. If this subgrouping turns out to be robust, it would suggest that the Tupí-Guarani ancestor of Omagua and Kokama is fairly old, which would help explain the large amount of change the language has undergone, changes so dramatic as to make some believe that it is not a Tupí-Guarani language at all, as discussed above.

4.4 Future Directions

In order to improve the phylogenetic analyses of subgrouping in the Tupí-Guarani languages, and to further determine the validity of observed trends within our extant dataset, we will include more languages in our dataset and we want to use a variety of characters for phylogenetic reconstruction.

In terms of language addition, we will be adding 7 more Tupí-Guarani languages as well as at least one more outgroup language from the Tupian stock. This way we will have at least 2 languages per proposed group so that we can test previous subgrouping hypotheses. We will also
be able to test the sister relationship of Awetí with Tupí-Guaraní.

In terms of characters and coding, we are planning to use the meaning-based method of character coding for our lexical characters to avoid non-independence problems. An additional possibility we are going to explore is keeping the etymon-based coding only for the semantic “complexes” that exhibited a lot of semantic shifts in order not to lose any subgrouping information. Apart from coding lexical characters per se, we will use our cognate sets to find and code phonological characters as well. As phonological characters are the most reliable line of evidence in the Comparative Method, it is important to include this type of data in our phylogenetic analyses.

Additionally, we plan to undertake the same analyses discussed in this study on a sample of morphosyntactic functional items in the Tupí-Guaraní languages. This database will include data from the same languages included in the lexical database, but instead of gathering forms based on an established set of items, we have chosen to cull all available functional morphemes from each language, and then organize them by type. The result of this more holistic data collection technique is that we can be more confident in obtaining a full sample of the forms within any given language. As is the case with phonological characters (discussed above), functional items are less prone to superficial effects like contact (Thomason 2001). Therefore, one would expect that any true trends in our data would be more likely to manifest themselves in an analysis of less volatile data, such as the functional items. Furthermore, because functional morphemes can follow grammaticalization trajectories, and might therefore be less likely to disappear from a language due to the entrance of a new morpheme, our assumptions of independence violated above may not be as problematic for the functional list.

In addition to our extension of the current cognacy analysis to functional items and the other types of coding based on semantics and phonology, we will also be reconstructing the Tupí-Guaraní family using the traditional Comparative Method. Despite the valuable insights garnered from phylogenetics, there is currently no viable alternative to the Comparative Methods for achieving respected and believable results. Therefore, our “by hand” reconstruction, as discussed in the introduction, will be invaluable. It will be an indispensable tool for ancestral state reconstruction, as well as providing us with a valuable basis on which to calculate the overall utility and accuracy of these exciting new methods.

References


Harrison, Carl H. 2009. Pedagogical Information and Drills for the Asurini Language.


Subgrouping in the Tupí-Guaraní Family


Subgrouping in the Tupí-Guaraní Family


Natalia Chousou-Polydouri
University of California, Berkeley
Department of Environmental Science, Policy, and Management
130 Mulford Hall # 3114
Berkeley, CA 94720
nataliacp@berkeley.edu

Vivian Wauters
University of California, Berkeley
Department of Linguistics
1203 Dwinelle Hall
Berkeley, CA 94720
vmw@berkeley.edu
A ‘Perfect’ Evidential: The Functions of -shka in Imbabura Quichua*

JESSICA CLEARY-KEMP
University of California, Berkeley

1 Introduction

This paper investigates the functions of the verbal suffix -shka in Imbabura Quichua (IQ), a Quechua IIB language spoken in the northern highlands of Ecuador. Cole (1982:148) analyzes -shka as — among other functions — a perfect aspect inflection on main verbs, noting that it indicates in addition “a degree of surprise.” Cognate forms in other Quechua languages have been labelled ‘sudden discovery tense’ and translated as ‘it turned out that...’ even though they are reportedly not restricted to past tense (see Adelaar and Muysken 2004:223). These characterizations suggest that -shka is a perfect aspect marker, with additional semantics of mirativity. I will argue in this paper that, although -shka does mark perfect aspect in periphrastic constructions (where the verb in -shka combines with copula ka- ‘be’), as a verbal inflection it marks ‘non-eyewitness past tense,’ contrasting in this function with -rka ‘eyewitness past tense.’ In addition, I will argue that the constellation of functions modern-day IQ -shka fulfils — including deriving nominals and passive participles — is explicable by assuming a series of typologically common diachronic changes, starting from a ‘resultative nominalization’ function.

The findings in this paper are based on data obtained from two speakers of IQ: MC, a 50-year-old woman from the area of Mariano Acosta, and ACO, a man from Otavalo. Both consultants are native speakers of IQ, having learnt Spanish as a second language, and English as a third language. MC has been living in the United States for approximately 25 years, and ACO for 14 years. The data, collected by members of the 2009-2010 UC Berkeley Field Methods class, is of two main types: recorded monologues and dialogues, and targeted elicitation. The data presented in this paper is from targeted elicitation, unless labelled otherwise. Some data is also taken from Cole (1982).

The paper is structured as follows: §1.1 briefly describes the linguistic and social context of the Imbabura Quichua language; §1.2 outlines the analytic framework adopted for describing tense and aspect; §2 identifies the major categories of tense and aspect in IQ; §3 describes and exemplifies the inflectional and derivational functions of -shka in IQ, and considers the uses of -shka in personal and traditional narratives. §4 presents a diachronic proposal to account for the synchronic polyfunctionality of -shka; and §5 briefly summarizes the findings of this paper.

*I would like to express my heartfelt gratitude to Mariana and Augusto for patiently and enthusiastically sharing their language and culture with us. I would also like to thank the members of the 2009-2010 UC Berkeley field methods class for very valuable comments and discussion relating to this topic. Thanks are also due to the audience at the 2012 winter meeting of the Society for the Study of Indigenous Languages of the Americas (SSILA), who gave valuable feedback on an earlier version of this paper. Finally, Lev Michael made numerous comments and observations that significantly improved this paper. My thanks go to him, but as usual, any errors remain my own.
1.1 The Imbabura Quichua Language and its Speakers

Imbabura Quichua (IQ) is a Quechua IIB language spoken by approximately 30,000–50,000 people in the northern highlands of Ecuador (Cole 1982:3). The Quechua languages have been in contact with Spanish for centuries, and the majority of IQ speakers are also fluent in Ecuadorean Spanish. The Quechua languages of Ecuador (and of Colombia) have lost much of the complex head-marking morphology exhibited in other Quechua languages. Subject marking on the verb is retained, but only the first person object marker remains, and possessive suffixes have been lost from nouns, possession now being marked on the possessor by genitive suffix -pa. As such, IQ is largely dependent-marking, especially compared to other varieties of Quechua (Adelaar and Muysken 2004:208). Argument alignment is nominative–accusative; case-markers attach to the NP phrase–finally, nominative being unmarked, and accusative marked by -ta. Constituent order is relatively free in main clauses, but SOV predominates, and is more strictly required in subordinate clauses. This correlates with general head-final order in constituents. Major word classes in the language are ‘noun’ and ‘verb,’ with small closed classes of demonstratives and quantifiers, plus a set of spatial terms that may be classed either as postpositions or relational nouns. No separate class of adjectives is apparent (although close analysis may reveal the need for a separate adjective class), and property-denoting terms essentially have the same distribution as referential nominals, being able to fill argument position and take case-marking, as well as functioning as modifiers and copula complements. A pervasive phenomenon in IQ discourse is the use of a set of so-called ‘validators’ — clitics that combine evidential, modal, and focus-marking functions. The main validators are =mi ‘direct evidential,’ =shi ‘indirect evidential~uncertainty,’ =chu ‘polar question~negative marker,’ and perhaps =ta ‘content question marker.’

The major previous work on IQ is Cole (1982), a grammatical description based on fieldwork undertaken with speakers of the Rinconada (including Mariano Acosta), Otavalo and San Roque dialects. Further references are listed therein.

1.2 A Neo-Reichenbachian Framework for Analyzing Tense and Aspect

Klein (1994), building on work by Reichenbach (1947), outlines a framework for describing and analyzing tense and aspectual distinctions cross-linguistically. The main elements of this approach are presented here in order to lay the groundwork for the analysis presented in §3.

1.2.1 The Building Blocks

The ‘lexical content’ (LC) of an utterance is the situation described in the clause, divorced from its finite temporal information. For instance, the LC of the English clause John was studying Quichua is {John study Quichua}. This is also the LC for the clauses John will study Quichua, John has studied Quichua, John studies Quichua, and so forth. In other words, the LC is atemporal. But it is “timeable,” that is, it can be linked to a temporal structure. This is done by means of tense and aspect marking. The temporal structure to which the LC is linked comprises three primary

---

1 See Kwon (this volume) and Cleary-Kemp (2010) for a more precise characterization of these elements and their functions.
elements: ‘utterance time’ (TU), ‘topic time’ (TT), and ‘situation time’ (TSit). The first two of these are relatively simple to define, while the third is somewhat more complex.

TU (Reichenbach’s S) is the time of utterance. In conversations this is easy to calculate — it is the moment of speaking. In non-immediate communicative situations, such as letters, email, books, graffiti, voicemail messages, pre-recorded television or radio programs, etc., the TU is less fixed, and more open to interpretation. The calculation of TU can be complex, but it is not crucial to the thesis of this paper, so I will not discuss it further here. TT (Reichenbach’s R) is the time about which a claim is being made. This can be overtly specified, as in [At 6:40 in the morning, on August 24, 2003]TT, Max was born, or it can be left to context. Occasionally, as in this example, TT is very precisely specified, but it is far more common for the exact span of the TT to be left open to be inferred from discourse context and world knowledge. Again, the details of how TT is determined are not vital to the arguments presented here, and so are set aside.

TSit (Reichenbach’s E) is the time for which the situation described by the LC holds. For instance, in the examples above, TSit is the temporal span during which it is the case that John is studying Quichua. {John study Quichua} is a one-state LC, because there is just one lexically-specified situation. In contrast, the LC {John leave the house} is two-state, since it encodes a lexically-specified change of state: from John’s being inside the house to John’s being outside. In such complex LCs, the initial state is referred to as the ‘source state’ (SS) and the final state as the ‘target state’ (TS). The TS in the case of {John leave the house} is clearly {John not be inside the house}, but the SS, in contrast, is not simply {John be inside the house}, rather it involves John’s being active in bringing about the TS. In other words, it is only felicitous to say John is leaving the house if John is in the process of bringing about the state of being out of the house, not, for example, if he is simply sitting inside the house reading a book. Klein (1994:105) notes that, for purposes of aspect marking, languages tend to choose one of the two states to treat as TSit. In English it is SS. This is evident when we consider how two-state LCs behave in the progressive construction, whose function is to indicate that TT is fully included in TSit (see Table 1). The sentence in (1) is felicitously uttered only if John is currently in the house, and is active in achieving the state of not being in the house.

(1) John is leaving the house

In other words, this utterance denotes the structure in (2).

(2) \{ ---[—]TT—\}_{SS}+++++++_{TS}

The sentence in (1) is not true if John is in the TS, having already left the house and being halfway down the street. Given that the construction in (1) locates TT within SS, and not within TS, this example clearly shows that English treats the SS of two-state predicates as TSit. With respect to this parameter, IQ behaves the same as English.

\[\text{2 Of course there is some leeway in the interpretation of this utterance, as with all natural language. If John has just stepped outside the house and is still on the front steps, perhaps the utterance in (1) will still be felicitous. But this flexibility does not detract from the overall point of the example.}\]
1.2.2 Defining Tenses and Aspects

Utilizing the concepts outlined above, it is possible to precisely formulate the functions of the major categories of tense and aspect in the world’s languages, in terms of how they situate the temporal elements with respect to each other. This is shown in Table 1. Prior to Klein (1994), definitions of tense and aspect had tended to be impressionistic, and therefore analytically inadequate. For instance, Comrie (1976:3) defines ‘aspects’ as “different ways of viewing the internal temporal constituency of a situation,” but later adds that “the perfect is rather different […] since it tells us nothing about the situation in itself, but rather relates some state to a preceding situation” (52). He defines perfect aspect as referring to “a past situation which has present relevance” (12). This definition captures an important intuition about uses of the perfect cross-linguistically, but it fails to provide diagnostics for identifying perfect constructions across languages. In contrast, Klein’s framework allows one to devise tests that categorically differentiate between the major categories ‘tense’ and ‘aspect,’ and between different tenses and aspects in a language. The distinction between tense and aspect under Klein’s approach is that tense locates TT in relation to TU, while aspect relates TT to TSit. For instance, as noted above, an imperfective aspect, such as the English progressive, locates TT within TSit. The utterance, When I walked into the room, John was reading a book indicates that the TSit of \{John read a book\} extends either side of the TT (which is here overtly specified with the adverbial clause when I walked into the room). In contrast, the use of past tense in this utterance indicates only that the TT precedes the TU, and says nothing directly about the TSit. This is demonstrated by the fact that it is perfectly felicitous to say When I walked into the room, John was reading a book, and he is still reading it now. If past tense situated TSit prior to TU, then it should not be possible to use past tense when the TSit is still ongoing, as in this example.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>situates TT with respect to TSit</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERFECTIVE</td>
<td>TT (partially) includes TSit {——[——]TSit}_TT or {——[——]TSit}_TT</td>
</tr>
<tr>
<td>IMPERFECTIVE</td>
<td>TT is fully included in TSit {——[——]TT——}_TSit</td>
</tr>
<tr>
<td>PERFECT</td>
<td>TT is after TSit {——}_TSit [_TT]</td>
</tr>
<tr>
<td>PROSPECTIVE</td>
<td>TT is before TSit [_TT] {——}_TSit</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tense</th>
<th>situates TT with respect to TU</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAST</td>
<td>TT is before TU [_TT] (_)TU</td>
</tr>
<tr>
<td>PRESENT</td>
<td>TT includes TU [_TT] (_)TU</td>
</tr>
<tr>
<td>FUTURE</td>
<td>TT is after TU (_)TU [_TT]</td>
</tr>
</tbody>
</table>

Table 1: Characterization of the major categories of tense and aspect (after Klein 1994)

2 Major Categories of Tense and Aspect in Imbabura Quichua

As background to the analysis of -shka, I briefly outline here the major tense and aspect distinctions in IQ. The description here should by no means be viewed as exhaustive; numerous other aspectual nuances, such as durative, habitual, and pluractional, can be expressed by verbal morphology in IQ, but these are not yet well-understood, and their analysis is beyond the scope of this paper (see
Cole 1982 for an overview).

### 2.1 Present Tense

Present tense in IQ is not marked by an overt morpheme. As shown in (3), it is indicated by the absence of overt past or future tense marking.³

(3) *Nyuka ri-xu-∅-ni yachachi-k wasi-man*

1SG go-IMPFV-PRES-1SG.SBJ teach-NMLZ house-ALL

‘I am going to school’

The subject agreement markers used in present tense are given in Table 2. They distinguish three persons, and two numbers in first and second person.

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-ni</td>
<td>-nchi</td>
</tr>
<tr>
<td>2</td>
<td>-ngi</td>
<td>-ngichi</td>
</tr>
<tr>
<td>3</td>
<td>-n</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: IQ subject agreement suffixes in present tense

### 2.2 Past Tense: -rka

The suffix *-rka* indicates simple past tense. It is compatible with imperfective and (null) perfective aspect marking.

(4) *Kayna nyuka ri-xu-rka-ni yachachi-k wasi-man*

yesterday 1SG go-IMPFV-PAST-1SG.SBJ teach-NMLZ house-ALL

‘Yesterday I was going to school’

(5) *Kayna nyuka ri-∅-rka-ni yachachi-k wasi-man*

yesterday 1SG go-PVF-PAST-1SG.SBJ teach-NMLZ house-ALL

‘Yesterday I went to school’

³ Abbreviations:

- ABL ablative
- ACC accusative
- AG agent
- ALL allative
- CAUS causative
- CL clitic
- COM/INSTR comitative/instrumental
- DIM diminutive
- DIST distal
- DS different subject subordinator
- EYE eyewitness
- FUT future
- GEN genitive
- HAB habitual
- IMPFV imperfective
- INDEF indefinite
- INDIR indirect evidential
- INF infinitive
- LIM limitative
- LOC locative
- LOB localizer
- NMLZ nominalizer
- PL plural
- PROX proximal
- PRES present
- PROX proximal
- PURP purposive
- SG singular
- SS same subject subordinator
- SBJ subject
- TOP topic
- NONEYE non-eyewitness
- ALL allative
- IMPFV imperfective
- PLUR plural actional
- PERF perfective
- PL proximal
As shown in Table 3, person marking in the past tense is the same as in present tense, except that third person subject agreement is null. This is illustrated in (6).

\[(6) \text{Pay} \text{ ri-xu-rka-0} \quad \text{yachachi-k} \quad \text{wasi-man} \]
\[3\text{SG} \quad \text{go-IMPFV-PAST-3SBJ} \quad \text{teach-NMLZ} \quad \text{house-ALL} \]
\[\text{‘S/he was going to school’} \]

The past tense suffix \(-rka\) is described by Cole (1982:144) as a simple past tense, but I will argue in §3.5 that it in fact has an additional entailment of ‘eyewitness’ evidentiality.

\[
\begin{array}{c|cc}
\text{SG} & \text{PL} \\
1 & -ni & -nchi \\
2 & -ngi & -ngichi \\
3 & -0 \\
\end{array}
\]

Table 3: IQ subject agreement suffixes in past tense

### 2.3 Two Future Tenses

There are two ways of indicating future tense in IQ. The synthetic future suffix \(-gri\), exemplified in (7), is homophonous with, and probably derived from, an analytic construction involving a nominalized verb in \(-k\) followed by an inflected form of the verb root \(ri-\) ‘go.’ This suffix attaches to the verb stem, and takes the agreement suffixes in Table 2.

\[(7) \text{Kaya} \quad \text{pay} \quad \text{ri-gri-n} \quad \text{yachachi-k} \quad \text{wasi-man} \]
\[\text{tomorrow} \quad 3\text{SG} \quad \text{go-FUT-3SBJ} \quad \text{teach-NMLZ} \quad \text{house-ALL} \]
\[\text{‘Tomorrow, s/he is going to go to school’} \]

The fusional future, in contrast, involves a separate set of agreement suffixes, shown in Table 4.

\[
\begin{array}{c|cc}
\text{SG} & \text{PL} \\
1 & -sha & -shun \\
2 & -ngi & -ngichi \\
3 & -nga \\
\end{array}
\]

Table 4: IQ fusional future subject agreement suffixes

The second person forms are identical in present and fusional future. Clauses with these forms and no other tense marking are therefore ambiguous between present and future interpretations, as shown in (8).

\[(8) \text{Kan} \quad \text{ri-xu-} \text{ngi} \quad \text{yachachi-k} \quad \text{wasi-man} \]
\[2\text{SG} \quad \text{go-IMPFV-2SBJ} \quad \text{teach-NMLZ} \quad \text{house-MAN} \]
a. ‘You are going to school’
b. ‘You will be going to school’
The Functions of -shka in Imbabura Quichua

The fusional future and the synthetic future have different modal overtones in terms of the speaker’s certainty about the future event. For example, the construction in (9) would be used if the speaker thinks that the ground will probably be dry tomorrow (for instance, because it has stopped raining and the sun has now come out), while the construction in (10) is appropriate if the speaker is certain that the ground will be dry tomorrow, for instance if he is taking action to dry it.

(9) Kaya-pa alpa chaki-shka ka-nga
tomorrow-GEN ground dry-PERF be-FUT:3 SBJ
‘By tomorrow, the ground will / should be dry’ (speaker is not certain)

(10) Kaya-pa alpa chaki-shka ka-gri-n
tomorrow-GEN ground dry-PERF be-FUT-3 SBJ
‘By tomorrow, the ground is going to be dry’ (speaker is certain)

In line with this, the first person plural form of the fusional future also has a hortative usage, while the synthetic future does not:

(11) Ri-shun yachachi-k wasi-man
go-FUT:1PL SBJ teach-NMLZ house-ALL
a. ‘We will go to school’
b. ‘Let’s go to school’

(12) Ri-gri-nchi yachachi-k wasi-man
go-FUT-1PL SBJ teach-NMLZ house-ALL
a. ‘We are going to go to school’
b. *‘Let’s go to school’

2.4 Imperfective Aspect: -xu

The suffix -xu is a marker of imperfective aspect (TT included in TSit). It co-occurs with all tenses, and is compatible with verbs of all Aktionsarten, including stative (though it seems not to co-occur with the copula ka-). This suggests that it denotes a very general imperfective aspect (and not, for example, a progressive aspect like English be X-ing, which is generally incompatible with stative verbs). There is no overt perfective marker in IQ. The absence of imperfective -xu is interpreted as perfective.

(13) Kayna nyuka chaya-kpi=ka, kan zhukshi-0-rka-ngi
yesterday 1SG arrive-DS=TOP 2SG leave-PFV-PAST-2SG SBJ
‘Yesterday when I arrived, you left’

(14) Kayna nyuka chaya-kpi=ka, kan zhukshi-xu-rka-ngi
yesterday 1SG arrive-DS=TOP 2SG leave-IMPFV-PAST-2SG SBJ
‘Yesterday when I arrived, you were leaving’
3 Analysis of -shka in Imbabura Quichua

Having outlined the general system of tense and aspect in IQ, we are now in a position to examine in detail the uses of the verbal suffix -shka. There are at least five distinct functions of -shka in IQ (several of which are structurally isomorphic): four nonfinite derivations and one finite verbal inflection. I will discuss these in turn, then in §4, I will propose a diachronic path of development that unites these functions.

3.1 Perfect Participle

The use of -shka to derive a perfect participle in IQ is illustrated in (15). The periphrastic perfect is created by attaching -shka to the verb stem, and combining this with the copula ka-, inflected for subject person and number, and tense. As shown, the copula can occur in past, present or future tense.

(15) a. Kayna nyuka shina-shka ka-rka-ni tanda-ta
    yesterday 1SG make-PERF be-PAST-1SG.SBJ bread-ACC
    ‘Yesterday I had made a cake’

b. Kunan nyuka shina-shka ka-0-ni tanda-ta
    today 1SG make-PERF be-PRES-1SG.SBJ bread-ACC
    ‘Today I have made a cake’

c. Kaya nyuka shina-shka ka-sha tanda-ta
    tomorrow 1SG make-PERF be-FUT-1SG.SBJ bread-ACC
    ‘Tomorrow I will have made a cake’

As defined in Table 1 above, tense locates TT with respect to TU. Since each simple clause has just one TT and one TU, the compatibility of -shka in the above constructions with overt past and future marking on the copula indicates that here -shka indeed encodes an aspect and not a tense. This is corroborated by the fact, illustrated in (16), that when the copula does not have past tense marking, the clause is ungrammatical with a past time adverbial, such as kayna ‘yesterday.’ A temporal adverb such as ‘yesterday’ restricts the TT of an utterance. Since ‘yesterday’ is deictic, and necessarily refers to a TT prior to TU, it is felicitous only if TT precedes TU. This is the case with past tense (defined as TT before TU), but not with present perfect aspect (where TT is after TSit, but includes TU). Hence the behavior of -shka in this periphrastic construction is fully compatible with an aspect marker.

(16) *Kayna nyuka shina-shka ka-0-ni tanda-ta
    yesterday 1SG make-PERF be-PRES-1SG.SBJ bread-ACC

That it marks ‘perfect’ aspect in particular is evident from the fact that it locates TT after TSit. For instance, the utterance in (17) is only true if it is the case that Maria arrived after Pepe left.

(17) Maria chaya-kpi=ka, Pepe zhukshi-shka ka-rka-0
    Maria arrive-DS=TOP Pepe leave-PERF be-PAST-3SG.SBJ
a. ‘When Maria arrived, Pepe had left’
b. *‘When Maria arrived, Pepe left’

If Maria’s arrival coincides with, or precedes, Pepe’s departure, the simple past must be used in the main clause:

(18) Maria chaya-kpi=ka, Pepe zhukshi-0-rka-0
Maria arrive-DS=TOP Pepe leave-PFV-PAST-3SG.SBJ

a. *‘When Maria arrived, Pepe had left’
b. ‘When Maria arrived, Pepe left’

The temporal structures of the two situations are illustrated below. The TT is Maria’s arrival, which is overtly set by the adverbial clause Maria chayakpi. IQ zhukshi- ‘leave’ is a two-state verb, therefore the TSit involves two elements: the SS, when Pepe is present and active in becoming absent, and the TS, when Pepe is absent. The structures in (19) correspond to the main clause in (17), while those in (20) correspond to the main clause in (18).

(19) PAST PERFECT
{ ——— }SS+++[++]TT++TS ( )TU or { ——— }SS+++++++TS [ ]TT ( )TU

(20) SIMPLE PAST (PERFECTIVE)
[ { ——— }SS+++]TT++TS ( )TU or { ———— }SS+++]TT+++TS ( )TU

It is clear, therefore, that the analytic construction with -shka expresses perfect aspect. However, due to a general tendency in IQ (and in many other Quechua languages) for copula ka- ‘be’ to be elided in present tense third person contexts, the present perfect in third person most commonly surfaces as in (21).

(21) Pay ri-shka
3SG go-PFV
‘S/he has gone’

In such cases it is isomorphic with the inflectional -shka construction discussed in §3.5 below. This fact is crucial for the diachronic account presented in §4 below, and will be discussed further there.

3.2 Anterior Aspect Subordinator

In subordinate clauses — including reason clauses, complement clauses, and relative clauses — the use of -shka indicates that the TSit of the subordinate clause is prior to the TT of the main clause. For instance, the -manda reason clause in (22) can only refer to a situation prior to the situation of the main clause.

(22) Mishki tanda-ta miku-shka-manda, nyuka kushi=zha ka-0-ni
sweet bread-ACC eat-PFV-ABL 1SG happy=LIM be-PRES-1SG.SBJ

a. ‘Because I have eaten cake, I am happy’
b. *‘Because I am eating / will eat cake, I am happy’
The verb form in -shka here is clearly nonfinite, as it does not take person marking (although the same construction can also occur with person marking; in that case I assume the verb in the -manda clause is finite).

A further example of anterior aspect -shka is in (23), where it attaches to the verb in a complement clause (which is marked with accusative case).

(23) Kayna chusku pacha-pi riku-θ-rka-ni pichari kaspi-ta
    yesterday four time-LOC see-PFV-PAST-1SG.SBJ somebody stick-ACC
    faki-shka-ta
    break-PERF-ACC
    ‘Yesterday at four o’clock, I saw that somebody had broken the stick’

3.3 Passive Participle

According to Cole (1982:133), IQ has two periphrastic passive constructions, in addition to at least one morphological passive.4 The first involves a non-finite verb in -y plus the auxiliary tuku- ‘become, finish.’ This construction implies a change of state, and is restricted to occurring with animate passive subjects. The -shka ka- passive, in contrast, does not necessarily imply change of state, and has no restrictions on its subject. These differences are illustrated in (24-25).

(24) a. Maria=ka Jose xuya-y tuku-rka-θ
    Maria=TOP Jose love-INF become-PAST-3SG.SBJ
    ‘Maria came to be loved by Jose’

b. Maria=ka Jose xuya-shka ka-rka-θ
    Maria=TOP Jose love-PERF be-PAST-3SG.SBJ
    ‘Maria was loved by Jose’

(25) a. *Aycha=ka (misi) miku-y tuku-rka-θ
    meat=TOP cat eat-INF become-PAST-3SG.SBJ

b. Aycha=ka (misi) miku-shka ka-rka-θ
    meat=TOP cat eat-PERF be-PAST-3SG.SBJ
    ‘The meat was eaten (by the cat)’

3.4 Resultative Nominalizer

In all three of the above constructions, -shka attaches to a verb root or stem, and derives a nonfinite form. But the derived form nonetheless has verbal properties, such as being able to assign accusative case to an object. -shka has a fourth derivational function, in which it derives an element that is clearly nominal, rather than verbal. Cole (1982:148) refers to this construction as

4 We have only been able to elicit the -y tuku- periphrastic passive, and have had difficulty getting consultants to accept or produce the -shka ka- passive. Nonetheless, it does occasionally appear in texts. It is likely that the very high functional load of -shka ka- forms, and the existence of several alternative passive constructions, makes the use of this passive dispreferred. Data in this section is adapted from Cole (1982).
The Functions of -shka in Imbabura Quichua

He notes that it emphasizes “the present result of a past situation,” in contrast to the perfect, which “focuses on the past action.” The deverbal -shka form, like other nominals in the language, can inflect for case (26), can modify other nominals (27), and can appear as copula complement (28).

(26) [Chinga-chi-shka-ta]NP maska-xu-∅-ni
get.lost-CAUS-PERF-ACC search.for-IMPFV-PRES-1SG.SBJ
‘I am looking for my lost (thing)’

(27) Kay [maki-pi awa-shka findu-ta]NP riku-xu-∅-ni
PROX hand-LOC make-PERF textile-ACC see-IMPFV-PRES-1SG.SBJ
‘I see this hand-made fabric’

(28) Kaya-pa alpa [chaki-shka]COPCOMP ka-nga
tomorrow-GEN ground dry-PERF be-FUT:3SG
‘By tomorrow the ground will be dry’

Although Klein’s framework was designed to analyze only the verbal domain, it can be extended to characterize the function of the IQ resultative nominal. The resultative can be described as situating TT within TS of the LC, in contrast to the perfect participle, which specifies only that TT is after SS, and is agnostic about whether TT falls within or after TS. These two respective temporal structures are diagrammed in (29) and (30) below.

(29) RESULTATIVE
   {——}SS++[++]TT++TS

(30) PERFECT
   {——}SS++[++]TT++TS or {——}SS+++++TS [ ]TT

This formulation reflects Cole’s intuitive characterization of the perfect as focussing on the past action, and the resultative as focussing on the resulting state. That -shka is indeed a resultative, and not simply a nominalizer, is evidenced by the fact that the utterance in (28) above is felicitous only if the ground has at some point been wet; that is, chakishka does not simply denote the state of being dry, rather it refers to the TS of the verb ‘to dry.’

A further distinction between perfect and resultative forms is their argument structure. While the perfect form maintains the argument structure of the verb — whether transitive or intransitive — resultative forms are necessarily intransitive, and therefore can retain only one of the arguments of a transitive verb. The retained argument, predictably, is the more affected participant, typically the object. Hence resultative nominalizations follow an absolutive pattern, denoting the subject of intransitive verbs and the object of transitive verbs. For example, the verb root shita- ‘throw’ takes an agent subject and a patient object. In (31), the resultative modifier shitashka refers (metaphorically) not to the ‘throwers,’ but to the ‘thrown.’ In contrast, the verb root kazara- ‘marry’ (presumably from Spanish casarse) is intransitive, and the resultative nominalization refers to the erstwhile subject.
(31) Kay shuxa shita-shka xari-kuna warmi-kuna…  
    PROX INDEF throw-PERF man-PL woman-PL  
    ‘Those homeless men and women…’

(32) Chayra=zha kazara-shka-kuna shamu-rka-nchi  
    DIST=LIM marry–PERF-PL come-PAST-1PL.SBJ  
    ‘When we were first married we came’

This pattern of argument structure with transitive verbs offers a clear link between resultative and passive constructions, and this will be discussed further in §4 below.

3.5 Finite Verbal Inflection

In all of the functions outlined so far, -shka derives a nominal or a nonfinite verbal element. But -shka can also function as a verbal inflection, in which case it attaches to the verb stem and immediately takes person marking. The person markers are the same as those used with past tense suffix -rka (shown in Table 3 above). This is the function that Cole (1982) labels ‘perfect aspect,’ noting that a verb inflected with -shka can also take past or future tense marking. While our consultants readily accept and produce verbs in -shka-rka, they do not accept forms with -shka followed by any future marking.

(33) a. Pay shina-shka-rka-0 tanda-ta  
    3SG make-SHKA-PAST-3SBJ bread-ACC  
    ‘S/he had baked a cake’

b. Pay shina-shka-0 tanda-ta  
    3SG make-SHKA-3SBJ bread-ACC  
    ‘S/he baked a cake’

c. *Pay shina-shka-nga tanda-ta  
    3SG make-SHKA-FUT:3SBJ bread-ACC

This does not, of course, invalidate Cole’s claim that -shka is a perfect aspect. It could be that -shka marks perfect aspect, but is restricted to occurrence in present and past tenses; and in fact, I argue in §4 that this is an important stage in the diachronic development of modern-day -shka. However, I will present data here that shows inflectional -shka in IQ is not in fact a perfect aspect; rather it is a past tense that is restricted to descriptions of situations which the speaker did not witness.

In Table 1 above are presented Klein’s cross-linguistic definitions of the major classes of tense and aspect. According to this typology, perfect aspect locates TT after TSit, as diagrammed in (34).

(34) PERFECT ASPECT  
    {———}TSit  [ ]TT

If inflectional -shka is a perfect aspect, therefore, its function should accord with the diagram in (34). This was shown to be the case for the periphrastic -shka ka- construction in (17–20) above.
The Functions of -shka in Imbabura Quichua

In contrast, constructions with inflectional -shka do not conform to the temporal structure for perfects.5

(35) Maria chaya-kpi=ka, kan zhukshi-θ-shka-ngi
    Maria arrive-DS=TOP 2SG leave-PFV-SHKA-2SG.SBJ
    a. *‘When Maria arrived, you had left’
    b. ‘When Maria arrived, you left’

The utterance in (35) is not felicitous if TT (Maria’s arrival) follows TSit (the leaving event). In contrast, it is felicitous if TT overlaps with TSit; that is, if Maria’s arrival coincides with the addressee’s departure. Recall from Table 1 that this temporal structure constitutes ‘perfective’ aspect. That -shka itself is not a perfective marker, however, is shown by its ability to freely combine with imperfective -xu.6

(36) Maria chaya-kpi=ka, kan zhukshi-xu-shka-ngi
    Maria arrive-DS=TOP 2SG leave-IMPFV-SHKA-2SG.SBJ
    ‘When Maria arrived, you were leaving’

The relation between TT and TSit expressed in (36) is diagrammed in (37), where TT is Maria’s arrival, and TSit corresponds to the SS of the two-state predicate {leave}.

(37) {—-[—]TT—}TSit+++++++++++++++++

It is clear, therefore, that inflectional -shka encodes neither perfect nor perfective aspect. The clue to its actual function is given in the translations of (36) and (37), both of which are past tense. This past tense is not provided by the subordinate clause Maria chaya-kpi=ka, which is nonfinite and therefore does not express tense. This adverbial clause is fully compatible with present and future marking in the main clause:

(38) Maria chaya-kpi=ka, kan zhukshi-ngi
    Maria arrive-DS=TOP 2SG leave-2SG.SBJ
    ‘When / If Maria arrives, you leave / will leave’

It seems, therefore, that the past tense reference is expressed by -shka. We can test this by examining the co-occurrence of -shka with temporal adverbs. If -shka is a past tense marker, it should be compatible with past time adverbs, and incompatible with present and future time adverbs. The data in (39) show that this is indeed the case, confirming the status of inflectional -shka as past tense.

---

5 I present examples in second rather than third person here to avoid ambiguity. As noted above, in third person present tense contexts, the copula is often omitted from the periphrastic perfect (and from other constructions). Inflectional -shka in third person is identical to periphrastic -shka in third person present tense with the copula omitted, and therefore this surface form is ambiguous between the two interpretations. Inflecting for first or second person avoids this ambiguity.

6 Although (Comrie 1976:24) claims that “imperfectivity is not incompatible with perfectivity, and that both can be expressed if the language in question possesses the formal means to do so,” within the framework followed here, perfective and imperfective aspects are mutually exclusive.
Since, as shown in §2.2, -rka is also a past tense marker, the question remains how -shka and -rka differ from each other. Some Quechua languages (e.g., Ancash; Adelaar and Muysken 2004:224) have a distinction between recent and remote past and this is a plausible assumption for IQ, given that perfects are known to develop into recent past markers cross-linguistically. However, as Cole (1982:148) notes, -shka is especially frequent in traditional narratives, the events of which are temporally remote. This suggests that -shka is not a recent past marker. The data in (40) and (41) show that in fact the relevant parameter of variation between -rka and -shka is source of evidence (eyewitness — non-eyewitness). Verbs in -rka are felicitous only to report a situation which the speaker witnessed first-hand, while -shka must be used to describe a situation that was not witnessed first-hand.

(40) a. Maria chaya-kpi=ka, kan zhukshi-0-rka-ngi  (*ni-n)  
Maria arrive-DS=TOP 2SG leave-PFV-PAST.EYE-2SG.SBJ say-3SBJ  
‘When Maria arrived, you left (*it is said)’

b. Maria chaya-kpi=ka, kan zhukshi-0-rka-ngi.  Kan-ta  
see-PFV-PAST.EYE-1SG.SBJ  
‘When Maria arrived, you left. I saw you.’

(41) a. Maria chaya-kpi=ka, kan zhukshi-0-shka-ngi  ni-n  
Maria arrive-DS=TOP 2SG leave-PFV-PAST.NONEYE-2SG.SBJ say-3SBJ  
‘When Maria arrived, you left, it is said’

b. Maria chaya-kpi=ka, kan zhukshi-0-shka-ngi.  (*Kan-ta  
Maria arrive-DS=TOP 2SG leave-PFV-PAST.NONEYE-2SG.SBJ 2SG-ACC riku-0-shka-ni.)  
see-PFV-PAST.NONEYE-1SG.SBJ  
‘When Maria arrived, you left. (*I saw you.)’

This evidential division of labor between -rka and -shka neatly elucidates why -shka is prevalent in narrative discourse. We need no longer posit, as Cole (1982:148) must, that traditional narratives are somehow of greater ‘present relevance’ than other types of discourse referring to the past. In §3.6, I show that the distribution of -rka and -shka in different discourse types neatly follows from their functions as ‘eyewitness’ and ‘non-eyewitness’ past tenses respectively.
One puzzle remains: if inflectional -shka is a past tense, how can it combine with past tense -rka in verb forms such as shinashkarka in (33a) above? The answer is simply that in this construction, -shka is not a past tense, but rather retains its 'perfect' meaning. That is, -shkarka is a 'past perfect' inflectional suffix which locates TT before TU and after TSit. This is further illustrated in (42), which can only mean that TSit (the departure) occurred before TT (Maria’s arrival); it is infelicitous to describe a situation where TT and TSit overlap. Likewise, it cannot express a temporal relation where TT includes or follows TU; it explicitly situates TT before TU. The temporal structure denoted by a verb in -shkarka is diagrammed in (43).

(42) Maria chaya-kpi=ka, kan zhukshi-0-shkarka-angi
   Maria arrive-DS=TOP 2SG leave-PFV-PAST.PERF-2SG.SBJ
   a. ‘When Maria arrived, you had left’
   b. *‘When Maria arrived, you left’
   c. *‘When / If Maria arrives, you leave / will leave’

(43) \{ --- \}TSit [ ]TT ( )TU

Table 5 summarizes the IQ system of past tense and perfect aspect marking, as described and exemplified in this paper.

<table>
<thead>
<tr>
<th>Form</th>
<th>Gloss</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>-rka</td>
<td>‘simple past eyewitness’</td>
<td>TT before TU speaker witnessed TSit</td>
</tr>
<tr>
<td>-shka</td>
<td>‘simple past non-eyewitness’</td>
<td>TT before TU speaker did not witness TSit</td>
</tr>
<tr>
<td>-shkarka</td>
<td>‘past perfect’</td>
<td>TT before TU TT after TSit</td>
</tr>
<tr>
<td>-shka (ka-)</td>
<td>‘perfect aspect’</td>
<td>TT after TSit</td>
</tr>
</tbody>
</table>

Table 5: IQ past tense and perfect aspect constructions

3.6 The distribution of -rka and -shka in discourse

The IQ text corpus we compiled comprises a variety of discourse types, including traditional narratives, personal narratives, procedural texts, and conversations. On close examination of these texts, it is clear that Cole is correct in noting a preponderance of -shka constructions in traditional narratives. But I argue that it is not the discourse type itself that determines which verb form will be used; rather the discourse type tends to determine whether or not the events related were directly witnessed by the speaker, and this is what determines whether -shka or -rka is used. Supporting this hypothesis is the fact that the ‘non-eyewitness’ meaning of -shka is almost obligatorily reinforced in discourse by the reportative marker nin (‘say-3SBJ’), or the indirect validator =shi. This is illustrated in the following passage from a traditional text told by MC.
In contrast, -rka forms do not occur with either nin or =shi, as shown in (45), from a personal narrative by the same speaker.

(45) Bini-0-rka-nchi wagra-kuna-wan zhama-kuna-wan
grow.up-PFV-PAST.EYE-1PL.SBJ cow-PL-COM/INSTR sheep-PL-COM/INSTR
atal-pa-kuna-wan misi-kuna-wan. Bini-0-rka-ni
chicken-PL-COM/INSTR cat-PL-COM/INSTR grow.up-PFV-PAST.EYE-1SG.SBJ
sara-ta tarpu-shpa, papa-ta tarpu-shpa. Ri-ka-rka-ni nyuka
corn-ACC plant-SS potato-ACC plant-SS go-HAB-PAST.EYE-1SG.SBJ 1SG
yaku-man taksha-ngapa nyuka zhachapa-ta maki-wan larka-pi
water-ALL launder-SS:PURP 1SG clothes-ACC hand-COM/INSTR river-LOC
taksha-k ri-ka-rka-ni.
launder-NMLZ go-HAB-PAST.EYE-1SG.SBJ

‘We grew up with cows, with sheep, with chickens, with cats. I grew up planting corn, planting potatoes. I used to go to the water to wash my clothes, I would go to wash them by hand in the river.’

When -shka is used in its perfect function, rather than its non-eyewitness past function, it does not collocate with nin or =shi, but rather tends to collocate with nya ‘already, still.’ This is illustrated in the following passage, from a conversation between MC and ACO (the speaker here is MC). The analysis of -shka here as ‘present perfect’ and not ‘past non-eyewitness’ is supported by the fact that the surrounding discourse is in present tense, and that the events related were experienced first-hand by the speaker.

(46) Nyuka ushi-gu-kuna-ta nya na kawsa-xu-0-n=chu
1SG daughter-DIM-PL-ACC already NEG live-IMPFV-PRES-3SBJ=NEG.CL
nyuka-wan, nya pay-kuna=ka mutari-shka, shuk ludu-pi=mi
1SG-COM/INSTR already 3-PL=TOP move.house-PERF INDEF place-LOC=DIR
live-PLUR-3SBJ 3-PL already school-ACC finish-CAUS-PERF already
Pay-kuna-pa shuk rura-y-wan kawsa-naxu-n.
3-PL-GEN INDEF do-NMLZ-COM/INSTR live-PLUR-3SBJ
The Functions of -shka in Imbabura Quichua

‘My daughters don’t live with me anymore, they have already moved out, they live in other places. They have already finished school. They have jobs.’

The patterns of use and collocation found across the different text types in the corpus reflect different meanings and usages of -shka and -rka. The data presented in this section, which are representative of the patterns found throughout the corpus, support the analysis of -rka as a past eyewitness marker, and -shka as variously ‘past tense non-eyewitness’ and ‘perfect.’

4 A Diachronic Explanation

In this section, I will set forth a possible diachronic explanation for the range of functions fulfilled by -shka in IQ. A similar range of functions is described for other Quechua languages. For example, Weber (1989) reports that verbal inflection -shka (apparently truncated from *-shqa ka-) in Huallaga Quechua (HQ) functions as a perfect aspect and as a past tense. He reports that it is displacing -ra (cognate to IQ -rka) as a marker of simple past. In addition he notes that periphrastic -sha ka- (also from *-shqa ka-) functions as both a perfect and a passive. -sha is additionally reported to have a substantivizing function in HQ. Adelaar (1977) reports for Tarma Quechua (TQ) that the main function of -sha is to form stative participles, such as wanusha ‘dead’ (wanu- ‘die’), punusha ‘asleep’ (punu- ‘sleep’). It also is used in subordinate clauses, but not necessarily with an anterior meaning, and he notes a very occasional perfect use. Adelaar and Muysken (2004) note that the existence of a so-called ‘sudden discovery’ tense, commonly expressed with a cognate of IQ -shka, is an areal phenomenon, having even spread into Andean Spanish (though using native Spanish morphemes). They note that it is not restricted to past tense, which suggests they are describing a perfect aspect function. Likewise, Faller (2004:46) describes the function of Cuzco Quechua (CQ) -sqa as “a spatio-temporal deictic which specifies that the described eventuality e [Klein’s TSit – JCK] is not located within the speaker’s perceptual field at topic time.” This seems to accord with the function of IQ -shka, in that it is restricted to contexts where the TSit was not witnessed by the speaker. In South Conchucos Quechua (SCQ; Hintz 2008), the suffix -sha ‘past perfective’ is restricted to third person contexts, and a suppletive form -ru occurs with local subjects and objects. Although Hintz does not mention any evidential overtones to these perfective markers, their distribution suggests that an eyewitness — non-eyewitness distinction may be, or have previously been, part of their meaning. Both -sha and -ru are used to express situations that are relatively recent in time, while -rqa (cognate with IQ -rka) denotes more distant events (Hintz 2007).

Given the variety of functions of IQ -shka, and considering the functions of its cognates throughout the Quechua family, I propose that its diachronic development proceeded as shown in (47). The source of this chain of developments is claimed to be either the passive or the resultative construction. There is no evidence I am aware of to favor either direction of change between passive and resultative. The development from resultative to perfect, on the other hand, is well-attested cross-linguistically. For example, this pathway is explicitly documented for English and Romance (Bybee et al. 1994:68). In the absence of documentation, I assume that the same direc-

---

7 I thank Tom Recht for pointing out that the resultative may have been the source for the perfect, and not necessarily the converse; and I thank Andrew Garrett for directing me to the relevant literature.
tionality applies for IQ -shka. In the following sections, I discuss each stage of grammaticalization in turn.

(47) PROPOSED DIACHRONIC PATH FOR THE DEVELOPMENT OF IQ -shka

Passive Participle
\[\uparrow\]
Resultative Participle
\[\downarrow\]
Perfect Participle
\[\downarrow\]
Inflectional Perfect
\[\downarrow\]
Inflectional Past Tense: Non-eyewitness

4.1 Passive and Resultative

As noted in §3.4, resultative nominalizations exhibit an absolutive alternation, where the denoted or modified argument is the object, rather than the subject, of a transitive verb. In addition, these nominalizations occur far more commonly with transitive verbs than with intransitive, presumably because the object of a transitive verb is typically more affected than the subject of an intransitive verb. Given this, the resultative nominalizations almost exclusively modify the erstwhile object of a transitive verb, and this, of course, is also the function of a passive. As shown in (48), a bridging context for the grammaticalization of passive to resultative (or vice versa) is therefore provided by predicative result nominalizations from transitive verbs.

(48) Aycha=ka yanu-shka ka-rka-θ
    meat=TOP cook-SHKA be-PAST-3SG.SBJ
a. Passive reading: ‘The meat got cooked’
\[
\{——\}_{SS++[TT++]}_{TS} ( )_{TU}
\]
b. Resultative nominalization reading: ‘The meat was cooked’
\[
\{——\}_{SS++} [TT+][TS] ( )_{TU}
\]

The temporal structures of these two interpretations of the surface string Aychaka yanushka karka differ in whether the TT includes SS or not. For resultative interpretation it does not; for passive it does. All that is needed for one of these constructions to be reinterpreted as the other, therefore, is for the leftmost boundary of TT to shift. If -shka originally derived a passive participle, and the resultative function was innovated, predicative modifiers in -shka must have subsequently been extended, by analogy with other, non-derived modifiers, to the attributive (49) and substantive (50) contexts they occur in today.

(49) Yanu-shka aycha-ta miku-xu-θ-ni
    cook-SHKA meat-ACC eat-IMPFV-PRES-1SG.SBJ
    ‘I am eating cooked meat’
The Functions of -shka in Imbabura Quichua

(50) Taksha-shka-kuna-ta angu-pi warku-xu-0-ni
      launder-SHKA-PL-ACC line-LOC hang-IMPFV-PRES-1SG.SBJ
‘I am hanging the laundered (clothes) on the line’

4.2 Resultative to Perfect

In contrast to the passive–resultative case just described, transitive constructions do not provide a bridging context between resultative and perfect interpretations. There are two reasons for this: first, adjectival predicates cannot be transitive; and second — as noted above — adjectives derived from transitive stems modify the object, not the subject, of the erstwhile verb, thus providing a bridging context with passive, but not perfect. Resultative–perfect bridging contexts instead arise in periphrastic perfect constructions with intransitive verbs. As shown in (51–52), such a construction is isomorphic with a predicative adjective construction. The two interpretations are differentiated only by the fact, as demonstrated in §3.4, that resultative adjectives place TT in TS, while perfects simply place it after SS. Since these potentially different temporal structures more often than not describe identical states of affairs, speakers may easily reanalyze one as the other.

(51) Nyuka kushi ka-0-ni
      1SG happy be-PRES-1SG.SBJ
‘I am happy’

(52) Nyuka zhukshi-shka ka-0-ni
      1SG leave-SHKA be-PRES-1SG.SBJ
a. Resultative adjective reading: ‘I am gone’
   {——}SS++[++]TT++TS
b. Passive participle reading: ‘I have left’
   {——}SS++[++]TT++TS or {——}SS+++++TS [ ]TT

Once semantic reanalysis had taken place, the construction could subsequently be extended from intransitive to transitive contexts, such as that in (53), via analogy.

(53) Tanda-ta miku-shka ka-0-ni
      bread-ACC eat-SHKA be-PRES-1SG.SBJ
‘I have eaten the bread’

4.3 Derivation to Inflection

The first stage in the development of -shka from periphrastic perfect to non-eyewitness past tense is the grammatical reanalysis of -shka as an inflectional suffix. The optionality of the copula in third person present tense contexts has been assumed to have been the catalyst for this reanalysis. Since third person agreement is null after past tense marker -rka, once the copula is elided from the perfect construction, as in (54a), it appears isomorphic with a third person past tense verb, as in (54b), (although in other persons and tenses it remains distinct).
On analogy with -rka forms, -shka was reanalyzed as a verbal inflection, and extended to all persons. At this stage, it is assumed that -shka was still an aspect marker, and therefore it could combine with other tenses. Since it developed from a present perfect construction, where TU follows TSit, it may have been restricted to present and past perfect, and been disallowed in future contexts, where TSit usually follows TU. Alternatively, this restriction may have developed later. A hypothesized partial paradigm for this stage is given in (55).

(55) a. *Punyu-shka*  
    sleep-PERF  
    ‘S/he has slept’  
b. *Punyu-rka-0*  
    sleep-PAST-3SBJ  
    ‘S/he slept’  
c. *Punyu-shka-rka-0*  
    sleep-PERF-PAST-3SBJ  
    ‘S/he had slept’  
d. *Punyu-shka-nga*  
    sleep-PERF-FUT:3SBJ  
    ?‘S/he will have slept’

4.4 Aspect to Tense

The final stage in this grammaticalization trajectory is semantic reanalysis of the inflectional construction from perfect aspect to past tense, a development that is commonly attested cross-linguistically (see Bybee et al. 1994:81). The motivation for this is clear once we examine the temporal make-up of these constructions, diagrammed in (56–57).

(56) **Present Perfect**  
    \[
    \{---\}_{TSit} \quad [ \quad (TU) \quad ]_{TT}
    \]

(57) **Past Tense (Perfective)**  
    \[
    [ \quad \{---\}_{TSit} \quad ]_{TT} \quad (\quad )_{TU}
    \]

It can be seen that both present perfect aspect and simple past tense situate TSit before TU, though they do this in different ways. A perfect aspect becomes a past tense when speakers reinterpret a construction denoting the relation in (56) as denoting that in (57); that is, they analyze the construction as situating TT before TU, and TSit within TT, rather than TSit before TT, and TU within
TT. Bybee et al. (1994:86) view this process as a generalization of meaning, whereby a construction denoting ‘past event with current relevance’ comes to denote simply ‘past event.’ Within the framework employed here, this path can likewise be characterized as semantic broadening. Since the source construction is ‘present perfect’ and the innovated construction ‘simple past,’ the construction goes from expressing a tense–aspect combination (situating TT, TSit, and TU in relation to each other), to simply expressing a tense relation (which only situates TT in relation to TU). However, a certain semantic narrowing accompanies this grammatical change in IQ, as the innovated past takes on a non-eyewitness entailment. Perfect constructions generally have a non-eyewitness implicature, which comes about because the post-time of the event is referred to, rather than the event itself. In IQ, this implicature was strengthened to an entailment in the new past tense form. Presumably this pragmatic strengthening was helped along by the fact that -shka is in competition with past tense suffix -rka, and there is pressure for different forms to have different functions. This leads to the division of labor between past tense eyewitness -rka and past tense non-eyewitness -shka in modern IQ.

5 Conclusion

Systems of tense and aspect in the Quechua language family seem to be particularly complex and challenging to analyze. As a result, their description has often been vague, and the meanings and functions of various constructions have been characterized only in the broadest terms. In this paper, I have attempted to fill in some of the gaps in our understanding of tense and aspect in Quechua, by providing a precise characterization of the functions of the IQ verbal suffix -shka. Five morpho-syntactically and/or semantically distinct functions of -shka are identified: (i) it forms perfect participles, which enter into a periphrastic perfect aspect construction with copula ka-; (ii) it expresses anterior aspect in subordinate clauses, including relative, adverbial, and complement clauses; (iii) it creates a passive participle, which, like the perfect participle, combines with copula ka- in a periphrastic construction; (iv) it derives resultative nominals, which, like other IQ nominals, can act as substantives, modifiers, or copula complements; and (v) it is a verbal inflection expressing non-eyewitness past tense. The first four of these uses function essentially as described by Cole (1982) (except that Cole characterizes the subordinate clause use as a tense). The main purpose of this paper has been to argue, using Klein’s framework for analyzing tense and aspect, that -shka as a verbal inflection is in fact a tense rather than an aspect as Cole (1982) claims. That this is the case is shown by the co-occurrence of -shka with past time adverbs such as kayna ‘yesterday,’ which should be incompatible with a present perfect aspect, and by the fact that -shka cannot be used to locate TT after TSit. A further aim of the paper is to show that inflectional -shka entails non-eyewitness evidentiality; that is, it can be used just in case the speaker did not witness the situation expressed in the predicate (TSit). The non-eyewitness past tense meaning of -shka neatly explains its distribution in discourse: verbs inflect in -shka in traditional narratives, the events of which were not witnessed by the speaker, and in eyewitness past -rka in personal narratives, which were experienced first-hand by the speaker.

Many of the functions fulfilled by -shka in IQ are likewise fulfilled by a single morpheme in numerous other languages. For instance, the English past participle in -ed or -en is a perfect, a

8 Faller (2004) stands out as a notable exception to this generalization.
resultative, and a passive, and in some non-standard varieties also a simple past. This suggests that the diachronic path of development that -shka has followed is typologically common. Bridging contexts for each stage of development are evident, but the directionality of some of the changes in Quechua is still open to investigation.

References


Jessica Cleary-Kemp  
University of California, Berkeley  
Department of Linguistics  
1203 Dwinelle Hall  
Berkeley, CA 94720-2650  
jessck@berkeley.edu
Hierarchies, Subjects, and the Lack Thereof in Imbabura Quichua Subordinate Clauses

CLARA COHEN
University of California, Berkeley

1 Introduction

One way to simplify grammatical descriptions and comparisons is to set up a hierarchy of properties associated with the constructions under analysis. Rather than specifying piecemeal the set of behaviors associated with constructions X and Y, the right hierarchy allows a linguist simply to say that construction X has behavior N and everything below it on the hierarchy, while construction Y has behavior N+2 and everything below it. These hierarchies can be implicational tendencies that describe cross-linguistic generalizations (for example, markedness hierarchies such as /t/ < /p/ < /k/, by which languages that have the phoneme /k/ will also have /p/ and /t/; Maddieson 1984), or they can be a language-specific hierarchy that simplifies description of a set of structures. This paper deals with two hierarchies of the second type. Specifically, I argue that two hierarchies that have been proposed to account for the behavior of non-canonical subjects in Imbabura Quichua do not uniformly hold true in the face of other data.

Imbabura Quichua (IQ) is a Quechua language spoken by perhaps a hundred thousand people in the Imbabura province of the Ecuadorean Andes.1 Of interest in this investigation are its multiple types of non-canonical subjects. Non-canonical subjects are arguments that behave similarly to subjects in some ways — e.g., in imperative constructions, or as antecedents for reflexive pronouns — but not all. Usually non-canonical subjects are distinct from canonical subjects at least in their morphological properties, taking different case-marking or governing different patterns of verb agreement, and often they diverge in some other syntactic properties as well (Onishi 2001). In a language like IQ, which has multiple types of non-canonical subjects, a hierarchical arrangement of subjecthood properties become extremely useful, because the right hierarchy allows the linguist to describe much more concisely the behavior of each non-canonical subject in a principled way. The linguist can simply refer to a subset of properties on the hierarchy, rather than giving a piecemeal listing of which behavior each argument type does or does not exhibit.

The first hierarchical description of non-canonical subjects discussed in this paper is proposed by Onishi (2001) and Hermon (2001). It describes which types of syntactic behaviors non-canonical subjects are most likely to display. A summary of Hermon’s proposal is given below in (1). Briefly, she proposes that an argument which displays behaviors later on the list will also display all the behaviors earlier on the list. For example, an argument that can be the target of coreferential EQUI deletion (step (b)) will also show a ban on WH-movement (step (a)), but not necessarily vice versa, and arguments that are morphologically coded as subjects (step (c)) will exhibit all the other behaviors listed in steps (a)-(b).

---

1 The exact amount varies: Lewis (2009) cites 300,000 from a 1977 SIL survey, while Gómez-Rendón (2007) points out that the entire population of the Imbabura province was barely 250,000 in 1982; he proposes a more conservative estimate of 150,000 speakers.
(1)  a. switch-reference controller / ban on WH-movement / subject-to-object raising / demolition in passivization
   b. Target of switch-reference deletion / coreferential EQUI deletion / target of subject-to-subject raising
   c. Morphological coding [here, nominative case and governing subject-verb agreement]

The second hierarchy, also proposed by Hermon (2001), accounts for the varying behaviors of different types of non-canonical subjects in Imbabura Quichua by positing a hierarchical progression of less subject-like arguments to more subject-like arguments. Specifically, she shows that lexical experiencer subjects are less subject-like than desiderative experiencer subjects, which are themselves less subject-like than canonical subjects. This second hierarchy is shown below in (2).

(2)    canonical subject > desiderative > experiencer > non-subject

These two hierarchies were devised to fit together closely. Non-subjects exhibit none of the properties on Hierarchy (1). Experiencer non-canonical subjects exhibit only the subject properties on step (a), desiderative subjects exhibit those properties of both step (a) and step (b), and canonical subjects exhibit the behaviors at all levels.

In this paper, I make two claims. First, I show that IQ has a third type of non-canonical subject that has not been discussed as such in any previous work: The demoted subject of a causativized verb. Next, I argue that when the caused subjects are included in IQ’s inventory of non-canonical subjects, the close-fitting relation between Hierarchies (1) and (2) falls apart. This second claim is supported by further evidence from my own fieldwork that provides a slightly more nuanced look at certain properties of IQ introduced by Hermon (2001).

This paper is organized as follows. In §2 I introduce the currently accepted two types of non-canonical subjects in IQ. In §3 I describe IQ’s switch-reference system, whose treatment of non-canonical subjects is a property that appears on two steps of Hierarchy (1). This description will include certain subtleties of the switch-reference system that Hermon does not discuss — specifically, the fact that there are two different switch-reference systems that do not behave identically, and the possibilities of using the systems to encode non-subject-like behaviors as well as subject-like behaviors. In §4 I propose the third type of non-canonical subject — the caused subject — that has as yet not been analyzed as any kind of subject in IQ, and show how the switch-reference system treats it as a subject-like argument. In §5 I show how adding the caused subject to the existing analysis causes the relationship between Hierarchies (1) and (2) to break down, thus robbing it of much of its explanatory power, before concluding in §6.²

² The following abbreviations are used: 1SG=first person singular subject agreement; 1.OBJ=first person object agreement; 2SG=second person singular subject agreement; 3=third-person subject agreement; ACC=accusative case; ADV=adverbial marker; ADV.DS=adverbial different subject marker; ADV.SS=adverbial same subject marker; CAUS=causative marker; INF=infinitival marker; DAT=dative case; DESID=desiderative marker; DIM=diminutive marker; DIST=distributive marker; INTER=interrogative marker; LIM=limitative marker; LOC=locative case; NEG=negative suffix; NOM=nominative case; NMLZ=nominalizer; POSS=possessive marker; PROG=progressive aspect; SUBJ.DS=subjunctive different subject marker; SUBJ.SS=subjunctive same subject marker; TOP=topic marker; VAL=validator
2 Non-canonical Subjects in IQ

IQ is generally analyzed as having two types of non-canonical subjects: Desiderative experiencers and lexical experiencers (e.g., Cole 1982; Cole and Hermon 1981; Hermon 2001; Willgoths and Farrell 2009). Desiderative experiencers (which I am calling here “desiderative subjects”) occur with verbs that have been marked with the desiderative suffix -naya, corresponding to the canonical subject of those same verb roots that lack that suffix. Lexical experiencers (which I am calling here “experiencer subjects”) occur as the single argument to verbs that express certain physical or emotional sensations, such as nana- ‘hurt,’ rupa- ‘be hot,’ chiri- ‘be cold,’ yarxa- ‘be hungry,’ and sometimes muna- ‘want.’

These non-canonical subjects are easily distinguished from canonical subjects on the basis of three morphological properties: case-marking, subject-verb agreement, and object-agreement. Whereas canonical subjects receive a null nominative case and control subject-verb agreement, as shown below in (3), non-canonical subjects uniformly appear with the accusative suffix -ta and appear with default third-person verb agreement. These properties are shown below in (4) for experiencer subjects and (5) for desiderative subjects.4

(3) a. Juan-0 trabaja-xu-n
    Juan-NOM work-PROG-3
    ‘Juan is working.’
b. Nyuka-0 chay-ta villa-rka-ni Marya-man-ka
    I-NOM that-ACC tell-PAST-1sg Marya-DAT-TOP
    ‘I told that to Maria.’ (Cole 1982)

(4) a. Nyuka maki-ta nana-xu-n
    My hand-ACC hurt-PROG-3
    ‘My hand hurts.’
b. Kan-ta rupa-xu-n
    you-ACC be.hot-PROG-3
    ‘You are hot.’
c. Nyukanchi-ta chiri-xu-n
    we-ACC be.cold-PROG-3
    ‘We are cold.’
d. Nyuka-ta yarxa-xu-n
    I-ACC be.hungry-PROG-3
    ‘I am hungry.’ (2010-03mar-18-MXC-CPC)

(5) a. Juzi-ta punyu-naya-n
    Jose-ACC sleep-DESID-3

3 The verb muna- is slightly different from the other lexical experiencer predicates in that the subject seems to be able to vary freely between a canonical nominative subject that governs verb agreement and a non-canonical accusative experiencer subject that cannot govern verb agreement. Thus, it is perfectly possible to say either nyuka munani ‘I NOM want.1SG’ or nyuka-ta munan ‘I ACC want.3.’
4 Citations that do not reference published works indicate data taken from fieldwork between August 2009 and May 2010.
'Jose wants to sleep.'

b. *Nyuka-ta-ka mishki-ta  miku-naya-rka-0
   I-ACC   candy-ACC   eat-DESID-PAST-3
   ‘I wanted (to eat) candy.’

c. Kan-ta-ka  ufa-naya-n-chu
   you-ACC-TOP   drink-DESID-3-INTER
   ‘Do you want to drink?’ (Cole 1982)

d. Nyukanchi-ta xatun yaku-ta  riku-naya-n
   we-ACC   big   lake-ACC   see-DESID-3
   ‘We want to see the ocean.’ (2010-4apr-01-MXC-CPC)

The third morphological property that distinguishes non-canonical from canonical subjects — that of compatibility with object agreement morphology — is not always observable. IQ has lost the full paradigm of object-agreement markers or portmanteau subject-on-object morphemes that are still present in other Quechua languages (Hermon 2001; Weber 1989). However, it does still retain an optional first-person object agreement suffix -wa. The examples in (6) below show how this suffix can mark agreement with a first-person direct (6a-c) or indirect (6d) object, but not with a subject (6e). Nevertheless, it is perfectly compatible with IQ non-canonical subjects, both desiderative (7) and experiencer (8).

(6) a. Kan-Ø hayta-wa-ngi
   you-NOM   kick-1.OBJ-2.SG
   ‘You kick me.’ (2009-09sep-22-ACO-RWK-01-wl)

b. Pay-kuna-Ø wakta-wa-naxu-n (nyukanchi-ta)
   he-PL-NOM   hit-1.OBJ-DIST-3 (us-ACC)
   ‘They hit us.’ (2009-11nov-24-MC-CPC-01)

c. (Nyuka-Ø uchila ka-xpi)  nyuka tayta-kuna ku-na-wa-xu-rka-Ø  nyuka
   (I-NOM   little be-ADV.DS)   my   father-PL   give-DIST-1.OBJ-PAST-3   my
   awla-man
   grandmother-DAT
   ‘(When I was a baby) my parents gave me to my grandmother [e.g., to hold].’ (2009-11nov-24-MC-CPC-01)

d. Kan-pa tayta kan-Ø  uchila ka-xpi  nyuka-man ku-wa-rka-Ø
   you-POSS   father   you-NOM   little be-ADV.DS   me-DAT   give-1.OBJ-PAST-3
   ‘When you were a baby your father gave you to me [e.g., to hold].’ (2009-11nov-24-MC-CPC-01)

e. *Nyuka-Ø-ka  kan-ta  maka-wa-rka-ngi/ni
   I-NOM-TOP   you-ACC   hit-1.OBJ-PAST-2.SG/1.SG
   Intended: ‘I hit you.’ (Hermon 2001)

(7) a. Nyuka-ta-ka punyu-naya-wa-rka-Ø
   I-ACC-TOP   sleep-DESID-1.OBJ-PAST-3

5 Note that IQ can freely drop pronouns — both subject and object.
Hierarchies and Subjects in IQ Subordinate Clauses

‘I wanted to sleep.’

b. *Miku-naya-wa-n-mi*
eat-DESID-1.OBJ-3-VAL

‘I want to eat.’

(8) a. *Nyuka-ta-ka chiri-wa-rka-θ-mi*
I-ACC-TOP be.cold-1.OBJ-past-3-val

‘I was cold.’ (Hermon 2001)

3 Switch-reference

3.1 With Canonical Subjects

A more complete discussion of all the syntactic subjecthood behaviors exhibited by desiderative and experiencer subjects can be found in Cole (1982) and Hermon (2001). The purpose of this section is to summarize the behavior of the switch-reference system in IQ subordinate clauses, because the evidence from this particular subjecthood diagnostic is instrumental in separating the first two steps in Hierarchy (1), as well as in separating desiderative and experiencer subjects in Hierarchy (2).

IQ’s switch-reference system is a way of marking certain subordinate clauses with one of two suffixes on the subordinate clause verb. One suffix is used when the subjects of the main and subordinate clauses are coreferential, while the other suffix is used when the two subjects are not coreferential. I will be calling the first suffix the SS, or same-subject, marker, and the second suffix the DS, or different-subject, marker.

Two types of subordinate clauses employ this switch-reference system. The first is the adverbial clause, which is a subordinate adjunct clause indicating the manner of an action, a temporal setting, or the condition under which the main clause holds. It usually corresponds to English subordinate clauses introduced by *while, when, having done* or *since*. The second type of subordinate clause is what I shall call the subjunctive clause. It can be either the complement clause of verbs like *muna-‘want’* and *tapu-‘ask,’* or a purposive adjunct clause, corresponding to English clauses introduced by expressions such as *in order to, so that.* The form of SS and DS suffixes for these two clauses is shown below in (9).

<table>
<thead>
<tr>
<th>SS</th>
<th>DS</th>
</tr>
</thead>
<tbody>
<tr>
<td>-sha</td>
<td>-xpi</td>
</tr>
<tr>
<td>-ngapax</td>
<td>-chun</td>
</tr>
</tbody>
</table>

Examples of the SS/DS pattern for adverbials is shown in (10), while the pattern for subjunctives is shown in (11). In all examples the subordinate clause has been bracketed, with an underscore indicating the position of the deleted subject.

---

6 Since only a small number of verbs can introduce a subjunctive complement clause, I will be examining subjunctive patterns of SS- or DS-marking in the context of a purposive adjunct clause, which by its nature as an adjunct can occur in a much more varied set of sentences. It should therefore be remembered that the pattern described here for subjunctive clauses might be restricted only to their use in adjunct purposives, and may not extend to their use as complement clauses.
    [plow-CAUSE-ADV.SS-TOP] that grass-PL what sort-INF be-2sg
    ‘After you’ve plowed, you have to sort through whatever grasses are there.’ (How to run a farm, line 14, 2009-11nov-10-MXC-CPC-1)

    b. [__ Kashna-zha tuku-\textit{xpi}-ka], nya alpa-gu-ta kimi-chi-na
    [like.this-LIM become-ADV.DS-TOP], already land-DIM-ACC pile-CAUSE-INF
    ka-ngi be-2sg
    ‘When it (i.e., the corn plant) has become about this big, you have to pile the dirt up (i.e., around the stalk).’ (How to run a farm, lines 39-40, 2009-11nov-10-MXC-CPC-1)

(11) a. Nyuka-\textit{\textordmasculine} uchila ka-shpa ri-rka-ni [__ kay zambu-ta apamu-ngapax]
    I-NOM little be-ADV.SS go-PAST-1sg [that pumpkin-ACC fetch-SUBJ.SS]
    ‘When I was little, I went to fetch a pumpkin.’ (Mariana loses her pumpkin, line 1, 2009-09sep-29-MXC-CPC-1)

    b. Yaku-ta chura-na ka-ngi [atalpa-\textit{\textordmasculine} ufia-\textit{\textordmasculine}chun]
    water-ACC put-INF be-2sg [chicken-NOM drink-SUBJ.DS]
    ‘You have to put out water so the chicken drinks.’ (2009-11nov-10-MXC-CPC)

In (10a) the subject of the subordinate clause, ‘you,’ is the same as the subject of the main clause. The person doing the plowing is the same person who is doing the sorting of the grasses. For this reason, the subordinate verb is marked with the adverbial SS suffix -shpa. By contrast, in (10b) the subjects are different: It is the corn which has grown to a certain size while it is ‘you’ who must pile dirt up around the stalks. Thus, the subordinate verb is marked with -xpi, the adverbial DS suffix. In (11a) the patterns are similar: The subject of the main clause ‘I went’ is first person singular, the same as the subject of the purpose clause ‘to fetch a pumpkin.’ Thus the subordinate verb \textit{apamu}-, ‘fetch,’ is marked with the subjunctive SS morpheme -ngapax.\footnote{Note that in addition to a pumpkin-fetching subjunctive clause, there is another, sentence-initial adverbial clause about being little, whose subject is again the first-person singular subject of the main clause. Thus the clause \textit{nyuka uchila kashpa} ‘when I was little,’ is marked with the SS adverbial -shpa on the verb ‘be’: \textit{ka-shpa}.} In (11b) the person putting out the water, ‘you,’ is not the person who will drink the water — that honor goes to the chickens. Thus, the subordinate verb \textit{ufia}-, ‘drink,’ receives the subjunctive DS suffix -\textit{chun}. Any other combination of markers would be ungrammatical. The subordinate verb in (10a) cannot be marked with -xpi, and in (10b) it cannot be marked with -shpa. In (11a) the subordinate verb would be ungrammatical if it were marked with -\textit{chun}, and likewise in (10b) if it were marked with -ngapax. When only canonical subjects are involved, there is no choice in the use of the switch-reference suffixes.

3.2 With Non-canonical Subjects

The point of interest of IQ’s switch-reference system lies in its use as a subjecthood diagnostic. SS-marking is licensed only when the subjects of the main and subordinate clauses are coreferential. Therefore, if sentences with a controller in the main clause and a coreferential target (usually
Hierarchies and Subjects in IQ Subordinate Clauses

deleted) in the subordinate clause allows SS-marking, then both of those arguments can be analyzed as subjects in the syntax. In this way, a non-canonical subject can be analyzed as subject-like to the extent that it is compatible with SS-marking, and as non-subject-like to the extent that it is compatible with DS-marking.

In adverbial or subordinate clauses, there are two possible positions for a non-canonical subject to be evaluated:

1. The non-canonical subject may be in the main clause, where it is evaluated as a possible controller for a coreferent canonical subject in the subordinate clause.

2. The non-canonical subject may be in the subordinate clause, where it is evaluated as a possible target for the coreferent canonical subject in the main clause.

According to Hierarchy (1), repeated below for clarity, it is “easier” for a main-clause controller to count as a subject and trigger SS-marking under identity with the deleted target in the subordinate clause than vice-versa. This is represented in the hierarchy by putting the subjecthood property of being a switch-reference controller on the lowest step (a), while being a target is on the middle step (b).

(1)  
(a) Switch-reference controller / ban on WH-movement / subject-to-object raising / passivization  
(b) Target of switch-reference deletion / coreferential EQUI deletion / target of subject-to-subject raising  
(c) Morphological coding [here, nominative case and governing subject-verb agreement]

Hermon (2001) observes that, although both experiencer and desiderative subjects can license the use of the SS-marker when they are switch-reference controllers, they cannot always be switch-reference targets. The sentences in (12) demonstrate that both desiderative (12a) and experiencer (12b-c) subjects are compatible with the use of the adverbial SS-marker -shpa, while the sentences in (13) show that the same is true of desiderative (13a) and experiencer (13b) subjects with the subjunctive SS-marker -ngapax. Importantly, in these sentences the non-canonical subjects are all switch-reference controllers in the main clause.\(^8\)

(12)  
a. [__Ashtaka-ta punzhata chura-shpa/xpi], nyuka-ta punyu-naya-n  
a.lot-ACC today do-ADV.SS/ADV.DS, I-ACC sleep-DESID-3  
‘Because I did so much today, I want to sleep.’ (2010-03mar-11-MXC-CPC)

b. [__Ashtaka-ta punzhata chura-shpa/xpi-ka], nyuka-ta yarxa-xu-n  
a.lot-ACC today do-ADV.SS/ADV.DS-TOP, I-ACC hungry-PROG-3  
‘Because I did so much today, I’m hungry.’ (2010-04apr-01-MXC-CPC)

c. [__Yaku-pi ka-shpa-ka], chiiri-wa-rka-mi  
[水-LOC be-ADV.SS-TOP], be.cold-1.OBJ-PAST-VAL  
‘While in the water, I was cold.’ (Hermon 2001)

(13)  
a. Nyuka-ta punyu-naya-n [__musku-ngapax/*chun]  
I-ACC sleep-DESID-3 [__dream-SUBJ.SS/*SUBJ.DS]  
‘I want to sleep in order to dream.’ (2010-04apr-29-MXC-CPC)

\(^8\) For clarity, the deleted subordinate clause argument has been represented with an underscore.
b. Nyuka-ta nana-xu-n [__ ali tuku-ngapax/*chun]
I-ACC hurt-PROG-3 [ good become-SUBJ.SS/*SUBJ.DS]
‘I hurt [i.e., suffer in the religious sense] in order to become good.’ (2010-05may-04-ACO-EEW)

Note that, although both SS -shpa and DS -xpi are permissible for desiderative (12a) and experiencer (12b-c) subject controllers in adverbial clauses, only the SS-marker -ngapax is permissible for desiderative (13a) or experiencer (13b) controller subjects in subjunctive clauses. The DS marker -chun is impossible in these sentences.9

When these same non-canonical subjects are switch-reference targets in the adverbial clause, the DS-marker -xpi is permissible in all cases, but only desiderative subjects allow the use of the SS-marker -shpa.10

(14) a. [__ Miku-naya-shpa/xpi], papa-ta randi-rka-ni
[ eat-DESID-ADV.SS/ADV.DS], potato-ACC buy-PAST-1SG
‘Wanting to eat, I bought a potato.’ (2010-04apr-01-MXC-CPC)

b. [__ Yarxa-*shpa/xpi], papa-ta randi-rka-ni
[ hungry*-ADV.SS/ADV.DS], potato-ACC buy-PAST-1SG
‘Being hungry, I bought a potato.’ (2010-04apr-01-MXC-CPC)

The situation with non-canonical subject targets in subjunctive clauses is slightly muddier. On the one hand, it seems straightforward that deleted desiderative subjects in subjunctive clauses allow the use of the SS-marker ngapax (15a-b) as well as the DS-marker -chun — exactly parallel to their behavior in adverbial clauses.

(15) a. [__ Punyu-naya-ngapax] pastizha-ta ufia-rka-ni
[ sleep-DESID-SUBJ.SS] pill-ACC drink-PAST-1SG
‘In order to desire to sleep, I took a pill.’ (Hermon 2001)

b. Nyuka-θ na miku-xu-ni-chu, [__ ashtaka-ta chishi-pi
I-NOM not eat-PROG-1SG-NEG, [__ a.lot-ACC evening-LOC
miku-naya-ngapax/chun]
eat-DESID-SUBJ.SS/SUBJ.DS]
‘I’m not eating today so that I’ll want to eat a lot later this evening.’ (2010-05may-04-ACO-DJM)

9 It should further be noted that the sentences in (13) are slightly unnatural sounding. This is probably due to the fact that there are very few circumstances in which a person can undergo a physiological experience with sufficient agency to justify the use of the following purposive clause.

10 The examples in (14) contain no overt subjects whatsoever — in the main clause because they have simply been deleted, as often happens in IQ, and in the subordinate clause because they are the intended targets of switch-reference deletion. However, the presence of the non-canonical subjects in the subordinate clauses — even after deletion — can still be deduced from the verbs in the subordinate clauses. In (14a) the subordinate clause verb miku-naya-shpa ‘want to eat’ has the desiderative suffix -naya, indicating a desiderative subject, while in (14b) the subordinate clause verb yarxa-*shpa/xpi ‘be hungry’ is one of those verbs of physiological experience which lexically requires an experiencer subject.
Hierarchies and Subjects in IQ Subordinate Clauses

By contrast, it seems that the possibility of using the SS-marker -ngapax with a deleted experiencer subject in the subordinate clause varies by speaker or by sentential context. Hermon (2001) gives the example in (16a) as evidence that it is impossible to use -ngapax with deleted experiencer subjects in subordinate subjunctive clauses. However, the elicited example in (16b) is a counterexample, where DS -chun is impossible, and only SS -ngapax is allowed.

(16) a. * [Ama __ chiri-ngapax] nina-ta rura-rka-ni
   [not be.cold-SUBJ.SS] fire-ACC make-PAST-1SG
   Intended: ‘In order not to be cold, I made a fire.’

   b. Nyuka-∅ bayta-ta chura-xu-ni [__ rupa-ngapax/*chun]
   I-NOM shawl-ACC put.on-PROG-1SG [ be.hot-SUBJ.SS/*SUBJ.DS]
   ‘I put on my shawl in order to be hot.’ (2010-05may-04-ACO-EEW)

3.3 Summary So Far

On the basis of these structures (desiderative and experiencer subjects, and the switch-reference system for adverbial and subjunctive clauses) it seems that IQ is entirely consistent with Hierarchies (1) and (2). As the table in (17) shows, the only situation when non-canonical subjects do not act like canonical subjects — i.e., do not allow the SS-marker -shpa or -ngapax even when they are coreferent with the subject in the other clause — is when they are the target of switch-reference deletion in the subordinate clause. Being a grammatical switch-reference target is at step (b) on Hierarchy (1), higher than the property of being a switch-reference controller in the main clause, and thus it is to be expected that SS-marking would be more restricted for switch-reference targets than for controllers. Further, the only non-canonical subject which does not allow the use of the SS-marker -shpa or -ngapax when it is the switch-reference target is the experiencer subject, which is lower on Hierarchy (2) than the desiderative subject.

(17) Possibility of using SS- and DS-marker

<table>
<thead>
<tr>
<th></th>
<th>Main-clause (i.e., controller)</th>
<th>Subordinate-clause (i.e., target)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adverbial</td>
<td>Subjunctive</td>
</tr>
<tr>
<td>DESIDERATIVE</td>
<td>SS/DS</td>
<td>SS only</td>
</tr>
<tr>
<td>EXPERIENCER</td>
<td>SS/DS</td>
<td>SS only</td>
</tr>
</tbody>
</table>

4 Caused Subjects

In this section I introduce the third type of non-canonical subject: Caused subjects, which appear when the causative suffix -chi is added to the main verb. Cross-linguistically, the causative is a construction that adds a new argument to a clause in order to represent a notional causer of some action. This notional causer becomes a subject, while the original subject is demoted to some oblique or peripheral status (Palmer 1994). In IQ this added notional causer behaves like a canonical subject in (almost) every way, receiving nominative case and governing verb agreement. It is the demoted original subject, which I call the caused subject, that I propose to analyze as IQ’s third type of non-canonical subject.
4.1 Properties of the Caused Subject

Morphologically, the caused subject behaves like a direct object. Examples (18-19) below show representative non-causative/causative pairs, while example (20) shows a canonical direct object for comparison.

(18) a. *Chay alku-* kalpa-xu-n
   That dog-NOM run-PROG-3
   ‘That dog is running.’

b. *Chay xari-* kalpa-chi-xu-n *alku-ta*
   that man-NOM run-CAUS-PROG-3 dog-ACC
   ‘That man is making the dog run.’ (2009-09sep-17-ACO-LDM)

(19) a. *Nyuka-* punyu-ni
   I-NOM sleep-1SG
   ‘I sleep.’

b. *Kikin-* nyuka-ta punyu-chi-wa-ngi
   You-NOM me-ACC sleep-CAUS-1.OBJ-2SG
   ‘You put me to sleep.’ (2010-04apr-15-MXC-CPC)

(20) *Kan-* nyuka-ta maka-wa-rka-ngi
   You-NOM me-ACC hit-1OBJ-PAST-2SG
   ‘You hit me.’ (Hermon 2001)

As (18b) and (19b) show, the caused subject receives the accusative case marker -ta, and as (19b) shows, it can also be cross-referenced on the verb with the first person object marker -wa. These same properties can be seen for the canonical direct object in (20), and contrasted with the behavior of the notional causers in (18b) and (19b), which receive null nominative case and govern subject-verb agreement exactly like canonical subjects (e.g., (3), (18a), (19a). In this way, the caused subject is very similar to a direct object, and indeed the demoted subject of Quechua causatives is often analyzed as a direct object (e.g., Cole 1982; Davies and Rosen 1988).

In one striking way, however, the caused subject also behaves like a canonical subject: When the caused subject is coreferent with the canonical subject of another clause, IQ allows the use of the adverbial SS-marker -shpa both when the caused subject is the switch-reference controller in the main clause (21), and also when it is the target in the subordinate clause. Sentences with this second configuration are shown in (22)-(25). The (a)-continuations show main clauses whose canonical subject is coreferent with the caused subject of the preceding subordinate clause. The (b)- continuations show main clauses whose canonical subject is coreferent with the notional causer, or similarly canonical subject, of the preceding subordinate clause.

(21) * [Wawa; ashtaka-ta punzha-ta waka-shpa/xpi], nyuka pay-* punyu-chi-ni
   babyi a.lot-ADV day-ADV cry-ADV,SS/ADV,DS, I him,-ACC sleep-CAUS-1.sg
   ‘Since the baby cried a lot during the day, I’m putting him (=causing him) to sleep.’ (2010-04apr-01-MXC-CPC)

(22) * [Alku-* misi-ta wanyu-chi-shpa/xpi] ... [dog-NOM cat-ACC die-CAUS-ADV,SS/ADV,DS] ...
'Since the dog killed the cat (=caused to die) . . . '

a. . . kunan-ka  misi  ismu-xu-n
   . . today-TOP  cat  rot-PROG-3
   ‘. . today the cat is rotting’

b. . . pay-∅  nali  alku
   . . he-NOM  bad  dog
   ‘. . he is a bad dog.’ (2010-04apr-08-MXC-CPC)

(23) [Nyuka-∅  wawa-ta  punyu-chi-shpa/xpi-ka]
    . . [I-NOM  baby-ACC  sleep-CAUS-ADV.SS/ADV.DS-TOP] . .
   ‘Since I made the baby sleep . . . ’

a. . . punzha-pi  pay-∅  kushi  ka-n
   . . day-LOC  he-NOM  happy  be-3
   ‘. . in the afternoon he is cheerful.’

b. . . punzha-pi  nyuka-∅  kushi  ka-ni
   . . day-LOC  I-NOM  happy  be-1SG
   ‘. . in the afternoon I am cheerful.’ (2010-04apr-08-MXC-CPC)

(24) [Nyuka-∅  wagra-kuna-ta  miku-chi-shpa/xpi-ka]
   ‘Since I fed the cows (=caused them to eat) . . . ’

a. . . pay-kuna-∅  ali-mi  ka-n
   . . he-NOM-PL  good-VAL  be-3
   ‘. . they are good (=healthy).’ (2010-04apr-08-MXC-CPC)

b. . . kikin-∅  you-ACC  kushi  ka-ni
   . . day-LOC  I-NOM  happy  be-1SG
   ‘. . you are happy.’ (2010-04apr-08-MXC-CPC)

(25) [Kikin-∅  nyuka-ta  runa  shimi-ta  yacha-chi-shpa/xpi-ka]
   ‘Since you taught me (=caused me to know) how to speak Quichua11 . . . ’

a. . . chay-mi  ali  rima-y-ta  usha-ni
   . . that-VAL  good  speak-NMLZ-ACC  can-1SG
   ‘. . now I can speak well.’

b. . . kikin-∅  kushi  ka-ngi
   . . you-NOM  happy  be-2SG
   ‘. . you are happy.’ (2010-04apr-08-MXC-CPC)

As the (a)-continuations of (22-25) show, it is possible to use either the SS-marker -shpa or the DS-marker -xpi on the subordinate verb. The compatibility of the caused subject with the use of the DS-marker -xpi is to be expected, given the other object-like properties of the caused subject. It is their compatibility with the use of SS -shpa that forms the basis for my claim that caused subjects are a type of non-canonical subject. Indeed, this claim is not entirely inconsistent with previous analyses. Both Cole (1982) and Davies and Rosen (1988) analyze IQ causatives as a monoclausal

---

11In addition to kichwa ‘Quichua,’ the language can also be referred to with the term runa shimi, literally ‘(indigenous) person mouth.’
surface structure derived from a biclausal underlying structure in which the underlying subordinate clause subject — my caused subject — becomes a surface object. Under this analysis, it is not implausible for the underlying subject-status of the caused subject to survive in the form of compatibility with the SS adverbial marker -shpa.

Interestingly, although the caused subjects are compatible with the adverbial SS-marker -shpa, they are not compatible with the subjunctive SS equivalent -ngapax. When a caused subject in a subjunctive clause is coreferent with a canonical subject in the main clause (26a), or vice versa (26b), the only possible marker is the DS subjunctive marker -chun.

(26) a. Pay-ta tushu-chi-ni [asi-*ngapax/chun]
   him-ACC dance-CAUS-1SG [laugh-*SUBJ.SS/SUBJ.DS]
   ‘I make him dance so that he’ll laugh.’ (2010-04apr-29-MXC-CPC)

b. Wawa-kuna yachachik wasi-man ri-n, [nyuka-0 pay-kuna-ta
   child-PL teacher house-DAT go-3, [I-NOM they-PL-ACC
   yacha-chi-*ngapax/chun]
   know-CAUS-*SUBJ.SS/SUBJ.DS]
   ‘Children go to school in order for me to teach them (=cause them to learn).’ (2010-05may-04-ACO-DJM)

4.2 Possible Non-canonical Properties of the Notional Causer

In addition to showing the possibility of using the SS-marker -shpa when the intended coreference is between a canonical subject and a caused subject, sentences (22)-(25) are striking in another way. It is not only caused subjects that are compatible with both the SS-marker -shpa and the DS-marker -xpi when they are the intended coreferent argument with the subject of another clause. The same seems to be true for the notional causers, which in every other way behave like canonical subjects. For example, in (23), the seemingly canonical subject of the subordinate clause is the notional sleep-causer nyuka ‘I,’ which is marked with null nominative case and in main clauses would control agreement on the verb (see (19b)). When the main clause subject is also nyuka ‘I,’ as in (23b), then the two subjects of the two clauses are entirely canonical and coreferent, and we would expect the verb to be marked with the SS-marker -shpa, as is the case in every sentence with coreferent canonical subjects (e.g., (10a) and (11a)). Indeed, this is grammatical — yet it is also grammatical for the subordinate verb to be marked with the DS-adverbial suffix -xpi, something that is impossible for non-causer canonical subjects. In other words, not only can a caused subject count as a subject when it is itself the switch-reference target, as seen by the possibility of using the SS-marker -shpa with all of the (a)-continuations of (22)-(25), it can also interfere with the subjecthood status of the causer when it is the causer that is the switch-reference target. Since compatibility with DS-marking is a property of non-subjects, then the possibility of using the DS-marker -xpi with all of the (b)-continuations of those sentences shows that the notional causer has at least one non-subject-like characteristic.

It should be noted, however, that the possibility of using the DS-marker -xpi to mark intended coreference between a canonical subject and the notional causer seems to be restricted to constructions where the notional causer is the target. When it is the controller in the main clause, as is the case with the first-person singular subject in (27) below, it behaves exactly as a canonical subject, requiring the use of SS -shpa and disallowing the use of DS -xpi.
Hierarchies and Subjects in IQ Subordinate Clauses

(27) [Na ali mama ka-shpa/*xpi-ka], wawa-ta irki-chi-rka-ni
[Not good mother be-ADV.SS/*ADV.DS-TOP], baby-ACC sick-CAUS-PAST-1SG
‘Since I am a bad mother, I made my baby sick.’ (2010-04apr-08-MXC-CPC)

If this distinction is not simply an artifact of a confusing elicitation session, then it is consistent
with Hierarchy (1), which states that arguments show more subject-like properties when they are
switch-reference controllers (step (a)) than when they are switch-reference targets (step (b)). If
we reverse that claim and make predictions about when arguments should show non-subject-like
properties, then we would predict that this would occur in step (b) constructions before step (a)
constructions. This is exactly what happens with the notional causer: It allows the use of the DS-
marker xpi, a non-subject-like property, when it is the switch-reference target (step (b)), but not
when it is the switch-reference controller (step (a)).

5 Problems Posed by Caused Subjects

5.1 Summary So Far

So far I have introduced two accepted non-canonical subjects in IQ — the desiderative and expe-
riencer subjects — and argued for the existence of a third: The caused subject. The specific sub-
ject properties exhibited by each of the three non-canonical subjects vary, but Onishi (2001) and
specifically Hermon (2001) have proposed that this variation can be analyzed hierarchically. Cer-
tain syntactic subject-properties are more readily exhibited by non-canonical subjects than other
properties, and on the basis of how many of these properties each type of non-canonical subject
exhibits, it is possible to arrange them on a scale of more to less subject-like. In this way Hermon
proposes the two interdependent hierarchies that I summarized in (1)-(2). Canonical subjects ex-
hibited the behaviors of all three steps (a)-(c) on Hierarchy (1), desiderative subjects exhibited the
behaviors of the first two steps (a)-(b), and experiencer subjects exhibited the behaviors of the first
step (a) only.

In this paper I have been focusing on the properties of subjects in IQ’s switch-reference system.
In sentences containing subjunctive or adverbial subordinate clauses with two coreferent subjects,
the switch-reference system requires the use of the SS-suffix -ngapax or -chun to mark the sub-
ordinate clause verb. This property of the language can be used as a subjecthood diagnostic in
switch-reference constructions that contain one canonical subject and one non-canonical subject.
To the extent that the SS-marker is possible, the non-canonical subject is subject-like, and to the
extent that a DS-marker is possible, the non-canonical subject is non-subject-like. It is the compati-
bility of caused subjects with SS-marking that leads me to propose that they should be considered a
third type of non-canonical subject. A summary of the data presented so far in this paper regarding
the subjecthood of all three types of non-canonical subject is given in (28).

(28) Possibility of using SS- or DS-marker

12Strictly speaking, of course, the very first place where arguments should show non-subject-like properties is in the
morphology, which is the highest step, step (c), of Hierarchy (1). Thus, the fact that notional causers can act like
non-subjects with respect to switch-reference but not with respect to morphology could be taken as a case where
Hierarchy (1) breaks down. However, I am not sufficiently confident of the behaviors of notional causers to make
this claim outside of a footnote.
5.2 Problems with the Hierarchies

As the table in (28) shows, caused subjects are compatible with the use of a SS-marker in an adverbial switch-reference construction, but not in subjunctive switch-reference constructions. This distinction causes problems for Hierarchies (1)-(2).

Hierarchy (1) claims that non-canonical subjects are more likely to be controllers in a switch-reference system than targets. In other words, it is easier to use the SS-marker when the non-canonical subject is a controller in the main clause than when it is in the subordinate clause. This claim is supported by the behavior of experiencer subjects, which are compatible with SS-marking as controllers, but not as targets. However, the distinction for caused subjects seems to be not whether they are targets or controllers in the switch-reference system, but whether they are being used in an adverbial or subjunctive switch-reference system. In the subjunctive switch-reference system, caused subjects cannot license SS-marking even as controllers, the lowest subject-property on Hierarchy (1), while in the adverbial system they can easily be both controllers and targets, the lowest and middle steps on Hierarchy (1).

One approach to account for this data would be to separate the switch-reference systems, such that one is more willing (i.e., lower on Hierarchy (1)) to accept a non-canonical argument as subject-like than the other. For example, it is impossible to use the subjunctive SS-marker -ngapax for caused subjects, while the adverbial SS-marker -shpa is entirely compatible with them. This suggests that the adverbial switch-reference system is more “lenient,” or is lower on Hierarchy (1) than the adverbial switch-reference system. However, the ability of some non-canonical subjects to appear with DS-markers makes this conclusion less satisfactory. Although the adverbial switch-reference system seems more willing to allow the use of the SS-marker, suggesting that it accepts more arguments as subject-like, it is also more willing to allow the use of the DS-marker, suggesting that it likewise accepts more arguments as non-subject like. Specifically, the adverbial switch-reference system allows the use of the DS-marker -xpi for both desiderative and experiencer controllers, whereas the subjunctive system requires the use of the SS-marker -ngapax. This property suggests that it is the subjunctive system, and not the adverbial system, that is more lenient, as the subjunctive system is willing to accept desiderative and experiencer subjects as entirely unambiguous subjects. According to the subjunctive system, desiderative and experiencer subjects are so subject-like that they require SS-marking when they are controllers in the main clause, and cannot license DS-marking. These observations lead to the conclusion that the different behaviors of IQ’s two switch-reference systems cannot be accounted for in terms of leniency. Rather, it seems to be the case that the adverbial system simply allows more variation in the choice of SS- or DS-marking than the subjunctive system.

With respect to Hierarchy (2), the behavior of caused subjects is similarly problematic. Hermon (2001) shows strong evidence that desiderative subjects are more subject-like than experiencers. Specifically, she demonstrates that desiderative subjects exhibit every subjecthood property on steps (a) and (b) of Hierarchy (1), while experiencer subjects exhibit only those properties on step
Hierarchies and Subjects in IQ Subordinate Clauses

(a) However, if IQ has a subjecthood hierarchy, then where do caused subjects fit in? Caused subjects are as subject-like as desideratives with adverbial switch-reference, allowing SS-marking both as controllers and as targets, but they are the least subject-like of all of them when evaluated by the subjunctive switch-reference system.

5.3 Evidence from Other Areas of the Grammar

Although this paper has focused on the switch-reference system, there are a number of other syntactic patterns which Hermon describes as separating subjects from objects in IQ. Two that I have investigated with respect to caused subjects are those of raising and passivization. Briefly, if a language has raising predicates, it is more likely to raise subjects than objects. Hermon gives data indicating that experiencer subjects cannot be raised, while desiderative subjects can. I, however, have found that my consultant accepts the raising of both desiderative and experiencer subjects. Regardless of this disagreement, however, she strongly rejects the raising of caused subjects. Thus, (29)-(30) below are good, while (31) is bad.

(29) Kikin-θ\(^{13}\) yarxa-y yari-ngo
you-NOM hungry-NMLZ seem-2SG
‘You seem to be hungry.’

(30) (Kikin anfa-shpa), punyu-naya-y yari-ngo
(you yawn-ADV.SS), sleep-DESID-NMLZ seem-2SG
‘(Because you are yawning,) you seem to want to sleep.

(31) * Wawa punyu-chi-y yari-n
baby sleep-CAUS-NMLZ seem-3
‘The baby seems to have been put to sleep.’ (2010-04apr-15-MXC-CPC)

Thus, although Hermon’s and my investigations thus yield different results with respect to experiencer subjects, the behavior of these raising constructions clearly puts caused subjects in the non-subject category.

The non-subject-like status of caused subjects can further be seen with passivization. Briefly, like objects and unlike experiencer and desiderative subjects, caused subjects can be easily passivized: (32a), which promotes a caused subject under passivization, is perfectly fine, while (32b-c), which attempt to promote an experiencer and a desiderative, respectively, are unacceptable.

(32) a. Wawa-ka nyuka miku-chi-shka ka-rka-θ
Child-TOP I eat-CAUS-NMLZ be-PST-3
The child was fed (=made to eat) by me.

b. * Nyuka-ka chiri-shka ka-rka-ni
I-TOP cold-NMLZ be-PST-1.SG
Intended: I was colded.

c. * Nyuka-ka punyu-naya-shka ka-rka-ni
I-TOP sleep-DESID-NMLZ be-PST-1.SG

\(^{13}\)Note that here the subject must be nominative, not accusative. Evidently, when it has been raised from the domain of influence of the experiencer or desiderative predicate, the accusative case-marking requirement disappears.
Intended: I was desired to sleep. (Hermon 2001)

On the basis of these syntactic tests, then, it seems that caused subjects are more object-like — passivizable but unraisable — than experiencer and desiderative subjects. Simply by tallying of syntactic properties, then, it seems appropriate to put caused subjects lowest on the subjecthood hierarchy. A revised version of hierarchy (2) is given below.

\[
(2') \text{ canonical subject} > \text{ desiderative} > \text{ experiencer} > \text{ caused} > \text{ non-subject}
\]

The problem with this particular hierarchy, however, is that, with the addition of caused subjects, it is descriptive but not predictive. Although arguments that are higher on the hierarchy possess more subject properties than lower arguments, it is not the case that higher arguments exhibit the same properties as the lower arguments. Although caused subjects are the lowest on the hierarchy, they are compatible with SS-marking in the adverbial switch-reference system when they are targets of that system. Experiencer subjects, which are higher than caused subjects, cannot do this. Even if the evidence from subjunctive clauses and DS-marking had not caused problems for Hierarchy (1), the addition of caused subjects to Hierarchy (2') means that there is no longer a principled relation between those two hierarchies. It is no longer the case that each new step of Hierarchy (1) corresponds to the new set of properties exhibited by each new argument type on Hierarchy (2). Their explanatory power is thus vastly weakened.

6 Conclusion

Using hierarchical structures in the description of a language’s grammar allows the linguist to account for cooccurrences of particular properties in some principled way. In Imbabura Quichua, it has been proposed that there are hierarchies both of subjecthood properties and of non-canonical subjects, such that the varying behaviors of non-canonical subjects can be explained according to the step on which the subjecthood properties fall on Hierarchy (1) and where the non-canonical subjects themselves fall on Hierarchy (2). In this paper, I have shown that, although much of the behavior of IQ non-canonical subjects does mostly align with these proposed hierarchies, the addition of caused subjects to the inventory of non-canonical subjects is problematic for Hierarchy (2). I have further shown that the two different switch-reference systems do not seem to diagnose subjecthood in exactly the same way, which suggests that using behavior in a switch-reference system as a subjecthood diagnostic in Hierarchy (1) is an overly simple approach. Thus, although the hierarchical analysis of IQ non-canonical subjects accounts for most of the data in the language, it follows the noble tradition of all descriptive grammars, and leaks with the best of them.

References


Hierarchies and Subjects in IQ Subordinate Clauses


Clara Cohen
University of California, Berkeley
Department of Linguistics
1203 Dwinelle Hall
Berkeley, CA 94720
cpccohen@berkeley.edu
One -mi: An Evidential, Epistemic Modal, and Focus Marker in Imbabura Quechua

IKSOO KWON
Hankuk University of Foreign Studies

1 Introduction

The Quechua language family has received much attention for its well-attested evidential/epistemic modal system (e.g., Faller 2003, Weber 1986, Nuckolls 1993, and Floyd 1999). This paper focuses on the multiple functions of the particle -mi in Imbabura Quechua (IQ). Some of the functions are exemplified in (1)-(3).1

(1) Juan-mi punyu-rka
Juan-mi sleep-PST
‘Juan slept.’2

(2) a. Juan-mi Berkeley-man ri-gri-n
Juan-mi Berkeley-to go-FUT-3
‘Juan will go to Berkeley.’ [after Faller 2003; elicited example]

b. Juan-mi izhakizha ka-ni
Juan-mi sad be
‘Juan is sad.’ [after Faller 2003; elicited]

(3) A: pita wanyuchi-rka pirkuti-ta
who kill-PST rat-ACC
‘Who killed the rat?’

B: Pepe-mi wanyuchi-rka pirkuti-ta
Pepe-mi kill-PST rat-ACC
‘Pepe killed the rat.’

In (1), the use of -mi has (by default) a firsthand evidential function; it indicates that the origo has herself observed the focal situation of Juan’s sleeping. In (2a), which includes a future tense marker, i.e., which is irreals, -mi has a somewhat different function. There, -mi has a positive epistemic modal sense; it indicates the speaker’s commitment to the validity of the information. In (2b), -mi can also be said to be functioning more as an epistemic modal marker than as an

1 Except where otherwise marked, all the data in this paper were elicited from two language consultants (called here MC and AO) during a field methods class at the University of California, Berkeley in 2009-2010. My gratitude goes to the consultants for their patience and kindness.

2 Boldface is used throughout this paper to represent the firsthand evidential, strong-positive epistemic modal, and discourse-focus meanings conveyed by -mi that cannot be expressed in English with a single morpheme (but are expressed periphrastically or by word order or intonation).
evidential marker, as the origo cannot have direct access to the internal state of the subject. In (3), 
\(-mi\) marks the focused element in the utterance, the answer to the preceding question. It is
important to note that these contexts of use are not necessarily disjunctive; as a result, the
utterances in (1) and (4) are ambiguous.

(4) Juan shamu-rka-\textit{mi} 
Juan come-PST-\textit{mi} 
‘Juan came.’

The sentence in (4) could be used in situations where, for example, the speaker witnessed Juan’s
arrival (a firsthand evidential reading), Juan told the speaker he had come to the place in question
(a positive epistemic modal reading), or where the speaker wished to stress that it was Juan, as
opposed to someone else, who had come (a contrastive focus reading).

The aim of this paper is to examine the possibility that these three functions are conceptually
interrelated. It explores two major questions that bear on this issue. The first question is related
to the ongoing debate between so-called conflationists and non-conflationists about whether
evidentials and epistemic modals should always be analyzed as belonging to a single functional
category, i.e., whether the two categories can ever be separated or not (Michael 2010). This
paper shows that there are quite a few situations in which it is not a simple matter to identify
which role \(-mi\) is playing in the context (for example, in performatives and irrealis clauses). I
suggest that, to better understand how \(-mi\) is licensed in both firsthand evidential and positive-
epistemic modal contexts in Imbabura Quechua, we should consider the conceptual causal event
structure of epistemicity (Sweetser 1990), which is evoked by both evidentials and epistemic
modals. This paper argues that \(-mi\) cannot be classified as a pure evidential marker or as a pure
epistemic modal marker, that it rather conveys both notions, and further, that to identify which is
the more salient function in any given case, we must consider the whole causal event structure of
epistemicity, which is evoked by both categories; this event structure involves a series of sub-
event structures linked by the origo’s inference processes. I further suggest that the separation of
the two categories is not, in fact, the most important issue, but that characterizing each in relation
to the other is key to grasping the nature of each.

The other major question is whether the core function of \(-mi\), whether one defines it as
primarily an evidential marker or primarily an epistemic modal marker, is being functionally
extended where \(-mi\) is used as a focus marker or whether the focus marking function is
completely unrelated (i.e., that \(-mi\) is synchronically polysemous). This study explores the
possibility that the functions of encoding firsthand evidentiality/positive epistemic modality and
focus are conceptually related, and concludes that they do indeed share a cognitive motivation in
that in both cases, the origo’s attention is explicitly directed to the cognitively more salient
portion of a given stimulus. In the case of evidentiality and epistemic modality, the target event
itself is picked out as being the most cognitively salient aspect of the situation while the origo’s
mode of access to the event or her beliefs about the likelihood of its occurrence are presupposed
and backgrounded. In the case of focus marking, the focused element is picked out as being
cognitively salient while the other elements in the construction are explicitly given less attention.

Section 2 of the paper reviews previous approaches to evidentials and epistemic modals in the
Quechua languages, summarizing points of agreement and disagreement in the literature. Section
3 explores the evidential and epistemic modal aspects of \(-mi\), analyzing a variety of relevant
eamples, and begins to develop a comprehensive description of its function. Section 4 examines
the morpheme’s focus marking function, providing a further basis for the later discussion of how the multiple functions of -mi may be conceptually related. Section 5 sets for the hypothesis that the firsthand evidential/strongly-positive epistemic modal and focus marking functions of -mi are conceptually related; they all explicitly foreground stimuli that are cognitively more salient and background stimuli that are presupposed or part of the common ground in the given context. This section discusses the theoretical implications of that relationship. Section 6 concludes the discussion.

2 Previous Approaches

Quite a few linguists (e.g., Chafe 1986, Willett 1987, de Haan 1999) have discussed the relationship between evidentials and epistemic modals. This relationship has been investigated with regard to the Quechua languages in particular Faller (2003), Weber (1986), Nuckolls (1993), and Floyd (1999), among others.

The outstanding question in this research is whether it is possible to disentangle the notion of evidentiality, which relates to the speaker’s source of information about the focal event, from the notion of epistemic modality, which relates to the speaker’s epistemic stance towards the focal event. For example, the usual interpretation of the utterance in (1) is both that the speaker has direct evidence that the focal situation occurred, for example having seen Juan sleeping, and that she is convinced it is a fact that the focal situation occurred.

The authors mentioned above take different positions as to whether both of these semantic properties are lexically encoded in -mi. Weber (1986, summarized in Faller 2003) argues that, lexically, evidentiality is encoded as a semantic property of -mi and that the ‘validational’ (i.e., positive-epistemic modal) meaning derives from the extralinguistic cultural axiom that one’s own experience is reliable. Nuckolls (1993) takes the opposite view, asserting that -mi is primarily ‘validational,’ i.e., that it lexically encodes positive epistemic modality, while the evidential meaning is implied. However, as Faller (2003) points out, neither of these approaches is satisfactory, as there are more than a few examples for which it is quite difficult to determine which of the two functions is more salient (for example, in a situation in which the speaker has physically witnessed something, but is not sure what it was, or in a situation in which the speaker was sure that something would happen in the future).

To account for these functions that are not easily categorizable, Floyd (1999) argues that the functions of -mi in Wanka Quechua form a radial category, with the evidential function at the center of the radial network. Similarly, Faller (2003) proposes that -mi in Cuzco Quechua should be analyzed as an evidential/epistemic modal element whose semantic and pragmatic function is parameterized by evidential licensing conditions. She treats -mi not as a pure epistemic modality marker but as an emphatic illocutionary force marker because it does not weaken assertions in the way that epistemic modal markers usually do (Speas 2008:956). Faller also argues that -mi is licensed if the speaker has access to the most direct source of information possible about the described event.

This paper shares Floyd’s and Faller’s position that maintaining a strict distinction between evidentiality and epistemic modality does not allow us to fully account for the target phenomena from prototypical cases to less prototypical ones. However, I further argue that neither evidentiality nor epistemic modality can be fully understood without considering the causal structure shared by the two functional categories (Sweetser 1990), i.e., the chain of causal events linked by an origo’s inference process. (This concept is described in detail in section 5.)
The other major issue discussed in this paper is whether -mi’s focus marking function is a functional extension of its other uses. There has been some descriptive work on the use of -mi as a focus marker (e.g., Weber 1989, Muysken 1995, Cole 1982), but the relationship between the evidential/epistemic modal and the focus marking functions has not received much attention.

3 Evidentiality vs. Epistemic Modality

In this section, I show that -mi is licensed in both evidential and epistemic modal context, indicating that it cannot be described either as a pure evidential marker or as a pure epistemic modal marker, but rather should be described as marking both.

3.1 Direct Evidentiality and Spatial Specification

According to my IQ language consultant, MC, -mi can be used to encode firsthand direct evidentiality. According to her, the sentences in (5) and (6) are subtly different.

(5)  chaypi  alku
     there  dog
     ‘There’s a dog there.’

(6)  chaypi-\textit{mi}  alku
     there-\textit{mi}  dog
     ‘There’s a dog there.’

It is notable that MC accompanied the utterance in (6) with a pointing gesture to indicate the location of the hypothetical dog; it seems that -mi narrows down the space being referred to in the utterance to indicate an entity within the range of the experiential origo’s vision. This suggests that -mi can function as a spatial specifier, encoding adjacency or immediacy (like right or just in English), which is conceptually contiguous to the directness of the origo’s mode of access encoded by -mi.

The direct evidential function of -mi is demonstrated by the contrast between the examples in (7).

(7)  a.  \textit{Iksoo}  shamu-r\textit{ka}\newline     \textit{Iksoo}  come-PST\newline     ‘Iksoo has come.’

     b.  \textit{Iksoo}  shamu-r\textit{ka}-mi\newline     \textit{Iksoo}  come-PST-\textit{mi}\newline     ‘Iksoo \textbf{has come}.’

According to MC, a sentence like that in (7a) might be used when the subject was outside the room, for example, if Iksoo had just knocked on the door and the speaker wanted to suggest someone should open it. In contrast, a sentence like that in (7b) might be used when Iksoo was already in the room and the speaker wanted to tell someone else Iksoo was with her; in this latter case, the speaker would, of course, have visual access to Iksoo.
The direct evidential function of -mi can be tested by adding a sentence to the utterance that explicitly denies that the speaker had visual access, as in (8).

(8) ??Ines-ka kayna pay-pa nyanya-ta-mi tupa-ri-rka
    Ines-TOP yesterday 3SG-GEN sister-ACC-mi meet-REF-PST
    nyuka na riku-rka-ni-chu
    1SG NEG see-PST-1SG-NPI
    ‘Ines visited her sister yesterday. I didn’t see it.’

AO said that (8) is “contradictory,” indicating that -mi carries an indefeasible evidential implicature.3

These examples show that -mi definitely has a direct evidential function, encoding that the speaker perceived the focal information via a direct mode of access (MOA; Hanks 1990).

3.2 Strongly Positive Epistemic Modality

The previous section demonstrates clearly that -mi has an evidential function. However, it is also licensed in some non-evidential contexts, where the origo did not obtain the focal information directly. For example, the utterance in (9) might be licensed in a context where Juan had told the speaker about having run.

(9) Juan kalpa-rka-mi
    Juan run-PST-mi
    ‘Juan has run.’

If -mi were only a direct evidential marker, it would not be licensed in (9), as the MOA (report) is not direct.

-mi is also licensed in utterances such as those in (2) (repeated here as (10)) that refer to irrealis events or to the emotional state of some third party, neither of which the origo can directly observe or experience at the speech time.

(10) a. Juan-mi Berkeley-man ri-gri-n
    Juan-mi Berkeley-to go-FUT-3
    ‘Juan will go to Berkeley.’ [after Faller 2003; elicited]

b. Juan-mi izhakizha ka-n
    Juan-mi sad be-3
    ‘Juan is sad.’ [after Faller 2003; elicited]

In these examples, -mi is being employed as a positive epistemic modal element, encoding the

---

3 When I first asked directly what the difference was between Ineska kayna paypa nyanya-ta tuparirka and Ineska kayna paypa nyanya-ta-mi tuparirka, AO said the latter means the speaker physically saw Ines, indicating that -mi is closely related to direct evidentiality.

4 However, it should be noted that the situation may be somewhat more complex, as I discuss in section 5.1.
speaker’s commitment to the validity of the information.

There are a number of other contexts in which -mi can play an epistemic modal role. First, -mi is licensed in performative utterances. As performatives are essentially about the speaker’s intent to induce the addressee to conform to her intention, they cannot be said to pertain to the origo’s (previous) experience, but rather refer to irrealis events that have not yet been assimilated into experience; therefore, they are generally not compatible with a direct evidentiality. When I asked AO about the performative utterances with -mi in (11) and (12), he said they sounded natural.

(11) nyuka ni-ni-mi kan-kuna kosa warmi ka-angi-chi
    1SG say-1SG-mi you-PL husband wife be-2PL-CAUS
    ‘I pronounce you husband and wife.’

(12) kay wasi-ta shuti-chi-rka-ni-mi Maria-ta
    that building-ACC name-CAUS-PST-3SG-mi Maria-ACC
    ‘I name the building Maria.’

Secondly, -mi is licensed in contexts in which the speaker is talking about events that she could not have been present for, for example because they occurred before her birth, if she is sure that they actually occurred.

(13) A: maypita wacha-ri-rka kan-pa mama
    where be.born-REF-PST 2SG-GEN mother
    ‘Where was your mother born?’

    B: nyuka mama-mi wacha-ri-rka Seoul-pi.
    my mother-mi born-REF-PST Seoul-LOC
    ‘My mother was born in Seoul.’

For example, if B had been told about the focal event by a reliable source such as his grandfather and was sure it was the truth, according to AO, the response in (13) with -mi would be licensed.5 This supports the hypothesis that -mi has an epistemic modal function in which it indicates that the speaker is committing herself to vouching for the validity of the focal information.

Lastly, -mi can be used in situations where the speaker wishes to indicate certainty about a proposition that might be supposed to be in question; for example, AO produced the utterance in (14) when I asked what a speaker might say if she saw someone from a distance, but was sure it was Josh.

(14) nyuka-mi riku-rka Josh-ta
    I-mi see-PST Josh-ACC
    ‘I saw Josh.’

5 According to AO, if the indirect evidential marker -shi had been used instead in the response in (13), it would be likely taken to indicate the speaker’s doubt and be interpreted as a confirming question. (It could be translated periphrastically as something like ‘I heard my mother was born in Seoul, right?’) In other words, the indirect evidential -shi is not licensed in such contexts unless the focal information is shared knowledge in the interlocutors’ common ground.
-mi would not have been licensed if the speaker had not been sure that the person she saw was Josh. In this situation, a direct evidential might not be appropriate because the evidence was not robust, but -mi can nonetheless be used, indicating that the constraints on the epistemic modal sense can override the constraints on the evidential sense.

The strongly positive epistemic modality of -mi can be confirmed by testing whether it is compatible with the weak epistemic modal adverbial nyarazha ‘maybe.’

\[
\text{(15) } \text{nyarazha tamya-gri-n(*-mi) mayb}e \text{ rain-FUT-3SG(-mi)} \\
\text{‘Maybe, it must be going to rain.’}
\]

\[
\text{(16) } \text{nyarazha tamya-xu-n(*-mi) mayb}e \text{ rain-IMPF-3SG(-mi)} \\
\text{‘Maybe, it must be going to rain.’}
\]

In (15) and (16), the strongly positive epistemic modality of -mi conflicts with the weakly positive epistemic modal adverb nyarazha. (While the utterance in (16) could also have a direct evidentiality reading, which would in itself conflict with nyarazha, the utterance in (15) is irrealis, and therefore cannot be construed as being based on direct evidence.) The ungrammaticality of these examples clearly shows that -mi conveys strongly positive epistemic modality.

4 Focus Marking

This section explores the focus marking function of -mi in detail. I show that information that has not yet been assimilated into the addressee’s knowledge is regarded as novel and that -mi is licensed to mark that novelty.

4.1 Argument Focus

One test that is commonly used to identify topic/focus contrasts is to set up a question-and-answer pair. The information that was explicitly included in the questioning utterance becomes the topic of the response utterance, while the new information (that corresponds to the interrogative pronoun in the question) is in focus. Some examples demonstrating the use of -mi in response to a question are given in (3) (repeated here as (17)) and in (18).

\[
\text{(17) a. A: pita wanyuchi-rka pirkuti-ta?} \\
\text{who kill-PST rat-ACC} \\
\text{‘Who killed the rat?’}
\]

\[
\text{B: Pepe-mi wanyuchi-rka pirkuti-ta} \\
\text{Pepe-mi kill-PST rat-ACC} \\
\text{‘Pepe killed the rat.’}
\]

---

6 In my other consultant’s dialect (MC’s), the equivalent expression to nyarazha is nachari.
b. B: ?? pirkuti-ta-mi wanyuchi-rka Pepe
    rat-ACC-mi kill-PST Pepe
    ‘Pepe killed the rat.’

In (17a), -mi marks the argument, Pepe, that gives new information. In this same context, according to MC, -mi cannot be attached to pirkuti ‘the rat,’ as in (17b).

In (18), a different argument is focused; a different argument, pirkutita is new information, and is therefore marked with -mi.

(18) a. A: imata-ta Pepe wanyuchi-rka
    what-ACC Pepe kill-PST
    ‘What did Pepe kill?’

    B: pirkuti-ta-mi wanyuchi-rka Pepe
    rat-ACC-mi kill-PST Pepe
    ‘Pepe killed the rat.’

b. B: pepe-ka wanyuchi-rka pirkuti-ta
    Pepe-TOP kill-PST rat-ACC
    ‘Pepe killed the rat.’

c. B: Pepe-ka pirkuti-ta-mi wanyuchi-rka
    Pepe-TOP rat-ACC-mi kill-PST
    ‘Pepe killed the rat.’

In fact, in the versions of this response in (18b) and (18c), the topic marker -ka is attached to the old information Pepe, highlighting the topic-focus contrast.

4.2 Sentential/Predicate Focus

The question-and-answer examples presented in the previous section suggest that -mi can be used to mark an argument in focus. This section and the next examine other focus environments. Firstly, the example in (19) shows that -mi can also be used to mark an argument within a larger predicate that is in focus.

(19) a. A: imata-ta Pepe ra-rka
    what-ACC Pepe do-PST
    ‘What did Pepe do?’

    B: pirkuti-ta-mi wanyuchi-rka Pepe
    rat-ACC-mi kill-PST Pepe
    ‘Pepe killed the rat.’

b. B: ?? pirkuti-ta wanyuchi-rka-mi Pepe
    rat-ACC kill-PST-mi Pepe
    ‘Pepe killed the rat.’
In (19), the only presupposed information in speaker A’s question is *Pepe rarka ‘Pepe did.’*

According to MC, the versions of the response in (19b) and (19c), where the particle appears on the predicate, are not preferred. However, MC described the example in (20) as sounding natural, and produced the predicate-focus example in (21) in a spoken text.

(20)  
A: *imata tuku-rka-ngi*  
what happen-PST-2SG  
‘What happened?’

B: *nyuka alku kishpi-rka-mi*  
My dog run.away-PST-mi  
‘My dog ran away.’

(21) *kunan-ga kan-pa kan-pa ya-shka-ta-ka*  
now-TOP 2SG-GEN 2SG-GEN think-PERF-ADVZ-TOP

*apa-sha-mi ni-shka nin*  
take-FUT-mi say-PERF QUOT  
‘“Now your things, your thoughts, I will take,” he said.’

The response in (20) is a good example of sentence focus, given that all of the information in speaker B’s utterance is new (not mentioned in A’s question); it shows that sentence-focus can also be marked with -mi. In (21), there are two topic-marked elements, and the predicate *apa ‘take’* is in focus, marked by -mi.

4.3 Contrastive Focus

-mi also seems to be used as a contrastive focus marker, as in the example in (22).

(22)  
A: *pita miku-rka atalpa-ta Juan o Laura*  
who eat-PST chicken-ACC John or Laura  
‘Who ate the chicken, John or Laura?’

B: *Laura na miku-rka-chu Juan-mi miku-rka*  
Laura NEG eat-PST-NEG John-mi eat-PST  
‘Laura didn’t eat it; John ate it.’

In this case, there are only two candidates for the element in focus; if one, *Laura*, is negated, it is natural that the other, *Juan*, should have contrastive focus, and it is *Juan* that is marked with -mi.

The example in (23) similarly demonstrates contrastive focus.
5 Discussion

5.1 Evidentiality, Epistemic Modality, and Immediacy

The examples presented in Section 3 showed that -mi is licensed in both evidential and epistemic modal contexts. This sub-section presents examples showing that some utterances with -mi have characteristics of both — or, in other words, could be described as hybrid.

First, the obvious conclusion one might draw from the result of the defeasibility test for evidentiality in (8) is challenged by the example in (24), which the same consultant (AO) described as grammatical.

(24) Juan-mi kalpa-rka nyuka na riku-rka-ni-chu
Juan-mi run-PST I NEG see-PST-1SG-NPI
‘Juan ran. I didn’t see it.’

This inconsistency may be due to the fact that I asked about (24) immediately after asking whether the example in (25) could be used in a situation in which Juan had told the speaker about the running. AO said it could, indicating that -mi in (25) is being used more as an epistemic modal marker.

(25) Juan-mi kalpa-rka
Juan-mi run-PST
‘Juan ran.’

It therefore seems likely that AO may have been thinking of (24) as being possible in a similar situation. In both cases, therefore, the utterance is licensed if the information was obtained via report rather than via firsthand direct evidence; this is not the prototypical type of situation where one might expect to find a direct evidential. It seems that the reason (24) and (25) are licensed is the speaker’s belief in and willingness to vouch for the validity of the information; this might be taken to indicate that -mi is functioning in these cases as an epistemic modal marker rather than

---

7 I asked AO in a way similar to the one for (8): first, “How can I say ‘Juan ran’?” Second, “Can I put -mi on Juan and what does that mean?” Then, “Can I say nyuka na rikurkanichu right after that?” In this case, unexpectedly, AO said that the utterance is okay. Presumably, it was prompted right after I had elicited an utterance in EM context, which can be shown in (25).
Imbabura Quechua -mi

an evidential marker.

It should be noted, though, that Faller (2003) argues that -mi cannot be a canonical epistemic modal in Cuzco Quechua, because it does not weaken an assertion. When an epistemic modal — even a strongly positive one — is used in an utterance, it indicates that the speaker is making an assessment about the situation rather than asserting that it pertains. An utterance containing an epistemic modal marker therefore always encodes a weaker speaker validation than a corresponding utterance that does not contain an epistemic modal marker (for example, English *He must have been there* vs. *He was there*). As the presence or absence of -mi does not indicate anything about the speaker’s assessment of the propositional content, Faller’s characterization seems to hold in IQ.

Given that any direct evidential reading conveyed by -mi may be defeasible when the context provides information about an indirect source of evidence, and that it can function as either an evidential or an epistemic modal, we can see that -mi can never be clearly assigned to either functional category, evidential or epistemic modal. This observation conforms to Faller’s (2003) definition of -mi in Cuzco Quechua as a best-possible-evidence marker. The best possible evidence is not always direct visual evidence, and so flexible evidential-licensing conditions allow -mi to assume both epistemic modal and evidential functions.

Secondly, -mi may also have some tense properties. MC produced some examples in which -mi was used as an immediate-tense marker, such as that in (26b), which might be licensed in a situation in which the speaker had seen a few minutes ago that it was raining and was certain that it still was.

(26)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tamya-xu-n</td>
<td>rain-IMPF-3SG</td>
<td>‘It’s raining.’</td>
</tr>
<tr>
<td>b. tamya-xu-n-mi</td>
<td>rain-IMPF-3SG-mi</td>
<td>‘It’s raining.’</td>
</tr>
</tbody>
</table>

MC’s intuition was that -mi in (26b) may be marking the information as having been obtained in the immediate or recent past or it may be marking speaker certainty. In contrast, the sentence in (26a) does not convey that the evidence about the focal information was obtained in the recent past. The recent past meaning of -mi is also relevant to the contrast between the utterances in (27).

(27)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tamya-shka-Ø</td>
<td>rain-PERF-3SG</td>
<td>‘It’s rained.’</td>
</tr>
<tr>
<td>b. tamya-shka-mi</td>
<td>rain-PERF-mi</td>
<td>‘It’s rained.’</td>
</tr>
</tbody>
</table>
According to MC, the sentence in (27b) would likely be used when the speaker had obtained some relevant piece of evidence, for example seeing that the ground was wet, quite recently. On the other hand, -mi would not be licensed when the information had been obtained further in the past, for example the previous evening; utterances like that in (27c) are therefore infelicitous. Presumably, the utterance in (27b) could be used to talk about an event in the more distant past, while the example in (27a) could not.

Interestingly, -mi can also be used to mark recency in combination with a non-past tense marker, as in (28b), which might be licensed in a situation where the speaker had just noticed dark clouds in the sky.

(28)  

a. *kaya tamya-gri-n*  
   tomorrow rain-FUT-3SG  
   ‘It will rain tomorrow.’

b. *kaya tamya-gri-n-mi*  
   tomorrow rain-FUT-3SG-mi  
   ‘It will rain tomorrow.’

In this case, it is possible that -mi is being used to mark the speaker’s certainty, i.e., that it has an epistemic modal function.

In general, I would like to suggest that -mi in the preceding examples also has a mirative function, i.e., that it encodes novel information in the context that has not yet been assimilated into the common ground. In the examples in (27b) and (28b), the speaker has just observed some piece of evidence from which she has inferred the occurrence of the focal event.

Following on my earlier argument that the spatial deictic function and direct evidential function of -mi are naturally related to the positive epistemic modal function. I would like to suggest that the temporal immediacy use is related to all of these other functions. In the utterances in (29) with action verbs, MC emphasized that the function of -mi was to indicate temporal immediacy.

(29)  

a. *nyuka pungu-ta tanga-sha*  
   *I* door-ACC push-FUT  
   ‘I’ll push on the door.’

b. *nyuka pungu-ta tanga-sha-mi*  
   *I* door-ACC push-FUT-mi  
   ‘I’ll push on the door.’

It seems plausible to claim that this use is related to spatial specification and direct evidentiality as well as epistemic modality; firsthand evidentials indicate that the evidence for the focal information is within range of the speaker’s perception (spatial contiguity), and speaker certainty inherently involves the metaphorical extension of spatial contiguity to temporal contiguity. In
Imbabura Quechua -mi

(29b), it is understood that the speaker intends to perform the action of pushing on the door in the very near future. In other words, the course of action is conceptually contiguous to the speaker at the speech time; temporal contiguity are often semantically linked to spatial contiguity by means of the TIME IS SPACE metaphor. This paper suggests that the contiguity evoked by -mi stems from its original evidential function, which entails that the focal information be deictically contiguous to the speaker. This conceptual contiguity is also related to -mi’s epistemic modal function, in that the more contiguous to the speaker the stimulus is in various dimensions, the stronger an epistemic stance the speaker can take. This relationship between functions is employed strategically in the textual example in (30).

(30) \textit{wanyu-gri-ngi-mi}
\textit{die-FUT-2SG-mi}

‘You \textbf{will die}!’

In this example, -mi can be described as an evidential/epistemic modal marker that encodes temporal/conceptual/causal contiguity.

The preceding argument suggests that the functions of -mi in Imbabura Quechua as a direct evidential marker, as an epistemic modal marker, as a marker of spatial contiguity, and as a marker of temporal contiguity are all related. As I noted above, Floyd (1999) suggests that the uses of -mi in Wanka Quechua form a radial category of multiple functions. It seems quite plausible that -mi in IQ has hybrid functionality, describable as a direct evidential/epistemic modal/deictic marker.

5.2 Evidentiality and Epistemic Modality as a Conceptual Package

The previous sections have demonstrated that the particle -mi in IQ is licensed in a variety of contexts, including evidentiality, epistemic modality, spatial and temporal immediacy, and discourse focus; these facts present strong counterevidence for any hypothesis about -mi that assumes it is either a pure evidential or a pure epistemic modal marker that is used exclusively for only one of those possible functions. Given that Quechua speakers use -mi freely in both contexts, and if the pragmatic implicatures of the marker in any context are not strictly constrained to the implicatures typical of evidentiality nor the implicatures typical of epistemic modality, but are flexibly interpreted in the given context, distinguishing precisely between the two categories may simply not be relevant to understanding the linguistic structure of Quechua.

Every epistemic modal construction has some evidential properties, whether they are explicit or implicit in a given context, and every evidential construction carries some epistemic implicatures (Kwon, to appear). For instance, in English, for someone to say, “You must be Seth Sweetser’s sister,” they must have obtained some kind of solid evidence, such as the name on the addressee’s nametag (Sweetser 1990). Without any evidence to induce her conclusion, the speaker would not have been certain. If it is invariably the case that each of these functional categories carries some of the semantics or pragmatics of the other, it would be more meaningful to investigate how each of the categories is construed in the context of the conceptual structure evoked by both.

If we assume that both categories evoke the same causal event structure, we do not need to worry about the separability of the categories (or their labels). The semantic and pragmatic properties of evidentiality and epistemic modality are not
determined each in isolation; rather, they are determined relative to each other, because they relate to different conceptual portions of the same causal event structure. This causal event structure involves an event in which the origo perceives some relevant piece of evidence, an event in which the origo assesses the situation based on that evidence, and an event in which the origo encodes her perception and/or the assessment of the evidence in a speech act about the situation. These three events are connected by the speaker’s inference processes. This structure is schematized in Figure 1.

![Diagram of causal event structure](image)

Figure 1: A schema of the causal event structure of epistemicity

In Sweetser’s (1990) force dynamics approach to epistemic modality, epistemic modality is conceptualized as a tension between the origo’s reasoning process (the force) and any potential constraints (barriers), such as counter-evidence. Figure 1 represents the schematic structure of evidentiality and epistemic modality in a more fine-grained way in that it breaks the force down into smaller pieces (or frame elements), including the origo’s perception of the evidence, the belief she forms based on that evidence, and her reasoning about the focal event, which is induced or “pushed” by her perception and/or inference.

This approach supports the contention that separating the two functional categories according to strict criteria is not the important issue. In fact, it implies that neither of these notions can be fully defined independently of the other; rather, I would like to argue that what phenomena should be considered to belong to the category of evidentiality, what should be considered epistemic modality, and what should be considered to belong to some hybrid evidentiality/epistemic modality category is determined in a given language by which portion of the causal event chain is profiled and how that profiling is semantically encoded in linguistic expressions. In IQ, both evidentiality and epistemic modality are marked with -mi, with flexible overlap between the two functions.

5.3 Similarity Between Evidentiality/Epistemic Modality and Focus Marking

This sub-section explores the second major hypothesis of this paper, that evidentiality/epistemic
modality and focus marking are functionally related to each other. It does not seem implausible to suggest that a single linguistic item could be used for all of these functions, as the categories share the cognitive function of foregrounding the portion of the linguistic content that the origo believes has not yet been assimilated into the addressee’s knowledge nor the common ground. Information that is new in information structure terms and the focal information in an evidential construction are both likely to be conceptually salient in their contexts and set apart from information that is already shared among and assimilated by all of the interlocutors.

Faller (2003) has shown that evidential constructions generally convey information that is new to the addressee, and a number of authors have pointed out that evidential constructions introduce presuppositions that the origo obtained the focal information in a particular way (Matthewson et al. 2006, Speas 2008, Kwon 2009). When a speaker produces an utterance with an evidential marker, the addressee expects that the marked information will be something she has not yet assimilated into her knowledge. The addressee automatically accommodates the presupposition that the speaker has obtained the marked information through some particular mode of access. For example, in the sentence Pepe killed the rat-\textit{mi}, the focal information is the event of Pepe’s killing the rat, not the presupposed event of the speaker’s obtaining that information. In other words, the focal event marked by -\textit{mi} is more novel and more cognitively salient and is therefore foregrounded.

If evidentials in general mark new information, the question arises of whether other types of evidentials can appear in focus position. The indirect evidential -\textit{shi} seems to occasionally appear in focus position in IQ, as in the example in (31b); Weber (1989:419) describes a similar pattern in Huallaga Quechua.

\begin{enumerate}
\item a. A: \textit{pita wanyuchi-rka pirkuti-ta} \quad \textit{who kill-PST rat-ACC}
\textit{‘Who killed the rat?’}

\item B: \textit{Pepe-mi wanyuchi-rka pirkuti-ta} \quad \textit{Pepe-mi kill-PST rat-ACC}
\textit{‘Pepe killed the rat.’}

\item b. B: \textit{Pepe-shi wanyuchi-rka pirkuti-ta} \quad \textit{Pepe-shi kill-PST rat-ACC}
\textit{‘Pepe killed the rat.’}
\end{enumerate}

Although the response in (31b) is not considered perfectly felicitous, AO does not judge it completely ungrammatical. It seems natural that evidentials in general should be used with novel information, as they are usually used in situations in which a speaker wants to disseminate information that (as far as she knows) has not yet been assimilated into the addressee’s knowledge.\footnote{The fact that -\textit{mi} is used more frequently than -\textit{shi} in focus position is not surprising given than speakers who are answering questions are likely to vouch for the validity of the information they are covering.}

More supporting evidence for the claim that the evidential/epistemic modal and focus marking

\footnote{Underlining is used here to represent the indirect-evidential meaning conveyed by -\textit{shi} that cannot be expressed in English with a single morpheme, but is rather expressed periphrastically.}
functions of -mi are related is that only one -mi is licensed in a simple clause (Cole 1982 and Muysken 1995), as demonstrated by the ungrammaticality of the example in (32).

(32) *nyuka mama-mi wacha-ri-rka-mi Seoul-pi.
    my mother-mi born-REF-PST-mi Seoul-LOC

‘My mother was born in Seoul.’

If the evidential/epistemic modal and focus marking functions of -mi were completely distinct (i.e., if -mi were simply polysemous), the sentence in (32) should have been acceptable in a context where the speaker had strong evidence that her mother, rather than some other person relevant in the context, had been born in Seoul.

6 Concluding Remarks

In this paper, I have analyzed the functional properties of the enclitic -mi in Imbabura Quechua; as in other Quechua languages, -mi in IQ has multiple functions, marking direct evidentiality, strongly positive epistemic modality, and discourse focus. I discussed a number of contexts in which it is not possible to define with clarity which function -mi is performing, including performatives, irrealis events, and temporally or spatially immediate events. Based on these observations, I explored the possibility that the three listed functions of -mi are not distinct, but are conceptually related to one another. I argued that -mi cannot be classified as a pure evidential marker or as a pure epistemic modal marker, but performs both functions, and positioned my argument within the general debate about whether evidentials and epistemic modals can belong to a single functional category. I suggested that, to better understand why -mi is licensed in both evidential and epistemic modal contexts in Imbabura Quechua, we should consider the causal event structure of epistemicity as a whole, including how an experiential origo perceives evidence for information and assesses that information. I further suggest that making a strict distinction between the two categories is not an essential goal, but that characterizing each in relation to the other is key to better grasping the nature of the debate and the nature of evidentials and epistemic modals themselves.

The second major question discussed in this paper was whether the evidential/epistemic modal meaning of -mi can be functionally extended to explain its use as a focus marker. I argued that the evidential/epistemic modal function is conceptually related to the function of encoding focus in that, in both cases, the origo’s attention is directed to the cognitively more salient portion of the relevant stimulus. In evidential/epistemic modal contexts, the target event is cognitively salient, whereas the origo’s mode of access to the information and her degree of belief in its validity is presupposed and backgrounded. In discourse focus contexts, the focused element is cognitively salient while the other elements in the construction are given less attention.

References


Iksoo Kwon
English Linguistics Department
Hankuk University of Foreign Studies
107 Imun-ro, Dongdaemun-gu, Faculty Building #512
130-791 Seoul
South Korea
kwoniks@hufs.ac.kr
The Stops of Tlingit

IAN MADDIESON\textsuperscript{1,2} and CAROLINE L. SMITH\textsuperscript{2}
\textsuperscript{1}University of California, Berkeley and \textsuperscript{2}University of New Mexico

Introduction

One aspect of the maintenance of languages is the maintenance of the appropriate pronunciation. A language whose speakers use it only a minority of the time is often strongly affected by the phonetic patterns of the dominant language, particularly as speech habits natural to the ‘new’ language are carried over to the traditional one. Documenting the phonetic patterns of those speakers who were the last to have learned the language in question as their first language can provide an appropriate source of guidance to the pronunciation to adopt for later generations. In this paper we aim to present a good description of certain important aspects of the pronunciation of the stops of Tlingit. We hope that this data will be of value to those interested in the continuation of Tlingit as a spoken language. Of course, all languages change over time, and pronunciation will evolve in natural ways, so this information should be used in ways that accommodate natural changes. This data will also have value for comparative phonetic studies as illustrating details of a language whose phonetics has been previously little studied.

Tlingit ([\textipa{tʃǐn̩ɪt}]) is a Na-Dene language (Krauss 1979) spoken in South-East Alaska and the Yukon. There are now few fluent speakers under the age of 60 (Dauenhauer and Dauenhauer 1995). The language has some dialect variation, with particular differences noted between Inland or Yukon Tlingit and Coastal or Alaskan varieties, and further differences between Northern, Central, and Southern varieties along the Coast. In this paper, all data is taken from what we are calling the Central dialect. Like its Athabaskan relatives, Tlingit has a quite complex consonant inventory, in particular having many affricates and fricatives. It also has a striking absence of plain labial consonants. More details on Tlingit phonetics are given in Maddieson, Bessell, and Smith (2001).

The basic stop inventory is outlined in Table 1 using standard phonetic symbols. There are stops at three places of articulation — alveolar, velar, and uvular, and with three manners of articulation — voiceless unaspirated, voiceless aspirated, and ejective. In addition to these nine stops there are also labialized velar and uvular stops in all three categories. The corresponding representations of these stops in Tlingit orthography are shown in Table 2. The voiceless unaspirated stops are written orthographically with the letters that would normally represent voiced stops, and the voiceless aspirated ones with the letters regularly used for voiceless stops in general. (This convention is also common in orthographies for Athabaskan languages.) Uvulars are written with the same letters used for velars but with an underline.

<table>
<thead>
<tr>
<th></th>
<th>Alveolar</th>
<th>Velar</th>
<th>Uvular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiceless Unaspirated</td>
<td>t</td>
<td>k</td>
<td>q</td>
</tr>
<tr>
<td>Voiceless Aspirated</td>
<td>t\textsuperscript{h}</td>
<td>k\textsuperscript{h}</td>
<td>q\textsuperscript{h}</td>
</tr>
<tr>
<td>Ejective</td>
<td>t\textsuperscript{'}</td>
<td>k\textsuperscript{'}</td>
<td>q\textsuperscript{'}</td>
</tr>
</tbody>
</table>

Table 1: Inventory of non-labialized stops in Tlingit
I AN MADDESON and CAROLINE L. SMITH

<table>
<thead>
<tr>
<th></th>
<th>Alveolar</th>
<th>Velar</th>
<th>Uvular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiceless Unaspirated</td>
<td>d</td>
<td>g</td>
<td>g</td>
</tr>
<tr>
<td>Voiceless Aspirated</td>
<td>t</td>
<td>k</td>
<td>k</td>
</tr>
<tr>
<td>Ejective</td>
<td>t'</td>
<td>k'</td>
<td>k'</td>
</tr>
</tbody>
</table>

Table 2: Tlingit orthography for stops

The Distribution of Tlingit Stops

The three series of stops contrast very straightforwardly in initial position, as in the triplet in Table 3.

<table>
<thead>
<tr>
<th>Transcription</th>
<th>Orthography</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>/qákw/</td>
<td>gakw</td>
<td>‘tree spine’</td>
</tr>
<tr>
<td>/qʰákw/</td>
<td>kakw</td>
<td>‘basket’</td>
</tr>
<tr>
<td>/qákw/</td>
<td>k’akw</td>
<td>‘screech owl’</td>
</tr>
</tbody>
</table>

Table 3: A minimal triplet illustrating contrast between initial stop manners

The same three-way contrast also occurs in syllable-onset position in the middle of words. However, there are only two contrastive possibilities in word- or stem-final position, ejective and non-ejective. We will describe some aspects of the three-way distinction as it is produced in initial position, and then ask the main question that we wish to address in this paper, which is: what are these non-ejective stops in final position? We may entertain three obvious possible answers to this question:

A. They are the same as the aspirated stops in initial position.
B. They are the same as the unaspirated initial stops.
C. They are unlike either of the initial series, perhaps intermediate between the two.

The standard analysis accepted by previous writers is to treat them as aspirated, i.e., to accept possibility A above. This is clearly stated or quite expressly implied in one way or other by Boas (1917), Pinnow (1966), Leer (1978), and Dauenhauer and Dauenhauer (1991). The Tlingit orthography also treats them as aspirated in that the final non-ejective stops are written with the letters “t, k, k” rather than “d, g, g”.

The Definition of Aspiration

It may be useful at this point to clarify what we understand by aspiration. Aspiration is an interval after the release of a voiceless consonant during which noisy airflow is sustained. The vocal folds remain in the voiceless position — that is, slightly apart — after the articulatory movement in the mouth is over, and the flow of air through the space between the vocal folds,
The Stops of Tlingit

the glottis, creates the characteristic noise. The acoustic and articulatory patterns involved are illustrated in Figure 1, which shows an annotated spectrogram of the beginning part of the Tlingit word /tʰaan/ ‘sea lion’ spoken by a male speaker. A time interval of about 450 milliseconds is shown in the figure. The approximate timing of the major articulatory phases involved is indicated by the boxes below the spectrogram. The precise moment at which the tongue makes contact with the alveolar ridge cannot be determined, but it is likely to be about 100 ms before the release, which is visible as the moment when the first acoustic signal is registered at the end of the ‘tongue closure’ box. The vocal folds come together into the voicing position a little after the 300 ms time point. The interval between the tongue release and the onset of voicing for the following vowel is filled with relatively high amplitude noise, with energy quite broadly distributed over the frequency range shown. On the spectrogram this is apparent as the heavy gray shading during this interval.

![Spectrogram](image)

Figure 1: Acoustic and articulatory timing of aspiration in word-initial position

The interval between a stop release and the onset of voicing for a following segment is often referred to as the ‘voice onset time,’ abbreviated VOT. In Figure 1, the voice onset time is about 130 ms. It is sometimes suggested that an appropriate way to define a stop as aspirated is that there is a relatively long delay between release and voice onset. There are at least two reasons why this is not appropriate. First, this definition could only apply to stops which were in fact followed by a voiced segment, such as vowel. Hence, the definition itself would disallow word-final aspirated stops, or aspirated stops before another voiceless segment. Secondly, a delay in the onset of voicing can be due to a number of different causes. Of particular relevance is the fact that delayed voice onset can occur because the vocal folds remain fully closed for some time
after the oral release of a stop. This is the case with the ejective stops of Tlingit, where the length of the interval between the release of the oral closure and the onset of a following vowel is about the same length as in word-initial aspirated stops. However, in the case of an ejective stop, the interval between the short burst of noise created by the oral stop release and the onset of voicing is filled with (near-)silence, rather than with the sustained noise that is characteristic of aspirated stops. This is illustrated in the annotated spectrogram in Figure 2, which shows the beginning of the word /t'aaw/ ‘(large) feather’ spoken by a female speaker. Here, the release of the tongue contact for /t/ occurs at about the 150 ms time-point, but the vocal folds maintain a complete closure for a considerable period after this release. Because no air is flowing out during this interval it is largely silent — that is, white on the spectrogram — rather than noisy. The onset of voicing, which occurs about 160 ms after the tongue contact is released, is due to relaxing a complete closure of the vocal folds, rather than to narrowing the glottis from the open position usually associated with voicelessness.

Figure 2: Acoustic and articulatory timing of an ejective stop in word-initial position

Across a substantial number of tokens, the mean value of the VOT for both the aspirated and the ejective stops in initial position is over 100 ms, as is shown in Figure 3. The VOT is considerably longer (four or five times) for both these categories of stops compared to the VOT for the voiceless unaspirated stops in the same position. The means in the figure represent 25 tokens in each category, taken from two (occasionally three) repetitions by four speakers of words with initial stops at all three places of articulation. The words were spoken in isolation as part of a recorded wordlist, and so are always utterance initial.

Since the ejective stops share with the voiceless aspirated stops the property of having a long
period of VOT, this property alone is clearly not sufficient to characterize aspirated stops. Another description of aspirated stops as being followed by ‘a puff of air’ is also unhelpful, since a puff of air will be expelled following any released stop. What is crucial about the identification of aspirated stops as aspirated is that their release is followed by a sustained period of noisy airflow, rather than only a short noise period (as for unaspirated stops) or a sustained (near-)silence (as for ejective stops).

![Figure 3: Mean durations of VOT in word-initial stops in Tlingit (in milliseconds)](image)

We are now ready to consider whether it is appropriate to regard the final voiceless stops in Tlingit as exemplifying the aspirated or the unaspirated category, or if they are different enough from both of the voiceless stop categories that appear in initial position to be considered a different class from either. Both phonological and phonetic evidence will be taken into account.

**Vowel-Initial Suffixes**

There is a good phonological argument in favor of treating the final non-ejective stops as belonging to the unaspirated category: they are clearly not aspirated when they are followed by one of the relatively few vowel-initial suffixes. One such suffix is the genitive suffix whose base form can be considered to be /-i/ with polar tone. This suffix occurs in the phrase [αχχαατί] ‘my fish,’ orthographic ax xaadi, which is illustrated by the spectrogram in Figure 4. The stem-final stop release occurs a little after the 800 ms time-point in this figure, and the onset of voicing for the following vowel occurs only about 20 milliseconds later.
A consequence of orthographically representing the final non-ejective stops as aspirated, as in the word *xaat* ‘fish’ is that the spelling of the stem must be changed when it is suffixed, as in the possessed form shown here, *ax xaadi* ‘my fish.’ The pattern in such cases was noted by Boas (1917) who explicitly chose to analyze it as a process involving replacing final aspirated stops with their unaspirated counterparts (actually, in his terminology, as replacing surds with sonants). That is, for Boas the stem-final stop consonant contrast was between ejective and aspirated stops when it occurred in coda position, but between ejective and unaspirated stops when the stem-final stops shifted to onset position. This analysis is accepted by all later writers. In particular, this is the case in the morphological analysis of Story (1963). Although she opts to write stem-final non-ejective stops with the symbols she uses for unaspirated stops, she clearly states that this is simply for the convenience of maintaining a constant transcriptional shape for the stems, which she considers to be phonologically alternating.

Cross-linguistically an alternation which introduces aspiration in coda position is somewhat surprising, since in many of the languages with aspiration contrasts in onset position, aspiration is disallowed in coda position (e.g., Thai, Korean, Cantonese, Navajo, Hupa, Tanana, and other Athabaskan languages). A simpler and more easily motivated phonology would posit no phonological alternation between stem-final non-ejective stops in coda and onset positions. We suggest that this is in fact the case, and that these stops are always unaspirated. The shape of the stem is faithfully preserved in all forms. Preliminary phonetic support for this claim is provided by the spectrogram of the unsuffixed form of the word for ‘fish’ (orthographic *xaat*) in Figure 5. The final stop in this word is similar in all major aspects to the intervocalic stop shown in Figure 4. In particular its release, at about the 680 ms time-point is followed by only a short noisy interval, suggesting that the appropriate transcription is indeed */χaat/* not */χaatʰ*/.
Are Final Voiceless Stops Aspirated?

In order to compare initial aspirated stops and utterance-final non-ejective stops more systematically, a set of measurements was made of the amplitude of the noise following release in the words given in Table 4. We would expect aspirated stops to have a higher amplitude for a longer period of time following their release than would be observed in unaspirated stops. Data from four speakers, three female and one male, was examined. To eliminate possible effects of vowel quality and word length only monosyllabic words with the long vowel /aa/ were used. Uvular stops were excluded from this analysis, because they often have the complicating factor of a substantial duration of affrication following their releases. Therefore only alveolar and velar stops are shown. Typically two repetitions of two words from each of the speakers were measured, but there are fewer tokens of initial /kʰ/ than of the other cases. The data measured includes a total of 33 final cases and 25 initial ones.
Table 4: Words measured for comparison of initial and final stops

<table>
<thead>
<tr>
<th>Transcription</th>
<th>Orthography</th>
<th>Gloss</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial /tʰ/</td>
<td>tʰaan</td>
<td>‘sea lion’</td>
<td>F1, M1, F2, F3</td>
</tr>
<tr>
<td></td>
<td>tʰaat</td>
<td>‘night’</td>
<td>F1, M1, F2, F3</td>
</tr>
<tr>
<td>Initial /kʰ/</td>
<td>kʰaak’</td>
<td>‘forehead’</td>
<td>M1, F2, F3</td>
</tr>
<tr>
<td></td>
<td>kʰaa</td>
<td>‘measuring stick’</td>
<td>F2</td>
</tr>
<tr>
<td></td>
<td>kʰáax’</td>
<td>‘chicken, grouse’</td>
<td>F1</td>
</tr>
<tr>
<td>Final /t/</td>
<td>tʰaat</td>
<td>‘night’</td>
<td>F1, M1, F2, F3</td>
</tr>
<tr>
<td></td>
<td>xaat</td>
<td>‘roots’</td>
<td>F1, M1, F2, F3</td>
</tr>
<tr>
<td>Final /k/</td>
<td>saak</td>
<td>‘eulachon’</td>
<td>F1, M1, F2, F3</td>
</tr>
<tr>
<td></td>
<td>ɬ’aak</td>
<td>‘dress’</td>
<td>F1, F2, F3</td>
</tr>
<tr>
<td></td>
<td>jaak</td>
<td>‘mussels’</td>
<td>M1</td>
</tr>
</tbody>
</table>

An amplitude curve and the way that measurements were made is illustrated by Figure 6. The amplitude of a speech sound is the amount of acoustic energy it has. A speech waveform, as at the bottom of this figure, displays the pressure variation in the signal both positive and negative as a series of instantaneous amplitude values. But the measure that corresponds to the total energy of a signal is a measure of the absolute value of the amplitude, usually called the RMS (root mean square) amplitude. To relate to sensations of loudness in the signal, this needs to be calculated over a certain window of time. For this analysis, a window 10 ms long was selected, which was moved along in 1 ms steps, so that values are reported at 1 ms intervals for overlapping time slices. The resulting amplitude curve — on a greatly enlarged scale — is shown in a lighter shade above the waveform. The highest amplitude occurs during the vowel, with lesser peaks after the two stop releases, whose times are marked by vertical gray lines. Note that because of the length of the analysis window, the amplitude curve appears slightly delayed with respect to the speech waveform.

In Figure 6 the amplitude remains quite high between the release of the closure of the initial stop and the onset of the following vowel, that is, during the aspiration interval for /tʰ/. In final position, although there is a short peak with nearly the same maximum, the amplitude drops off rapidly. A systematic analysis was performed in order to quantify such differences. The amplitude values were averaged over an interval of 80 ms beginning 20 ms after each stop release. The first 20 ms were not included in order to exclude the high amplitude of the release burst itself. The next 80 ms was measured so as to include most of the aspiration, whose duration is over 100 ms, as shown by Figure 3.
Amplitude is measured on the logarithmic scale of decibels (dB), and a difference of 6 dB corresponds roughly to an impression of double the loudness to a listener. In aggregate, in the data measured the mean amplitude over the 80 ms interval after initial stops was 25.0 dB, but the mean amplitude after final stops was only 15.8 dB. That is, it was well under half as loud on average. In statistical terms we may say that the initial and final amplitude values are significantly different (in a two-way analysis of variance with position and speaker as main effects, $F (1,50) = 14.26$, $p < .0001$ for position). The results are shown for each individual speaker in Figure 7. Although the levels for different speakers are also significantly different, in each case the noise amplitude after final non-ejective stop releases is at least 7 dB lower than after initial aspirated stops.


Suggested Interpretation

Both the phonological and phonetic evidence presented above suggest that the final non-ejective stops are in fact unaspirated. Phonologically, this means that there is no alternation between an unaspirated form in onset position (preceding a suffix) and an aspirated form in coda position. Such an alternation would be surprising in view of the general disfavoring of aspiration in final position noted in many of the world’s languages. Phonetically, the rapid reduction in amplitude following the release of final stops confirms that they are unaspirated. The RMS amplitude does rise briefly to a fairly high level immediately at the release of a final stop, but this peak is of very short duration, and corresponds to the audible release of the stop closure.

We suspect that the description of these Tlingit final stops as aspirated arises from a property of the phonetic grammar of the language which specifies that the release of utterance-final stops must always be audible. This is in quite marked contrast to English, where the final stop in a word such as hat is often pronounced without an audible explosion. It is likely that the invariable presence of an audible release struck English (and German) speaking linguists as making the final stop releases in Tlingit saliently stronger than those they were familiar with. However, as was emphasized earlier, the presence of an audible release is not the same as aspiration, which involves a sustained period of outward airflow.

The tendency to interpret final released non-ejective stops as aspirated may have been reinforced by an effect which can be noted on some occasions when such stops occur in words spoken in isolation. When an utterance is finished, a speaker usually opens the vocal folds wide into the position used for normal respiration. This posture is considerably wider than that for voiceless segments in speech. When respiration follows closely on the heels of an audible stop release, the release may be followed by some noise due to exhalation. This noise is of much lower amplitude than true aspiration noise, generated when the vocal folds are held in the voiceless position.
What About Word-Final Stops Not Utterance-Final?

The amplitude measurements presented above were for stops in utterance-final position. In this position, an audible release occurs without exception. To investigate how general the requirement for an audible release is in Tlingit, we also examined word-final stops that were not utterance-final. That is, we examined non-ejective stops occurring at the end of a word in the middle of a phrase, preceding another word. Because the data collection protocol had not originally been designed with this comparison in mind, suitable recordings were available only from one female speaker, identified by the code F2. In the phrases selected for this analysis, the second word always began with a voiceless stop or affricate, since a release of the first consonant can be unambiguously seen in such sequences. An example of the phrase /tɬeet qʰáa/ (dleit káa) ‘white person’ (literally ‘snow person’) is shown in Figure 8 (note that the lateral component of the initial /tɬ/ segment in this particular utterance was pronounced as an approximant, not a fricative). In this example, the release of the /t/ at the end of the first word is clearly visible shortly after the 300 ms time mark, followed by a separate closure for the uvular stop /qʰ/ about 50 ms later.

In all the examples examined, when the word-final and word-initial consonants are at different places of articulation, as in Figure 8, the first consonant always has an audible release before the silent period for the voiceless stop closure at the beginning of the second word. Thus the two closures never overlap in time and the silent periods for the two stops are separated by the release noise, making it easy for a listener to tell that there are two distinct stops being pronounced. In particular, the release noise provides information about the place of articulation of the first stop, which would otherwise be more difficult to extract from the signal.

In order to determine whether these word-final, phrase-internal voiceless stops are more like the aspirated or unaspirated stops that occur in syllable onsets in their timing, the durations of the
releases of the word-final stops were measured and compared with the durations of voice onset times for word-medial aspirated stops in words such as /qaatʰaa/ ‘steel trap’ for the same speaker. The release duration is the interval from the beginning of the release burst for the first stop until the onset of the closure for the second (as deduced from the acoustic displays). Comparing the word-final stops with the word-medial aspirated stops, rather than the word-initial, may be more appropriate because in English, and likely in many other languages, consonants in utterance-initial position are longer than consonants later in an utterance. Mean results are given in summary form in Figure 9.

The mean release duration of the word-final stops is notably longer than the voice onset time for utterance-initial voiceless unaspirated stops reported in Figure 3. It also depends considerably on the place of articulation, being longer for consonants articulated further back in the mouth. The release duration of the word-final stops was shorter than word-medial aspiration for /t/ and /k/, and longer for /q/. As mentioned above, the release of /q/ is often accompanied by a substantial amount of affrication, which accounts for the longer release duration at this place of articulation. On the other hand, word-medial aspiration duration, measured as the Voice Onset Time, was relatively uniform across different places of articulation; hence a single mean is shown.

![Figure 9: Duration of release for phrase-internal word-final stops and aspiration for word-medial aspirated stops](image)

Because the number of tokens varied for different combinations of consonants, a more detailed breakdown of the results is given in Table 5. The data is not extensive, but it is sufficient to suggest that while the place of articulation of the first consonant is important, the place of the following consonant has no major impact on the length of the release interval. Table 5 confirms the relative stability of VOT measures for medial aspirated stops at different places. The difference in place sensitivity between release duration and aspiration duration can be taken as a further reason not to identify the final stops as belonging to the aspirated category.
The Stops of Tlingit

<table>
<thead>
<tr>
<th>Word-final stops</th>
<th>Followed by</th>
<th>Number of tokens</th>
<th>Mean duration of release (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/t/</td>
<td>velar</td>
<td>24</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>uvular</td>
<td>7</td>
<td>46</td>
</tr>
<tr>
<td>/k/</td>
<td>coronal</td>
<td>5</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>uvular</td>
<td>1</td>
<td>68</td>
</tr>
<tr>
<td>/q/</td>
<td>coronal</td>
<td>3</td>
<td>112</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Word-medial aspirated stops</th>
<th>Mean duration of aspiration (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/tʰ/</td>
<td>8</td>
</tr>
<tr>
<td>/kʰ/</td>
<td>14</td>
</tr>
<tr>
<td>/qʰ/</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 5: Duration measures and number of tokens for word-final phrase-medial voiceless stops and word-medial aspirated stops

When a word-final stop is at the same place of articulation as a following word-initial stop or affricate in a phrase, creating what is known as a homorganic sequence of consonants, there is a quite different pattern to observe. In this case, an audible release of the first stop is not required. For example, in the /t t/ sequence from the phrase /tlee ttaak wusitʰán/ ‘it is snowing’, shown in Figure 10, the first /t/ is not separately released. Instead, the two stops are produced with a single long, geminate-like, closure.

Figure 10: Sequence of stops [t t] at the same place of articulation
To examine how often a release was present or absent in such cases, a set of phrases in which a word-final /t/ or /k/ occurred before a word-initial consonant at the same place of articulation was examined from the tape available for speaker F2. No tokens were available of a uvular stop followed by a uvular. The number of tokens found and the number of these sequences in which no release of the first consonant could be heard, or seen in a spectrogram, is given in Table 6. Overall, the word-final consonant was not released in almost two-thirds of the homorganic sequences examined.

<table>
<thead>
<tr>
<th>Word-final stop</th>
<th>Followed by</th>
<th>Number of tokens</th>
<th>Number of tokens with no release</th>
<th>Percentage of tokens with no release</th>
</tr>
</thead>
<tbody>
<tr>
<td>/t/</td>
<td>[t], [tʰ], or [tʃʰ]</td>
<td>19</td>
<td>13</td>
<td>68%</td>
</tr>
<tr>
<td>/k/</td>
<td>[k] or [kʷʰ]</td>
<td>7</td>
<td>4</td>
<td>57%</td>
</tr>
</tbody>
</table>

Table 6: Number of tokens for word-final voiceless stops followed by a word-initial stop at the same place of articulation

Since the place of articulation is the same in these sequences, the release of the first stop does not provide any non-redundant information to the listener, provided there is a long closure duration to signal the presence of two segments. Thus the release is not perceptually functional in the same way that it is in the sequences with different places of articulation. We may also consider that the freedom to suppress a release in such contexts supports the interpretation of the final non-ejective stops as unaspirated. This is because the absence of a release would be quite odd if these stops were distinctively specified as aspirated, as aspiration cannot be manifested unless a release is present.

In a follow-up study we hope to analyze the quality and amplitude of the release duration noise which occurs in phrase-medial position, and to look at a wider range of consonant sequences, with data from a larger number of speakers.

Summary

We have shown that Tlingit follows the pattern found in Athabaskan languages of having a contrast between unaspirated and aspirated voiceless stops in initial position but in having only the unaspirated series in final position. Both phonological patterning and detailed phonetic analysis support the conclusion that word-final non-ejective stops should not be considered as aspirated. Utterance-final unaspirated stops are produced with an audible release, but measurements show that this release is not comparable to aspiration. The release noise is shorter and has lower amplitude than the noise of aspiration. Furthermore, a release is found in almost all word-final stops in phrase-internal positions; the cases where it is not produced being just those contexts where the release would contribute little to perception — that is, when the stop is followed by another stop or affricate at the same place of articulation. This pattern suggests that rather than being an essential property (as aspiration noise is for aspirated stops), the release is instead a consequence of speakers maximizing the perceptual distinctiveness of place of articulation cues in word-final stops. One aspect of speaking Tlingit correctly is to respect the distribution of such cues in the appropriate places.
The Tlingit orthography is helpful in showing that only one type of non-ejective stop occurs in final position, but misleading as to the correspondence between the final stops and the contrasting stops that occur in initial position. But so long as speakers are aware of the appropriate pronunciations, as laid out in the ‘phonetic grammar’ of the language, this is not likely to be a problem.

Acknowledgements

The data reported here were collected in August 1996 in Juneau through the good offices of Richard and Nora Dauenhauer and the Sealaska Heritage Foundation. We are extremely grateful to the Dauenhauers and all the speakers of Tlingit who graciously shared their knowledge with us. We also owe a great deal to our collaborator Nicola Bessell who originated the idea of a research trip to study Tlingit phonetics in more detail. Major support for this project was provided by the National Science Foundation through grant SBR-9319705 to Peter Ladefoged and Ian Maddieson for research on the phonetic structures of endangered languages.

References


Leer, Jeff. 1978. Introduction to Tongass Texts by Frank Williams and Emma Williams, 6-17. Fairbanks, AK: Alaska Native Language Center.


The Plank Canoe of Southern California: Not a Polynesian Import, but a Local Innovation

Yoram Meroz

By nearly a millennium ago, Polynesians had settled most of the habitable islands of the eastern Pacific, as far east as Easter Island and as far north as Hawai‘i, after journeys of thousands of kilometers across open water. It is reasonable to ask whether Polynesian voyagers traveled thousands of kilometers more and reached the Americas.

Despite much research and speculation over the past two centuries, evidence of contact between Polynesia and the Americas is scant. At present, it is generally accepted that Polynesians did reach South America, largely on the basis of the presence of the sweet potato, an American cultivar, in prehistoric East Polynesia. More such evidence would be significant and exciting; however, no other argument for such contact is currently free of uncertainty or controversy.1

In a separate debate, archaeologists and ethnologists have been disputing the rise of the unusually complex society of the Chumash of Southern California. Chumash social complexity was closely associated with the development of the plank-built canoe (Hudson et al. 1978), a unique technological and cultural complex, whose origins remain obscure (Gamble 2002).

In a recent series of papers, Terry Jones and Kathryn Klar present what they claim is linguistic, archaeological, and ethnographical evidence for prehistoric contact from Polynesia to the Americas (Jones and Klar 2005, Klar and Jones 2005). At the core of their argument is the proposal that the sewn-plank canoe appeared among the Chumash and neighboring Gabrielino people of Southern California through the arrival there of Polynesians using similar boats.2 This work has generated interest among students of North American and Oceanic prehistory (Nicolay 2005, 2007; Rick et al. 2005:208; Clarke et al. 2006:894; Kirch and Kahn 2007:200; Weisler and Green 2008; Bentley et al. 2007:645; Matisoo-Smith 2009:160; Raab et al. 2009:220, Matisoo-Smith and Ramirez 2010:85), attention in the popular press (Edgar 2005; Davidson 2005; Smith 2011), and some criticism (Anderson 2006; Arnold 2007; Lawler 2010:1347).

In this paper, I give a comprehensive review of Jones and Klar’s arguments. I conclude that they fail to demonstrate prehistoric contact between Polynesia and Southern California. Instead, a review of the linguistic, technological, archaeological and ethnological evidence supports a new scenario in which the plank canoe was independently elaborated in California from earlier dugout boats, long before the settlement of East Polynesia.

---

1 Storey et al. (2007) have recently claimed that chicken remains found in Chile are genetically Polynesian and are in a pre-European context (see discussion in Gongora et al. 2008a, 2008b; Storey et al. 2008; Storey et al. 2011). There is preliminary evidence that human remains from near the Chilean coast may be Polynesian (Matisoo-Smith and Ramirez 2010; Matisoo-Smith 2011). Green (2000) and Clarke et al. (2006) present suggestive but not fully conclusive evidence for an American origin of some varieties of the Polynesian bottle gourd. Other than the word for sweet potato, there is no accepted linguistic evidence for early Polynesian-American contact.

2 In this paper, I use the name Gabrielino for the language now usually called Tongva by the descendants of its original speakers. I use the common ‘Hawaiian’ for the more correct Hawai‘ian, ‘Maori’ for Māori, ‘Samoan’ for Sāmoan and ‘Tubuai’ for Tupua‘i. For other Polynesian languages I follow common but inconsistent conventions: ‘Tongan’, not ‘Tonga’, but ‘Tikopia’, not ‘Tikopian’, etc.
1 Introduction

1.1 Jones and Klar’s Proposal

Jones and Klar’s proposal was presented in several papers, which I will refer to by single letter abbreviations. The proposal was first presented in Jones and Klar (2005), hereafter A, and, its linguistic arguments were elaborated in Klar and Jones (2005), hereafter B. Anderson (2006) presented a critique of A, which was followed by a rejoinder in Jones and Klar (2006), hereafter C. Arnold (2007) is a detailed critical review of A, with a reply by Klar and Jones (2008), hereafter D, another version of which appeared as Jones and Klar (2009), hereafter E. I summarize Jones and Klar’s arguments as follows:

- Planked canoe construction was practiced prehistorically in the Americas only by the Chumash and Gabrielino of Southern California and by the Mapuche of Chile, yet widespread among the Polynesians, who are known from other evidence to have reached the Americas.
- The Chumash and Gabrielino planked canoe appears in the archaeological record at about the time Polynesians first reached East Polynesia and the Americas, or soon afterwards.
- A certain style of fishhook, the curved-barb compound fishhook, is of a Polynesian form, and appears in the Chumash archaeological record at about the same time as the planked canoe.
- Several Native American words describing plank canoes have no apparent internal etymologies but can be derived from relevant Polynesian vocabulary:
  - The Chumash word for the planked canoe, reconstructed to the earlier form *tomolo or *tomoloʔo, can be derived from a Polynesian word, *tumuraʔau, meaning something like ‘useful wood’, and referring to the material from which the canoe is built.
  - The Gabrielino word for the sewn-plank boat, tiʔat, can be derived from a Polynesian word, *tia, ‘to sew’.
  - Another Gabrielino word for ‘boat’, taraina, can be derived from a Polynesian word,

3 At the time of this writing, most of Jones and Klar’s papers are available online at Terry Jones’s website, http://cla.calpoly.edu/~tljones/. In this paper I attribute linguistic arguments to ‘Klar and Jones’ and archaeological ones to ‘Jones and Klar’. Summaries of their arguments are also published in Jones (2010), Klar (2010), throughout Jones et al. (2011), and in Jones and Klar (2012).

4 Jones and Klar’s papers concentrate on contact between Polynesia and Southern California, but they also suggest contacts between Polynesia and the southern Chilean coast (A:461; D:93-94; E:179-180; Klar 2010; Klar 2011; Jones and Klar 2012). That topic is discussed more marginally; the linguistic part of it is brought up in paper D, but not in its revised version E. I will therefore not go into it in detail here. The linguistic argument suffers from similar weaknesses to the ones discussed here.

5 The single liquid consonant phoneme of Proto East Polynesian is sometimes marked as <R>, with an undetermined phonetic value reconstructed as either [r] or [l]. In the attested East Polynesian languages, [l] occurs only in Hawaiian, and early records show that [r] existed in Hawaiian as an allophone or a dialectal variant of that phoneme. I therefore reconstruct the PEP phoneme as [r]; this makes for clearer reading as well. Which liquid is reconstructed is not significant for this study, as has been noted by Klar and Jones (B:386), since all Chumashan languages have only one liquid, /l/, and either a Polynesian /r/ or an /l/ would be borrowed into a Chumashan language as /l/. The same argument applies to Gabrielino, which only has one liquid, /r/.
The Plank Canoe of Southern California

*taraina*, analyzed as *tarai*, ‘to hew, to carve’ + *-na* ‘nominalizer’, i.e., ‘carved object’.

- The above are explained by a scenario in which Polynesian voyagers have reached the American Pacific coast, and passed to the Native American populations the technology of planked boat construction along with related vocabulary, as well as particular fishhook styles.

I find that each of the arguments above is either flawed or entirely unsupported by the evidence, and that individually or together, they do not demonstrate Polynesian-American contact. I argue that the available linguistic, ethnographic and archaeological data point to a local origin for the planked boat of Southern California.

1.2 The Sweet Potato in Polynesia as an Example

That Polynesians have reached South America is established with certainty through the evidence of the sweet potato (*Ipomoea batatas*) and its native name (Yen 1974). The sweet potato, a South American cultivar, was present as a staple food crop throughout East Polynesia at the time of European contact, and archaeological evidence has indicated its presence centuries earlier (Hather and Kirch 1991; Higham and Gumbley 2001). The reconstructed Proto East Polynesian name of the sweet potato, *kumara*, is accepted as a borrowing of the form *kumar*, recorded in some dialects of Quechua, and more recently traced to the extinct Cañari language of the Ecuadorian coast (Scaglion 2005; Scaglion and Cordero 2011).

The certainty given to the evidence of the *kumara*, even without any other evidence of trans-Pacific contact, rests on two factors. The first is uniqueness: there is no possibility that a species could have independently arisen in two different places, and the sweet potato, a cultivated plant, would need to be purposefully transported and planted to get from one place to another. Secondly, the linguistic argument is straightforward. The meanings of the South American *kumar* and of the Polynesian *kumara* are identical. The only formal change in the word is the addition of the final -a to the Polynesian form, where closed syllables are prohibited, a process ubiquitous in borrowings into Polynesian languages. The length of the word argues against chance similarity.

Taking the case of the sweet potato as a standard for establishing such prehistorical contacts, I examine the evidence given by Jones and Klar. Here the material evidence of boat construction and fishhooks does not meet the standard of uniqueness, in that the technologies were innovated independently elsewhere. The linguistic evidence given by Jones and Klar requires several unattested or unlikely formal and semantic changes, and so opens more questions than it answers. And finally, the material and linguistic evidence can all be better explained through a scenario of local development within California.

1.3 Plan of the Paper

I will first examine the claim for the uniqueness of the plank canoe and show that planked boat construction is more widespread in the Americas and elsewhere than Jones and Klar suggest, and will argue for independent innovation as the preferred explanation for the appearance of planked boat construction. I find that each of the arguments above is either flawed or entirely unsupported by the evidence, and that individually or together, they do not demonstrate Polynesian-American contact. I argue that the available linguistic, ethnographic and archaeological data point to a local origin for the planked boat of Southern California.

---

6 Rensch (1991b) proposes that the Hawaiian form of the word, ʻuala, may indicate a separate introduction of the sweet potato to Polynesia from a source further north on the South American coast.
canoe construction in Polynesia and the Americas. 

Next, I will compare the techniques of planked boat construction by the Chumash and in East Polynesia. I will demonstrate that there is no clear evidence to link the two, and that differences in technique favor separate origins.

Next, I will examine issues of chronology, first of the appearance of the plank canoe in Southern California, then of Polynesian settlement in the eastern Pacific; I argue that the California plank canoe has either predated the Polynesian entry the Pacific, or else occurred very soon thereafter. This point is important in evaluating the linguistic arguments to follow.

I will then examine briefly the evidence for relating the Polynesian two-piece fishhook to the Chumash one, and will show that the two are unlikely to be related, based on chronological and stylistic arguments.

Moving to the linguistic evidence, I will first review the relevant issues in Polynesian historical phonology. I will then examine each of the American forms, and show that their claimed Polynesian sources are unlikely as such, on grounds of phonology, semantics, or both.

Finally, I will offer alternative etymologies for the Gabrieleno and Chumash forms, and discuss other scenarios relating to the appearance of the planked boat in Southern California.

2 Technologies of Boat Construction in California

Broadly, three types of boat were built and used in California before European contact: tule bundle boats, wooden dugout canoes, and planked boats. Their distribution has been reviewed elsewhere (Cunningham 1989; Heizer and Whipple 1951:11-14; Heizer and Massey 1953). I will mention a few salient points and add more details toward the end of this paper.

The planked canoe, around which this discussion turns, has been described in great detail by Hudson et al. (1978), based on all known sources, but especially the notes of John P. Harrington. This boat, most commonly known by the Barbareño Chumash name *tomol*, was constructed of planks, which were usually split from logs of redwood which had drifted south from this tree’s range in northern California. The edges of the planks were carefully glued together by a heated mixture of asphaltum and tree pitch, and then lashed tight by cords passed through holes drilled along the edges of the planks. The *tomol* was a large and seaworthy vessel, capable of reliably transporting people and goods between the mainland and the Channel Islands. These boats were used by the southern Chumash, on the coast facing the Channel Islands, by the Gabrieleno further south along the coast, and by the Channel Islanders themselves. No similarly constructed boat appears elsewhere in California.

Tule boats (balsas) were widespread in California, and their distribution roughly complements that of dugout canoes (Kroeber 1922:267-269). They were constructed of several bundles of reeds (common tule, *Scirpus acutus*), each bound tight; one bundle would serve as a keel, and one or more bundles would serve to build up the boat on either side. Tule boats are relatively easy to construct in a short time, and the materials for their construction were easily available in the wetlands near where the boats would eventually be used, including the former Buena Vista Lake and Tulare Lake in the San Joaquin Valley, rivers throughout the Central Valley, and estuaries on the southern and central California coasts.

Tule boats were a significant form of water transport in coastal Southern California, though their significance has been overshadowed in the literature by the more elaborate and better-attested plank canoe. They were utilized for ocean travel by the Luiseño, Gabrieleno, Chumash and Salinan people, and on beyond to the north and south. Tule boats were seaworthy enough to travel between the coast and the Channel Islands. The Chumash, and perhaps others,
sealed tule boats with asphaltum to waterproof them, which increased the time they could spend in the water before needing to be taken out and dried.

Dugout canoes are even more sparsely documented in the region. The Luiseño and Channel Chumash built dugout canoes, and likely the northernmost Chumash as well. Dugout canoes were historically used mostly for near-coast travel and fishing, but the Luiseño are said to have used them in earlier times for crossing to the southern Channel Islands. In historical times, at least, dugout canoes in Southern California were never as large or as ubiquitous as those of far northern California or the Pacific Northwest.

The Channel Islands were occupied by humans since the early Holocene (e.g., Rick et al. 2005). Since the archaeological evidence for plank canoes does not reach back more than one or two millennia, tule boats, dugouts or both must have been the predominant mode of oceanic transportation in Southern California for the past 10,000 years or so.7

3 Sewn Boat Technology: Worldwide Distribution

Linguistic issues aside, the argument for an external origin of the Chumash sewn plank canoe depends on the claim for its uniqueness in the Americas, or at least its rarity. Jones and Klar (A:461) state that the California plank canoe is the only example of planked canoe construction in the Americas, except possibly the dalca of southern Chile, and that the Chumash tomol is similar in details to Polynesian canoes in details of its construction. In this section, I show that sewn plank canoes are distributed worldwide, indicating multiple independent inventions of the technique. In particular, dugouts with sewn-on strakes, using a similar technique to fully planked construction, were used elsewhere in the Americas. Finally, despite Jones and Klar’s claims, the Chumash tomol was significantly different in its construction from East Polynesian sewn plank canoes.

Sewn-plank canoes in the narrow sense — canoes built entirely of planks sewn together — existed in Ancient Egypt (the ship of Cheops, 2600 BC, McGrail 2004) and elsewhere in the Mediterranean, Western Europe and Northern Europe (the Ferriby boat, 1900 BC, McGrail 2004) well into northern Russia (Litwin 1985), in inland western Africa (Insoll 1993), the Indian Ocean, China and Japan (McGrail 2004), and possibly southern Brazil (M. Brindley 1924:129, following Bates 1873:36). The dalca of southern Chile (Cooper 1917:198-200; Latcham 1930; Finsterbusch 1934; Heizer 1941b; Edwards 1965:21-34, Medina 1984; Puente 1986) was constructed of three planks, one serving as keel and the others serving as sides, and apparently originated with the Huilliche of Chiloé Island (Lothrop 1932). Another type of three-planked sewn boat is the xodol or eksil’ of the Yukaghir of the Kolyma River, near the Arctic coast of eastern Siberia (Jochelson 1926:375-378; Mudge 1880:290), almost antipodally from the dalca.8

7 Fagan (2004) has suggested that plank canoes were present in Southern California for much of the Holocene. He does so by dismissing tule balsas as a viable means of transportation to the islands, which I consider unjustified, for the reasons discussed by Des Lauriers (2005). Cassidy et al. (2004:125-126) argue that Middle Holocene tool assemblages in the area are strikingly similar to those used for planked boat construction. Their argument is better, but still circumstantial and not conclusive, and is inconsistent with other evidence showing later appearance of the tomol (Arnold 2007:202).

8 Rousselot (1994:244-245) mistakenly states that the Yukaghir board canoe was an adaptation of Russian boats. In this he must refer to the qarbas, a sewn-plank canoe with a clearly European design and a borrowed name, rather than to the older 3-board design.
The technique of plank sewing was used to a less complete degree in built-up boats, that is, boats consisting of a dugout base or a keel with one or more rows of planks attached to its sides, so as to increase the boat’s freeboard. Such boats were used by the Ainu of northern Japan (Ohtsuka 1999) and elsewhere in eastern Siberia (H. H. Brindley 1919-1920, II:104, III:139-140); and, in the Americas, in the Pacific Northwest (Kwakiutl, Boas 1909:334-337, 446; Haida, Stewart 1984:50 and Durham 1960:57, see also Howay 1941:207-208; Tlingit, Emmons 1991:84, 91), where the same technique was also used to repair cracks (Stewart 1984:45-47; for Coast Salish, Lincoln 1991:30); and in the Caribbean (McKusick 1960:5-7) and the Orinoco basin (Roth 1924:612-614, after Gumilla 1791, 2:113-116). In the Old World they are recorded from West Africa (Durand 1806:111), Russia (Litwin 1985) and elsewhere.

Both fully planked boats and dugouts with raised sides existed throughout Oceania, and in East Polynesia in particular (Best 1925; Haddon and Hornell 1936; Bataille-Benguigui et al. 2008). Where both forms existed, the choice of boat form depended on balancing the additional labor involved in building plank boats with the necessity of obtaining large logs for dugout bases (Haddon and Hornell 1936:345; Kamakau 1976:118).

The technique of sewing flat pieces of wood together into boats was also used in the construction of bark canoes, which employs bark peeled from trees, often as thick as planks split from a log. Sewn bark canoes were used at least in East Africa, Australia, Borneo, the Solomon Islands, northeastern North America, the Orinoco and Amazon basins, and Tierra del Fuego (Vairo 2002:97-125).

Lashing planks to each other side by side requires perfecting several techniques: truing the edges of the planks for a close fit; drilling holes; sealing the joints by calking them; and establishing a series of tight lashings which will not loosen or fall apart even after absorbing water. These techniques need to be established and relied on whether one attaches a single row of strakes to a dugout, builds a canoe of tree bark, or builds a fully planked canoe. Among American boat types, the Chumash plank canoe is hence much closer technologically to the built-up boats of the Pacific Northwest and the Caribbean than any of those are to a simple dugout, and the argument for the uniqueness of the tomol in the Americas is therefore weaker.

Anderson (2006:759-760), in his comment on Jones and Klar’s original paper, mentions the wide use of sewn plank boats elsewhere in the world. Jones and Klar did not address this issue in their reply (C). Anderson does not describe just how widespread sewn boats are, and his argument veers toward advocating a different external source for the Chumash canoe. My

---

9 Jones and Klar, referring to the addition of strakes to Pacific Northwest canoes, comment that “When strakes or gunwales were added to the sides of these craft to increase freeboard, they were generally attached by mortising, not by sewing” (A:461). This is clearly not true in general, as seen in Boas’s account of Kwakiutl techniques and in the other references mentioned here. On the other hand, Jones and Klar’s quote comments by the eighteenth century observers, Crespi and Peña, about Northwest canoes made of ‘several pieces’. These may well have referred to separate bow and stern pieces, not to raised sides.

10 Anderson (2006:760) distances the Chumash boat from Polynesian designs by drawing a distinction between fully sewn-plank boats (like the tomol) and built up boats with strakes sewn to a dugout base. As I argue here, that is a minor distinction, since the technique is mostly the same for both. In any event, early Polynesian voyaging canoes may well have been fully planked boats (e.g., Kamakau 1976:118), like those built in the Tuamotus through historical times (Haddon and Hornell 1936:67, 131 and elsewhere; Bataille-Benguigui et al. 2008). Polynesia, like almost every culture with seagoing tradition, had a great variety of boat types, specialized for different purposes and different effort of production.
argument is different: the more evidence there is for an innovation occurring independently many times, the stronger the argument is for the innovation occurring yet once more, and the lesser the need for external introduction as an explanation.

4 Sewn Boat Technology: Comparing the Details of Construction

Jones and Klar claim that “...tools, and techniques used in the construction of Polynesian sewn-plank boats are remarkably similar to those associated with the Chumashan tomolo,” and enumerate what they consider parallels between the details of Chumash and East Polynesian canoe styles (A:465-466). I will examine these criteria, followed by the traits Prins (1986) uses in his typology of sewn plank canoes, and finally discuss several additional distinguishing technological traits.

Jones and Klar’s comparative traits are:

• Adze form. “...hand-held adzes of nearly identical design (a short handle to which was lashed a shell blade) used as the primary tools to work planks. In the Tuamotu group, adzes were commonly made with clam shells as they were among the Chumash.” This is not a significant trait. Short handles are universally necessary on carpenters’ tools used for smaller work, and vice versa; this has no special connection with boat construction. The ‘elbow adze’, with a bent-down handle, was used by the Chumash for shaping wood; a similar form is widely distributed in northwestern North America (Olson 1927:7). Both stone and shell adzes were used in Polynesia and by the Chumash (Kamakau 1976:122; Hudson and Blackburn 1987:52), showing that both people, reasonably, used all the materials available to them as they found them suitable; this provides no evidence at all for cultural transfer. The details of adze form provide a valuable archaeological tool, and have been studied closely in Polynesia and elsewhere, but Jones and Klar give no details for comparing adze forms in these two areas, nor compare them to adze forms elsewhere; their claim of a ‘nearly identical design’ is unsupported.

• Sandpaper. “Wood was finished with sandpaper — in Polynesia derived from a plant source, not the Chumashan sharkskin.” A tight seal between joined planks depends on a precise fit. The final shaping of the joined surfaces was achieved in Polynesia by fine adzing (Haddon and Hornell 1936:135). Though sharkskin and coral rocks were known in Polynesia as sanding materials, I am not aware of any account of the use of sanding for shaping the matching edges of boat planks.

In Chumash technique, boat planks were first assembled and glued in place with pitch, before the final sewing. That required a particularly tight fit and smooth joint surface, which was achieved by polishing with sharkskin (Hudson et al. 1978:73, 75). The Gabrielino also sanded the outer surfaces of planks by weighting them down and dragging them on wet sand (Alliott 1917:42-43). Smoothing the outer surface of a canoe with sharkskin was also practiced in the Pacific Northwest (Stewart 1984:54). In other words, the use of sharkskin as sanding material is neither exclusive to Polynesia and the Chumash coast, nor is it universal in these areas. Sanding is a general woodworking technique, not especially linked with boat construction. There is nothing in sanding technique to connect Chumash and Polynesian boat-building.

• Caulking tools. “As among the Chumash, caulking in Polynesia was done with wooden caulking tools, although those of Hawaii were of more complex design.” The Chumash caulking
tool was a wooden stick whittled at its end to a sharp edge, with which to force caulking material into the gap between planks. Similar tools were also made of bone (Hudson et al. 1978:41-42, and n. 56). There is nothing remarkable or unique about using wooden tools for such a purpose: the mere use of wood here does not imply a cultural connection.

• Canoe sheds. “Plank canoe construction in much of Polynesia was undertaken within a specially constructed canoe shed that protected the craft from the elements during its construction. This is very similar to the structure of mats and poles used by Chumash canoe builders for the same purpose.” To begin with, building shelters against the sun and the weather is a common activity in all human cultures. Their use in boat construction is not remarkable. More specifically, as Arnold (2007:203) has noted, Polynesian canoe sheds were large, permanent structures meant to completely enclose the boat under construction and protect it from the rain and the sun (at places evolving to the size of hangars, Haddon and Hornell 1936:328). The Chumash built small temporary frames of three poles and leaned a mat against them to protect the canoe from the sun, while the pitch used to glue its planks together was hardening (Hudson et al. 1978:44). The Chumash boat hut, as described, matches larger structures elsewhere in the area (e.g., Wallace 1978a:451) and indicates no external character. The purpose and form of the Polynesian and Chumash shelters were entirely different from each other. Jones and Klar (D, E) do not address this point as raised by Arnold.

In sum, none of the traits mentioned by Jones and Klar offer any support for Polynesian-Chumash contact. The traits they enumerate are either widespread, or are in fact not comparable.

Prins (1986) is an extensive comparative survey of sewn plank boat construction techniques worldwide. Although not quite complete in its coverage and details, it is the only work of its kind and scale. One of Prins’s aims was to select a small number of binary typological traits by which sewn planked boat traditions may be broadly distinguished, and use them to show the geographical distribution of different techniques. I note that Prins’s study focuses on highlighting world-scale patterns, and his traits are not always optimal for distinguishing boat building traditions within smaller areas. His four basic traits are: the presence (or absence) of continuous sewing; the presence of hole plugs; the presence of aligning dowels; and edge-to-edge versus overlapping plank construction. Their significance here is as follows:

• Continuous sewing. Continuous sewing is the practice of lacing a running cord back and forth through many pairs of drilled holes in adjacent planks. In discontinuous sewing, one short cord is passed through each pair of holes, tightened to pull the planks together, and tied off. In this regard East Polynesian canoes are clearly different from Chumash ones. The tomol was lashed with individual short cords, one for each pair of holes (Hudson et al. 1978:83-85), while continuous sewing was nearly universal throughout East Polynesia (Haddon and Hornell 1936 passim). Discontinuous sewing was used, however, elsewhere in Oceania, from Samoa westward, suggesting that continuous sewing was an East Polynesian innovation. Both continuously and discontinuously sewn boats occur in many parts of the world (Prins 1986:168).

• Plugs. In some plank sewing, a peg or plug is jammed into the hole after the cord was passed through it, in order to maintain the tension in the cord and provide additional sealing. This technique is not used in the Chumash canoe (Hudson et al. 1978:83-85). It was sometimes used in East Polynesia (Haddon and Hornell 1936:142), but not universally, and may be a later
innovation. It may have been developed in East Polynesia specifically for better tensioning of continuously sewn cords.

• **Dowels.** In some sewn plank boats, blind holes are drilled into the edges of the planks where they meet, and dowels are inserted, so as to align the planks and keep them from shifting past each other. Dowels were not used in East Polynesia, except perhaps in the largest Tuamotuan sewn plank canoes (Haddon and Hornell 1936:80). This technique may have been used in the construction of the Chumash *tomol* (Hudson et al. 1978:95), only in attaching the uppermost round of boards, and not always even then; the Chumash used dowels for other purposes, likely under European influence (Hudson et al. 1978:92-93). Horridge (1986:57-58, quoted in Pawley and Pawley 1998) notes that in western Oceania dowels were a later development, which generally followed the introduction of metal tools.

• **Plank positioning.** In overlapping (‘clinker’) construction the planks partly overlap each other, as they are joined face to face. Otherwise they are joined edge to edge. Polynesian and Chumash canoes are both edge-joined. This is not a diagnostic feature, since edge to edge construction is common worldwide, except mainly in Northern Europe and the Solomon Sea (Prins 1986:168).

Of Prins’s four traits, The use of non-continuous sewing in the Chumash canoe weakly argues against a Polynesian connection. However, continuous sewing could conceivably be a late innovation which spread through East Polynesia after the time in question; in that case, this trait is not diagnostic. The other three of Prins’s traits are not relevant here. In total, Prins’s traits do not offer evidence in favor of Polynesian-Chumash contact.

Other distinctive traits not discussed by Jones and Klar or by Prins include:

• **Battens under cords.** In East Polynesian canoes, a long batten — a flat strip of material — was placed so as to cover the seams between the planks, and the cords would pass over the batten and hold it tightly in place. This provided further sealing against leaks, kept the caulking material within the joint, and helped to keep the lashing taut. This technique was used at least in the Marquesas (Handy 1923:157-158), Tuamotus (Haddon and Hornell 1936:58, 68, 69, 71, 89), Societies (Nordhoff 1930:145), the Northern Cooks (Haddon and Hornell 1936:178) and New Zealand (Best 1925:77; Haddon and Hornell 1936:202). Though widespread, the technique might not have been universal in the area. Battens were not used in Chumash boats.

• **Recessed groove (countersink) for cord.** In the Chumash canoe, grooves were carved into the planks between the holes, in which the cord could pass without projecting above the surface. This kept the cords from being abraded, and on the inside it prevented the rough cords from chafing against the skin of the crew (Hudson et al. 1978:82). Clearly, countersunk cords cannot be wrapped over battens as described above, and in fact countersinking was not usually used in East Polynesia (but see Best 1925:72 for countersinking when lashing together hull sections in Maori canoes).

• **Bent planks.** In Chumash boat-building technique planks were first cut and formed, then bent using heat and moisture (Hudson et al. 1978:68-72). This technique was never used in East Polynesia, where planks were shaped entirely by splitting and carving (Handy 1923:157; Henry 1928:549; Fornander 1917, 5:612).
• **Frame.** The Chumash canoe is built up of planks held to each other, without a supporting frame of ribs and with a single thwart, or cross-brace (Hudson et al. 1978:92); this is an unusual and distinct form of sewn plank construction. Of the fully planked canoes of the Tuamotus, the largest had frames (Haddon and Hornell 1936:80, 83). These very large boats presumably were close in design to the voyaging canoes which had voyaged to the American coast. Smaller sewn-plank boats of East Polynesia, closer in size to the *tomol,* did not use ribs or a frame, though they utilized thwarts.

• **Caulking.** In Polynesia, caulking — sealing the gaps between the planks — was done by placing between the planks fibrous matter, typically coconut fibers with breadfruit juice, which would then be compressed as the planks were lashed together. The Chumash, on the other hand, used *yop,* heated asphaltum diluted with tree sap, which would fill the gap between the planks and then harden. Secondary caulking was of tule was added along the of the joints and sealed with more *yop.* Gamble (2002:307) has found traces of asphaltum on the edge of canoe planks from every archaeological context she studied, including one dated to the late first millennium AD. The Chumash use of *yop* goes beyond mere caulking, in that it has a significant structural function. Each round of planks of the Chumash boat was assembled by gluing the planks to the lower round using *yop,* and sewing them together only after it has hardened. This technique no doubt owes its origin to the availability of asphaltum in Chumash territory, and is possibly unparalleled anywhere else; certainly it is quite different from Polynesian technique.

The above five traits all represent techniques, some clearly beneficial, which are not clearly shared between East Polynesia and Southern California. Several of these characteristics can be noticed in a few minutes inspection of a finished boat, and could have easily been transferred to the Chumash even through brief and casual interaction. By the simplest interpretation, the difference in technologies argues against a Polynesian origin for the Chumash canoe. Of course, the Chumash and Polynesian canoes of AD 1800 are no doubt different than those of, say, AD 1000, and some of the technologies discussed here may be later developments. Even so, there are no distinctive traits shared by the two areas. At best, the evidence of boat-building techniques provides no proof of the Polynesian-Chumash contact hypothesis. At worst, the evidence disproves it.\(^\text{11}\)

Beyond the specifics of planked canoe construction, Anderson (2006:760) and Arnold (2007:203) have already pointed out other characteristics of Oceanic boat construction absent from the Chumash canoe, namely outriggers or double hulls, which add stability, and sails. Jones and Klar (C:766) deflect this argument by saying that the Chumash, for whatever reason, have chosen not to adopt these elements. Elsewhere (A:469) they suggest that perhaps the lack of suitable sail material kept the Chumash from adopting that technology, and the double hulled boats were too complex to copy. This is a weak argument, absent any convincing reasons why the Chumash would ignore these elements while adopting other complex technologies. As above, this argument at best trades counterevidence for lack of evidence.

---

\(^{11}\) Robinson (1943:17) also sees no connection between Chumash and Oceanic boatbuilding techniques, but provides no details. Heizer (1940:83-88) examines in detail several details of construction in the Chumash and other plank boat types of construction. He concludes that the Chumash boat was an independent innovation, which he believes evolved from the design of the tule boat.
At best, the inferred dates for the appearance of the Chumash canoe and the initial settlement of East Polynesia are uncomfortably close. The very closeness of the dates would fit nicely in a scenario of rapid settlement of the eastern Pacific, culminating with American contact soon thereafter. However, if the Chumash canoe turns out to have appeared even slightly before humans had reached East Polynesia, this would clearly rule out Polynesian contact as its source. Therefore, the knowns and the uncertainties in both dates have to be well understood. Additionally, if the Chumash canoe was indeed developed after the settlement of East Polynesia, the chronology turns out to narrow down the details of the language of the Polynesian populations which could have landed in California. This aids the analysis of the proposed Polynesian etymologies.

5.1 The Chronology of Settlement in East Polynesia

Polynesians were the first people to inhabit the islands of East Polynesia, also referred to as triangle Polynesia — most of the area encompassed by Easter Island, Hawaii, and New Zealand — and they reached them from the west, in the final stage of the Austronesian expansion. The dating of the initial settlement of East Polynesia, or of any of its islands, is an active field of research. Jones and Klar (A:461, 477; C:767-768; D:92-93; E:179), and the critiques of Anderson (2006:760), and Arnold (2007:202-203) touch on these issues; I review here the history of East Polynesian settlement chronology and its current status in greater detail.

Radiocarbon dating has been utilized in Polynesian archaeology since soon after its invention in the late 1950s, and remains the tool of choice for obtaining absolute dates. Early on, a sequence of settlement dates emerged for the major East Polynesian island groups, and was used together with archaeological and linguistic evidence to evolve what Kirch (1986), in a detailed review, called the “orthodox scenario”. In that scenario, East Polynesian settlement started with the settlement of the Marquesas from West Polynesia around AD 300, progressing to the rest of East Polynesia in the following few centuries, up to the settlement of New Zealand in AD 800-1000 or earlier. Some variations on the model called for even earlier East Polynesian settlement dates, as early as the first millennium BC. This model, with some modifications, was the predominant one from the 1960s to the 1990s, and still occasionally appears in the literature.12

Early models of East Polynesian settlement chronology were anchored by relatively few radiocarbon dates of the first millennium AD, and a few even earlier ones, in East Polynesia as a whole as well as in individual island groups, in contrast with much more abundant post-AD 1000 dates. This paucity of older samples was usually taken to show small initial populations growing slowly, and so producing fewer datable artifacts, which with greater age would also be less likely to have survived.

A decisive turn in Polynesian chronology came with Atholl Anderson’s work, beginning with Anderson (1991). In it he examined the entire corpus of radiocarbon dates existing for New Zealand, then thought to have been settled sometime in the first millennium AD. Anderson applied what is known as chronometric hygiene, systematically rejecting samples based on a set

12 Jones and Klar refer to “the era of greatest Polynesian exploration (ca. A.D. 500-1000)” (A:461) and “the era when Polynesian seafarers discovered the most distant outposts of the Pacific (A.D. 500-1100)” (A:477). They quote no source for these dates.
of internal criteria (such as materials prone to producing erroneous dates) and external ones (such as samples with aberrant ages among others from the same context). Without these questionable samples a large set of dates still remained, but showing no dates earlier than the twelfth century AD, and abundant later dates. This suggested a model significantly different than the ‘orthodox scenario’; here New Zealand was not settled until the twelfth century, and its population grew rapidly after settlement. Conversely, this study has shown that earlier New Zealand dates are all likely the result of technical errors. Subsequent studies have shown strong evidence for an even later settlement date, in the late thirteenth century (Hogg et al. 2003, Wilmshurst and Higham 2004, Wilmshurst et al. 2008), further weakening the remaining arguments for early settlement (Sutton et al. 2008; see also Butler 2008 and Matisoo-Smith et al. 2008).

The successful revision of New Zealand settlement chronology was extended by the same principles beyond New Zealand as well. Spriggs and Anderson (1993) applied the procedures of chronometric hygiene to the rest of East Polynesia, resulting in a similar rejection of many early dates, and tentatively estimated the settlement of East Polynesia at AD 600-950, based on several dates with large uncertainties. Since then, abundant additional work has consistently reinforced later chronologies throughout East Polynesia. Moreover, new techniques of sample selection, preparation, measurement and correction, not available in the first wave of East Polynesian chronometry (Spriggs 2010), have failed to turn up early first millennium dates, and several claimed early samples and sites were shown to be of a younger age (Kirch and Kahn 2007:198-201). Even some supposedly early sites which passed the criteria of Spriggs and Anderson (1993) turned out on fresh reanalysis to be much more recent (Anderson and Sinoto 2002; Kirch and Kahn 2007:199; Dye and Pantaleo 2010). Additional support for these shorter chronologies comes from recent geological studies (Pirazzoli and Montaggioni 1988, Dickinson 2003, 2009), which show that some islands were still submerged or otherwise not habitable during the times suggested by some earlier chronologies. A review of the literature by Kirch and Kahn (2007:201) puts initial East Polynesian settlement at no earlier than AD 800. Weisler and Green (2011) place it at ca. AD 800. Finally, Wilmshurst et al. (2011), following the chronological hygiene methodology of Spriggs and Anderson but employing more accurate and far more abundant data, places the earliest settlement of East Polynesia at about AD 1000.

While unknowns and controversies remain, the extent of the debate has shrunk. Initial settlement of East Polynesia not much earlier than AD 1000 is now generally accepted, and the controversies cover a range of a few centuries rather than a millennium. Data from some island groups are sparse, and chronometric data cannot yet be used to resolve the order in which they were settled, with a few exceptions. In addition, the interpretation of environmental proxies for early settlement, particularly pollen and charcoal records from wetland deposits, remain hard to interpret and less decisive. While earlier chronologies in the region continue to be argued (Sutton et al. 2008; Kirch and Ellison 1994, see also Anderson 1994), they are at this point on the decline. To call the shorter chronologies ‘controversial’, as Jones and Klar do (D:92, E:179) is an exaggeration, as is their claim that “there is no consensus on the proposed short chronology for

---

13 The term ‘chronometric hygiene’ is due to Wilfred Shawcross, and was popularized by Spriggs (1989).

14 Polach (1976) is an older but very useful survey of radiocarbon measurement and its pitfalls.

15 In one recent paper, the authors disagree among themselves whether a settlement date of ca. AD 800 or ca. AD 1000 is a better fit to the chronometric data for one island group (Anderson et al. 2003:137). This is fairly representative of the current range of opinions for the initial settlement date of East Polynesia.
eastern Polynesia” (C:768).  

In Hawaii, a departure point for California suggested by Jones and Klar, no clear evidence exists for human presence before AD 1000. Three recent regional syntheses — Athens at al. (2002) for the ‘Ewa plain in southern O‘ahu, Carson (2006) for Kaua‘i, and McCoy (2007) and Kirch and McCoy (2007) for Moloka‘i — independently find no certain dated evidence for human presence in their particular areas before ca. AD 1000, and Wilmshurst et al. (2011) and Rieth et al. (2011) place Hawaiian settlement at no earlier than AD 1200. To support an early date for Hawaiian colonization, Jones and Klar (C:768, D:92-93, E:179) quote a single datum from the compilation of Spriggs and Anderson (1993:202). That date, Beta-30860, is from a sample from the Honokaua site in Maui. Its calibrated date interval is AD 650-770 (1σ, uncalibrated date 1330±60 BP). There are no other accepted samples in Spriggs and Anderson’s compilation whose entire 2σ calibrated range falls earlier than AD 1000, or whose entire 1σ range falls earlier than AD 800; in other words, none of them decisively indicates a pre-AD 800 presence. Jones and Klar’s claim here rests heavily on this single sample. This sample comes from charcoal of an unidentified plant source (Theresa Donham, p.c. 2010), and could therefore be of a wood significantly older than its time of deposition. Jones and Klar’s argument, already chronologically precarious, can not safely depend on a revision of Hawaiian chronology based on this single questionable date.

In their later papers Jones and Klar suggest Central East Polynesia as another possible departure point for California, apparently to better accommodate the linguistic evidence. Central East Polynesia must have been reached before Hawaii, but apparently not by much. There is no direct unquestioned evidence for human presence anywhere in East Polynesia before ca. AD 800 in earlier scenarios (Kirch and Kahn 2007) or AD 1000 in the more recent ones (Wilmshurst et al. 2011). The overall picture is of rapid settlement of East Polynesia beginning not much earlier than AD 1000, over a period of a few centuries. This is comparable to the span of time over which the first European explorers reached all the Pacific archipelagos, though of course the circumstances of discovery were different in the two cases.

In sum, the emerging consensus, stated twenty years ago and bolstered and refined since by additional data, is that Central East Polynesia and Hawaii have been reached and settled not much earlier than AD 1000. Polynesians could not have been the source of the Chumash plank canoe if it appeared earlier than that time.

---

16 Even shorter chronologies for Easter Island settlement were proposed by Hunt and Lipo (2006, 2008), on the basis of chronological hygiene, and mentioned by Jones and Klar as examples of short chronology controversies. Hunt and Lipo’s Easter Island settlement date (ca. AD 1200) is indeed even younger than other recent estimates (pre-AD 1000, Martinsson-Wallin and Crockford 2001; Weisler and Green 2011). However their arguments do not directly affect the chronology for Central East Polynesia or Hawaii, and are separate from them.

17 Dye (2011) argues for moving the Hawaiian settlement date back, perhaps to ca. AD 1000. His argument relies on including in the radiocarbon corpus two dates rejected by Wilmshurst et al. (2011), one from a Kukui nut and one from a rat bone. Rat bones are notoriously hard to date correctly (e.g., Anderson 2000, Wilmshurst et al. 2008), which leaves the revision dependent on a single date.

18 1σ ranges were calculated from the data of Spriggs and Anderson (1993) using CALIB 5.1 (Stuiver et al. 2006).

19 Kirch and Kahn’s early date, AD 780-1000 (cal 1σ) has been presented for Henderson Island (Weisler 1993:210 and Weisler 1994); however, this is based on unidentified charcoal and may represent old wood.
5.2 Dating the Appearance of the Chumash Plank Boat

The accepted dates for the appearance of Chumash plank canoe derive from two sources: artifacts, in the form of drilled planks, and the remains of fish thought to have been caught from plank canoes.

Gamble (2002) is the most recent work on the dating of plank boat-related physical artifacts. The older of her two dates for clearly identified canoe planks is AD 625-700 (cal 2σ; CA-SMI-261, Daisy Cave, San Miguel Island). However, as she points out, this plank is from driftwood, and is subject to the old wood effect: if redwood (Sequoia sempervirens) or other large tree species was the source of this plank, the driftwood log used for making the plank could have been hundreds of years old when the plank was made. This date is thus of little value for distinguishing a definitely pre-Polynesian, pre-AD 800 *tomol* from one definitely later than Polynesian settlement. The other canoe plank dates to no older than the fifteenth century AD.

Other canoe planks, as well as stone drills possible used for plank drilling, have also been identified in a burial from the cemetery of Simo’mo. This burial was assigned to as phase 4 (M4) of the Middle period of the cultural sequence of King (1981), or possibly as early as M3. M4 is bracketed chronologically by dates known to belong to the preceding phase M3 and the younger phase M5a (King 1981: 47, 59, 64), and the chronology is further reinforced by dates associated with California beads found in the desert Southwest, hundreds of kilometers inland. King places the boundaries of M4 around AD 700-900. However, one artifact assigned to M5a yielded a radiocarbon date of AD 690-860 cal 1σ (1246±60 BP; UCLA 1886). If this date is correct, it would shift the M4/M5a boundary to the early ninth century, and make it likely that the M4-era Simo’mo planks are older than AD 800. The invention or introduction of the *tomol* would be older still.

The advent of the Chumash plank canoe can also be estimated using the proxy of fish remains (Bernard 2001, 2004; Arnold and Bernard 2005). In particular, swordfish (*Xiphias gladius*), shortfin mako shark (*Isurus oxyrinchus*) and tuna (*Thunnus* sp.) remains are considered to be a strong indicator of *tomol* fishing, as any other boat known to have existed in the area would be too small to handle these strong, determined fish or to haul them back to shore (Bernard 2001:21-29). Bernard (2001) has assembled records of remains of what she considers ‘*tomol*-acquired species’ and correlated them with dated strata from a variety of known sites.

In Bernard’s ‘high resolution’ data set (Bernard 2001:65-70) there are four sites with reasonable stratigraphic control containing apparently pre-AD 900 remains of *tomol*-acquired species. SMI-481 (Otter Point, San Miguel Island) yielded a few swordfish vertebrae, from a context dated to AD 730-800 (range of medians of calibrated dates), based on the data of Rick

---

20 The cultural seriation of King (1981) is based on the analysis of bead ornaments in funerary contexts.

21 Both dates used for bracketing M4 are on human bone collagen samples. They are thus free from issues of old carbon. However, human bone collagen may appear too old by several decades if it came from people with a significantly marine-based diet (Polach 1976; Walker and DeNiro 1986). King (1981) provides no other details of sample preparation protocol. The oldest M5a sample, UCLA 1886, is reported as “a radiocarbon date on collagen from Burials 35 and 36 from LAn-264” (King 1981:64); this was apparently a commingled burial (Chester King, p.c. 2010). The dates are calibrated here using CALIB 5.1 (Stuiver et al. 2006), assuming the raw dates are corrected for fractionation appropriate for bone collagen. If not, a fractionation correction of 80±35 years (Polach 1976:268) would make UCLA 1886 older by several decades. King (1981:64) gives the calibrated date of UCLA 1886 as AD 730-790 and comments on the discrepancy with his chronology.
LAN-52 (Arroyo Sequit, west of Malibu) has some remains of tuna from a layer dated to AD 600-750 (1σ, 1340±100 BP), with more at lower strata ca. 100 years older, by extrapolation using an estimate for the deposition rate. This date was obtained in 1963, and details of sample selection and preparation are not given; its accuracy is therefore questionable. Another site, LAN-227 (Century Ranch) yielded tuna remains from strata dated to the seventh century. This date, again, is obtained by interpolation based on depth and on imprecise and possibly uncalibrated radiocarbon dates from the early 1960s. A fourth site, SBA-72N (Tecolote Canyon, near Santa Barbara), yielded a single tooth of a shortfin mako, with an associated date of ca. AD 500; I consider this find too inconclusive. Bernard (2001:105) synthesizes her findings into a significant increase in tomol-caught species beginning around the 8th-9th centuries AD; she reads the the gradual slow increase in the remains of such fish as indicating a period of refinement in boat-building technology.

One particularly spectacular swordfish remnant is the swordfish dancer’s mask described by Davenport et al. (1993). The mask is assembled from a swordfish skull, and attached to a cape of abalone shells. An ‘ornament’ (apparently of mother-of-pearl) from the cape was used to obtain a corrected radiocarbon date of 2040±90 BP, which yields a calibrated date of AD 480-680 1σ or AD 380-780 2σ (Stuiver et al. 2006, ΔR=230±35); I note that the shells used for the cape may have been old ones collected inland, and so the cape may be younger than the shells. Jones and Klar (C:767) correctly point out that this cape is younger than the uncorrected radiocarbon age would indicate.

In sum, the canoe plank remains of Simo’mo indicate the presence of plank canoes ca. AD 900 or before, with some uncertainty. The evidence of fish remains points at an earlier time for appearance of the tomol, but depends on more uncertain interpretation and less reliable dates. The current data for the Simo’mo cemetery and for the fish remains sites are consistent with the appearance of the tomol, by a very rough estimate, no later than AD 800 and probably a century or more earlier.

5.3 Chronology: Discussion

To summarize the chronological issues: there is substantial direct evidence for human presence in Central East Polynesia and Hawaii after AD 1000, and only circumstantial and uncertain evidence for such before AD 900. In Chumash country, there is some evidence for the plank canoe existing by AD 800 or even AD 700.

Both these date estimates suffer from uncertainties. Earlier dates may yet be found in East Polynesia, though earlier remains of Chumash plank boats may be found as well. Much of Chumash boat chronology is based on uncertain dates and overreaching assumptions. Better data from both areas may confirm and strengthen the current chronologies, or provide new chronologies with dates moved either forward or back. For now, the likelier conclusion is that the Chumash plank canoe predated the presence of Polynesians in the east Pacific and the Americas by a century or two, and that the two are therefore unrelated. A less likely, but currently still tenable position, is that Polynesians settled East Polynesia at about the time the

---

22 See also Erlandson et al. (2005). At the time Bernard (2001) was written, these data were still unpublished. Rick (2004:147-149) identifies the remains of what are apparently three Swordfish or Marlin vertebrae from his unit 1a/1b. Three calibrated dates were obtained from the unit (Rick 2004:134, also Erlandson et al. 2005), roughly AD 970±70, AD 795±85 and AD 740±50 (1σ corrected).
tomol first came to be. That would require finding compelling evidence for pushing the dates associated with the plank canoe a century or more toward the present, and East Polynesian settlement a century or more back, to converge on a date of, say, AD 800-850. Jones and Klar have attempted to push both these chronologies in that way, but have not conclusively succeeded (C:768; D:92-93; E:178-179).

For the rest of my discussion here, I will adopt this second, less probable, position: that East Polynesia had been first settled just before the plank canoe appeared in California. By implication, any Polynesians arriving then in Chumash territory would manifest the culture and language reconstructed for ancestral East Polynesia. This point is used below in evaluating the evidence of fishhooks and the linguistic evidence.

6 Archaeology: The Evidence of Fishhooks

To support their scenario, Jones and Klar compare the forms of a particular style of fishhook, the two-piece fishhook, from East Polynesia and Southern California. This fishhook is assembled of two parts: a barb, which hooks the fish, and a shaft, which connects to the fishing line (Hudson and Blackburn 1982:179-181; Kirch 1985:200-203). The two parts are tied together at the bottom with cordage, and in California are also glued with asphaltum. The fishhook barb, often made of bone or shell, may be well-preserved in archaeological contexts. Jones and Klar (A:466-468, C:766-767) do not claim the Chumash two-piece fishhook is a Polynesian import; they only associate a more curved, s-shaped form of the barb with similarly shaped fishhook barbs of Polynesia, which they believe were the model for the Chumash ones.

Artifacts which may be interpreted as two-piece fishhook barbs appear relatively early in the Chumash archaeological record, in King’s phases M2b and M3 (200 BC-AD 300 and AD 300-700 respectively, King 1981:47), and are uncontroversially a local development. The curved form in question is recorded from as early as phase M5 (AD 900-1150). No two-piece fishhook barbs are recorded from M4 contexts. These data, though fragmentary, are consistent with a transition to the curved form of fishhook some time during M4 or early M5. Jones and Klar associate it with the appearance of the tomol, at roughly that time.

However, the East Polynesian two-piece fishhook did not yet exist then. It is recorded from the margins of East Polynesia: Easter Island, New Zealand, and Hawaii; in Hawaii, the curved form mentioned by Jones and Klar appears only late in the chronological sequence (Kirch 1985:205-207; Emory et al. 1959:26). It does not appear at all in the archaeological record from Central East Polynesia (Sinoto 1979:125). It is believed that it was innovated later, and separately, in those three marginal locations. The most common ancestral East Polynesian fishhook was made of one piece, carved from a single round shell. Kirch and Green (1987:173) cite the two-piece fishhook as an example of convergent technological evolution: it was developed in Polynesia where strong shell material was not available for producing the older one-piece style of fishhook. Other examples of the independent invention of similar fishhooks are known from the Baltic region (Anell 1995:195, fig. 20), even showing the curved barb which Jones and Klar regard as a distinctive trait linking Polynesia and Southern California. In sum, chronology and geographical distribution argue against the Chumash fishhook originating in Polynesia, and the independent historical emergence of formally similar fishhooks within Polynesia and worldwide agrees with their independent development in California, with a similar functional motivation.

Notably, the one-piece circular shell fishhook also appears in very similar styles in Chumash country and in East Polynesia. Several have suggested this striking similarity as evidence for
The Plank Canoe of Southern California

trans-Pacific cultural contact (Rau 1884:138; Olson 1930:21; Kroeber 1939:44; and others). However, as Jones and Klar have already discussed (A:459, 466), the oldest Chumash circular fishhooks predate East Polynesian colonization by millennia, and are therefore unrelated. The methodological lesson here is to use caution in equating artifacts based on formal similarities, however striking. This is especially true with items whose form is mostly functional, such as fishhooks; I believe that this applies to sewn-plank boats as well.

7 East Polynesia: Historical Linguistics

The Polynesian language family is a typologically close family of several dozen languages and dialects. It and its nearest relatives, the Fijian languages and Rotuman, make up the Central Pacific language family. The subgrouping of the Polynesian languages is mostly uncontroversial, and its accepted subgroups have been supported by a large and growing body of grammatical and lexical evidence (Green 1966, 1985; Pawley 1966; Howard 1981; Wilson 1985; Marck 1996, 2000).

The following sketch summarizes an accepted subgrouping of the Polynesian languages, omitting some languages irrelevant to this discussion and some less established subgroupings:

POLYNESIAN
  TONGIC
    Tongan
    Niuean
  NUCLEAR POLYNESIAN
    East Futuna, East Uvea, Rennellese, Tikopia, Pukapuka, various other languages of Western Polynesia and Vanuatu
  ELLICEAN
    Sāmoan, Tuvaluan [Tūvalu (Ellice islands)], Luangiua, Takuu, Sikaiana [Solomon Islands], various other languages of Vanuatu, the Solomon Islands and Micronesia
  EAST POLYNESIAN
    Rapanui [Easter Island]
  CENTRAL EAST POLYNESIAN
    MARQUESIC
      Hawai’ian
      Marquesan [dialect complex, divided into North and South Marquesan]
      Mangarevan
    TAHITIC
      Tahitian [Society Islands, including Tahiti]
      Tuamotuan
      Austral languages: Rimatara, Rurutu, Tupua’i (Tubuai), Ra’ivavae
      Māori [New Zealand. Several dialects]
  Cook Island Māori: Rarotongan, Mangaian, Aitutaki, Tongareva

23 Recent work (Walworth 2012) questions the validity of Marquesic and Tahitic as valid taxa. The arguments in this paper do not rely significantly on that part of the classification.

24 The position of the closely related Austral languages within Tahitic is unclear, as they are poorly documented and have been heavily influenced by Tahitian since European contact.
All Polynesian languages have a small phonemic inventory, of which only the stops and the nasals matter to this discussion.\textsuperscript{25} Proto Polynesian had the stop consonants \^{*}p, \^{*}t, \^{*}k and \^{*}?\textsuperscript{25}, and the nasals \^{*}m, \^{*}n, and \^{*}?\textsuperscript{25}. The PPN bilabials \^{*}p and \^{*}m are unchanged in all the Polynesian languages.\textsuperscript{26} The following discussion is based on Marck (2000), Biggs (1978), Hovdhaugen (1986) and Fischer (1999).

The three non-bilabial stops participate in a pull-chain shift, \(\emptyset < \^{*}? < \^{*}k < \^{*}t\), which has progressed to some stage in nearly every Polynesian language. Examples of the stages of this chain shift are given in table 1. Each set of reflexes here indeed corresponds to a stage in the pull-chain: no language has shifted \^{*}k to \^{*}? until after \^{*}? had been lost, and no language has shifted \^{*}t to \^{*}k until after \^{*}k had shifted to \^{*}?\textsuperscript{26}. The \^{*}t > \^{*}k shift in Hawaiian is late, and has not affected most of Hawaii until the nineteenth century, progressing from east to west (Blust 2004). This shift has a similar history in Samoan, where it is at present only reflected in the colloquial register (Blust 2004; Hovdhaugen 1986).

<table>
<thead>
<tr>
<th>Stage</th>
<th>^{*}p</th>
<th>^{*}t</th>
<th>^{*}k</th>
<th>^{*}?</th>
<th>Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>(t)</td>
<td>(k)</td>
<td>(\emptyset)</td>
<td>Rapanui, Tongan, East Uvea, East Futuna, Rennellese; Proto Polynesian (PPN), Proto East Polynesian (PEP)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>(t)</td>
<td>(k)</td>
<td>(\emptyset)</td>
<td>All Polynesian languages not elsewhere in this table, including North Marquesan, Mangarevan, Maori and Niuean; Proto Central East Polynesian (PCE)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>(p)</td>
<td>(t)</td>
<td>(\emptyset)</td>
<td>Hawaiian (modern Ni’ihau and most older dialects), South Marquesan, Tahitian (standard), Austral languages, Samoan (formal)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>(k)</td>
<td>(\emptyset)</td>
<td>All Polynesian languages not elsewhere in this table, including North Marquesan, Mangarevan, Maori and Niuean; Proto Central East Polynesian (PCE)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Reflexes of Proto Polynesian stops in the Polynesian languages. The stages of the chain shift are numbered here to highlight the chronological progression.

As for the nasals, PPN \^{*}n has persisted as \(n\) in all Polynesian languages except Luangiua and Colloquial Samoan, where its reflex is \(\eta\), merging with that of PPN \^{*}?\textsuperscript{25}. PPN and PEP \^{*}?\textsuperscript{25} is reflected as \(\eta\) in most Polynesian languages, with the following exceptions: \(\eta > n\), merging with \^{*}n, in Hawaiian, South Marquesan, the Bay of Plenty dialect of Maori, and the Austral languages of Rimatara and Tubuai; \(\eta > k\), merging with \^{*}k, in North Marquesan (except in Taipivai on Nuku Hiva), and in South Island Maori; and \(\eta > ?\) in Tahitian. It is possible that Tahitian \^{*}? had once merged with \^{*}k, as in North Marquesan, followed by the plosive shift \(k > ?\) as described

\textsuperscript{25} Other sound changes which affect various East Polynesian languages are \(s > h\) (all except Tongareva), \(f > h\) (various), \(l > ?\) (Mangarevan, Rarotongan, some Austral languages) and \(r > ?\) (Marquesan). Sound changes are of little use in the subgrouping of the Polynesian languages, since many of them recur independently in separate branches of the family.

\textsuperscript{26} I use the following standard abbreviations: PPN (Proto Polynesian); PEP (Proto East Polynesian); PCE or PCEP (Proto Central East Polynesian).
above.

<table>
<thead>
<tr>
<th>*m</th>
<th>*n</th>
<th>*ŋ</th>
<th>Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>η</td>
<td>All Polynesian languages not elsewhere in this table</td>
</tr>
<tr>
<td>n</td>
<td>k</td>
<td></td>
<td>North Marquesan (except Taipivai), Ngāi Tahu (S. Island) Maori</td>
</tr>
<tr>
<td>n</td>
<td>?</td>
<td></td>
<td>Tahitian, Rurutu</td>
</tr>
<tr>
<td>n</td>
<td></td>
<td></td>
<td>Hawaiian, South Marquesan, Bay of Plenty Maori, Rimatara, Tubuai</td>
</tr>
<tr>
<td>η</td>
<td></td>
<td></td>
<td>Colloquial Samoan, Luangiua</td>
</tr>
</tbody>
</table>

Table 2: Reflexes of Proto Polynesian nasals in the Polynesian languages

Some degree of documentation exists for nearly all Polynesian languages, and extensive dictionaries exist for many.\(^\text{27}\) POLLEX (Biggs and Clark 1993) is an extensive comparative lexical database of the Polynesian languages, maintained since the 1960s, and an invaluable tool in Polynesian comparative linguistics.\(^\text{28}\)

8 Terms for ‘boat’ in Southern California: The Documentary Evidence

In this section I will review the sources for the Californian words for ‘boat’ under discussion, so as to start from reliable and precise phonetics and semantics, as far as sources allow.

8.1 The Gabrielino Record

Gabrielino was spoken in what is now the Los Angeles basin, adjacent inland valleys, and the southern Channel Islands. It belongs to the Takic language family; Takic is a subgroup of Northern Uto-Aztecan, along with Numic, Tübatulabal, and Hopi. The Takic languages all are or were spoken in Southern California. The following chart omits a few poorly documented languages not germane to the discussion.

**TAKIC**

**SERRAN**

Serrano [San Bernardino Mountains and adjacent Mojave Desert and inland valleys]

Kitanemuk [Western Tehachapi Mountains, SE Central Valley, Antelope Valley]

**GABRIELINO**

Gabrielino proper [Los Angeles basin south past Newport Bay, and inland valleys]

Fernandeño [San Fernando Valley]

**CUPAN**

LUISEÑO [Coast from south of Newport Bay to Carlsbad, and adjacent mountains]

---

\(^{27}\) Standard dictionaries used in this paper are: Rapanui — Englert (1978); Fuentes (1960); Hawaiian — Pukui and Elbert (1986); Marquesan — Dordillon (1904); Mangarevan — Rensch (1991a); Tahitian — Andrews and Andrews (1944); Maori — Williams (1971); Cook Island Maori — Buse and Taringa (1995); Mauriati et al. (2006); Shibata (2003); Tuamotuan — Stimson (1964). Other sources are mentioned as necessary.

\(^{28}\) Klar and Jones refer to the 1994 version of POLLEX. The 1993 version, which I use, is not substantially different (Marck 2000:6). An updated version is available online at http://pollex.org.nz.
Luiseño proper [Southern coastal part of the above and mountains]
Juaneño (Acjachemem) [Northern coastal part, around San Juan Capistrano]
CAHUILLA-CUPEÑO
Cahuilla [Inland area, from San Bernardino to Salton Trough]. Three dialects:
Pass Cahuilla (Wanikik) [San Gorgonio Pass]
Desert Cahuilla [Coachella Valley]
Mountain Cahuilla [Santa Rosa and San Jacinto Mountains]
Cupeño [Around Warner Springs, near San Luis Rey River headwaters]

Fernandeño was one of several closely related dialects of Gabrielino, accorded its own name through the presence of the mission at San Fernando. Juaneño, likewise, is closely related to Luiseño. Serrano and Kitanemuk were mutually intelligible.

All Takic languages are largely suffixing languages. Isolated, non-possessed nouns take one of several language-specific and lexically determined suffixes, known in the Uto-Aztecan literature as absolutes. Possessed nouns take a possessive prefix and omit the absolutive suffix, e.g., Luiseño ’hu-la {arrow-ABS} ‘arrow’, no-‘hu: {1SG-arrow} ‘my arrow’ (Elliott 1999:22). The historical phonology of Takic is fairly well understood (Bright and Hill 1967; Langacker 1970; Munro 1990; Hill 2007; and others).

Gabrielino is poorly documented. Some of the what is known about the language comes from several early wordlists (Hale 1846; Taylor 1860c; Gatschet 1879 — all published in McCawley 1996; Kroeber 1907, 1909), of varying phonetic quality. The greatest amount of information of the language comes from two indefatigable linguists-collectors of the early twentieth century, C. Hart Merriam and John P. Harrington, both of whom documented otherwise barely-known languages. Merriam collected extensive wordlists for Gabrielino (and other languages), including words for precisely identified animals and plants.29 While his vocabulary is the largest existing lexical resource for Gabrielino, his orthography was phonetically naive, and his transcriptions were imprecise, inaccurate and inconsistent. Harrington’s notes, while covering somewhat less lexical material, are extensive, phonetically accurate, and detailed.30 Most of Harrington’s materials are unedited and unpublished, but Munro (2000) includes many elicited sentences from his notes, and Bright (1976) contains basic vocabulary based on his notes.

8.2 The Gabrielino Record: tarainxa

One Gabrielino word for ‘boat’ is known from the following records:

<table>
<thead>
<tr>
<th>Form</th>
<th>Meaning</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;trainxe&gt;</td>
<td>‘canoe; boat’</td>
<td>[Hale 1846; McCawley 1996:282]31</td>
</tr>
<tr>
<td>&lt;Ta-rin-ha&gt;</td>
<td>‘canoe, boat’</td>
<td>[Taylor 1860c; McCawley 1996:272]</td>
</tr>
<tr>
<td>&lt;Tah-ri’ng-hah&gt;</td>
<td>‘boat (bundles of tules)’</td>
<td>[Merriam 1903a; McCawley 1996:246]</td>
</tr>
</tbody>
</table>

29 Merriam’s notes, digitized from the microfilm, are available at http://www.archive.org. Finding lists are available through the Bancroft Library’s website.

30 Where Harrington’s notes are quoted here, they are referenced by microfilm series, reel and frame; e.g., Harrington (3:102:582) is reel 102, frame 582 of microfilm series 3 (Southern California). Harrington’s Gabrielino work was carried out in 1914-1917, and again in 1933.

31 Hale gives another Gabrielino form for ‘canoe, boat’ as well, <nikin>, not recorded elsewhere. I believe it is the result of an elicitation error. In Gabrielino, the independent possessive 1SG pronoun, ‘mine’, is given as <ne-hin’> by Merriam (1903a). It appears that what Hale actually elicited was not ‘boat’, but ‘my things’. A similar
/taˈrain xa/ (without specifying vowel lengths) is the most straightforward underlying form consistent with all three elicitations, where /x/ is realized as [χ] or [ʁ]. I read Hale’s <trainχ> as [ˈtrainˌχa], with reduced vowels in the final and initial unstressed syllables. I emphasize that the nasal is /n/, not /ŋ/; this point is important in the later analysis. For Hale’s exemplar, this is straightforward, as he consistently uses the symbol <η> for the velar nasal, e.g., <atōnjin> ‘mouth’, cf. Kroeber (1907:74) ni-thon. In comparing Hale’s wordlist to others, it appears that <η> and <n> are always transcribed correctly.

Taylor’s <Ta-rin-ha> is also consistent with /taˈrain xa/, within the limits of variation in his orthography. Taylor’s orthographic <i> usually corresponds to [ι], and [q] and [x] are usually represented by <k>. However, Taylor’s form for ‘blood’ is <a-hin>, representing [ʔa-χain] ‘his blood’. In other words, Taylor’s <h> may stand for [χ] and <i> may stand for [ai] (as in English). With these orthographic variants, Taylor’s <Ta-rin-ha> fits with the phonetic form tarainxa suggested here. Again, the nasal is not /ŋ/, which Taylor consistently transcribes <ng>.

In Merriam’s <Tah-rí-ŋ-hah>, <i> represents [ai], as it does elsewhere, following common English dictionary phonetic spelling. The first <i> in <hah> could represent [h] or [χ]. <ng> appears to be a nasal velar, in contrast to Hale and Taylor. I read the phonetic form as [taˈraiŋxa] or [taˈrainqxa], underlyingly /taˈrain xa/ with n assimilating to n under the influence of a following χ. The affrication of [q] to [qχ] also occurs in Serrano (Hill 1967:4). Merriam’s informant, Narcisa Rosemyre, had a Serrano father and a Gabriélino mother, and she spoke both languages (McCawley 1996:17); the pronunciation here might be influenced by her Serrano.

I emphasize again that the η or χ does not represent an underlying back consonant. If the phonetic form recorded by Merriam had been [taraɪŋa], he would have recorded it as something like <Tah-rí-ŋ-ah>, as in his <Ah-soo-ŋ-ah> ‘inside’, cognate with Luiseño -ṣun-ŋa ‘inside’, literally {heart-LOC} (Elliott 1999).

Merriam lists another Gabriélino form, <Hoo-pá-kah tar-rí-n-hah> ‘A kind of pointed instrument’ (Merriam 1903a [McCawley 1996:245]). I will discuss this form and its semantics below, but for now I will only mention that I believe <tar-rí-n-hah> in this compound is the very same word as <Tah-rí-ŋ-hah> ‘tule boat’, the differences reflecting either phonetic variation or Merriam’s inconsistent transcription.

Klar and Jones claim a different form of the word, taraina, <tarayna> in their orthography (A:474; B:388, 390, 396n10; C:766; D:89; E:175). They base it on a comment by Pamela Munro, but mention no primary sources or any other justification for it. They mention the form tarainxa only in B (p. 396, n. 10), basing it on a comment by Jane and Kenneth Hill, and again quoting no primary sources. Evidently, Harrington also considered taraina a possibility, as he

32 “The χ is, in those tongues, a somewhat deeper guttural than the Spanish jota.” (Hale 1846:535)
33 Orthographically <ni-tohín>. In Kroeber’s notation <i> is the velar nasal η; “h, nasal of k as n is to t” (Kroeber 1907:90). The stems here are preceded by the possessive prefixes ʔa-, 3sg, or ni-, 1sg.
34 Kroeber (1907:76) <mu-xain> ‘their blood’ (in the Fernandeño dialect); Hale (1846) <akχain>; Gatschet (1879) <akχain>. In Gatschet’s vocabularies, <kh> is “a surd guttural aspirate, the German ch...”; <ai> is “as in aisle (‘long i’ in pine)” (Gatschet 1879:423).
35 Compare also the affrication in Cupeño daraŋxa ‘orange’, from Spanish naranja (Hill 2005:180).
attempted to elicit both taraina and tarainxa from three separate informants while reeliciting Hale’s and Taylor’s wordlists. My guess is that he considered the possibility that the <h> in Taylor’s <Ta-rin-ha> was a record of non-phonemic aspiration, although there’s no other evidence for such aspiration in the language. In any event, taraina is inconsistent with Hale’s and Merriam’s forms, while tarainxa is consistent with all three, as I have shown.

Klar and Jones refer to taraina/tarainxa as the Gabrielino “word for ‘boat’ in general”. This is inaccurate. Both Hale and Taylor give an imprecise gloss for the word, without specifying the type of boat in question, but they did not attempt to elicit the words for different types of boats, and their informants may have been familiar with only one type. On the other hand, Merriam’s Gabrielino wordlist was based on a questionnaire tailored for work with California Indian languages. There are separate entries in Merriam’s questionnaire form for ‘Boat (log dugout)’, ‘Canoe’, ‘Kayak or bidarka’, and ‘Boat (bundles of tules)’. In the Gabrielino vocabulary the first three entries are left blank, but <Tah-rí’ng-hah> is recorded for ‘Boat (bundles of tules)’. This suggests that the word refers specifically to tule boats, and not to boats in general.

I have no details on Hale’s informant. Taylor’s informant, Juan de Parma, was born and raised near the San Gabriel mission, some 30 km from the coast. Merriam’s informant, Narcisa (Mrs. James) Rosemyre, grew up there as well (McCawley 1996:17). It may be that the inland Gabrielino were the first to lose the collective memory of the ocean-going plank canoe, while still remembering the more widespread tule boat.

8.3 The Gabrielino Record: \textit{tiʔaːt}

The presumed Gabrielino word for ‘planked canoe’ is recorded only once in Harrington’s notes, and apparently nowhere else. Its source is the informant José María Zalvidea:

\begin{quote}
tiʔát lancha, cayuco. Z[alvidea]. tiʔá†, canoe. Z. It was so called because it carried many people. \textquoteleft at, people. Made with boards, calked with mineral tar, and tied together with string made of horse nettle, he volunteers. [Harrington 1986, 3:102:582]
\end{quote}

In two later attempts to elicit the word, Harrington’s informants were not familiar with the form. Both times Harrington spells the word as <teʔaːt>; Either teʔaːt or tiʔaːt is an acceptable reading, since Gabrielino neutralizes the \textit{i:e} contrast in unstressed positions. Both informants were familiar with tule boats. I agree with Klar and Jones (B:388-389) that Zalvidea’s

\begin{footnotes}
\footnotetext[36]{Harrington apparently had no access to Merriam’s vocabulary.}
\footnotetext[37]{“N[escit] teʔaːt but ev[idently] sic. In[formant] heard that the island indians came over here to mainland in the tule fixed some way” (Harrington 1986, 3:103:515). “N[escit] G[abrielino] *teʔaːt, canoe. In[formant] supposes they are of tule, for he heard that tejian [they wove] these boats of tule. Pl[ural] tetfiʔatam” (Harrington 1986, 3:103:87). For the plural form, Harrington notes that the first \textit{a} is indeed long and the second is indeed [a], not [o], confirmed by repeated elicitations. Klar and Jones (B:389) quote Pamela Munro as saying this plural, teʔiʔatam, is anomalous, teʔiʔatam with a short \textit{a} being the expected form (and see Munro 1983:291-297), and Harrington obviously found the plural form odd as well. But since the informant did not know the singular form teʔat, he must have produced what was to him the regular plural inflected form, perhaps an idiosyncrasy of that informant’s speech. Harrington’s Serrano informant, on a visit to San Pedro, produced \textquoteleft tfaʔat\textquoteright as the word for ‘boat’, as well as ‘basket’ (Harrington 1986, 3:101:416). This Serrano word is recorded elsewhere with the meaning ‘basket’ alone. This might be an interpretation of the Gabrielino tiʔaːt by folk etymology.}
\end{footnotes}
explanation is unsatisfactory. As they point out, Zalvidea’s ‘<aptured’ (\textit{ʔat}) is obscure, and the usual Gabrielino word for ‘person, people’ is \textit{taxac: \textit{ʔat}} might have occurred to Zalvidea based on its phonetic similarity to the Luiseño \textit{ʔatax}, ‘person, people’. His explanation does not address the first part of \textit{tiʔat}, and appears to be a folk etymology.

I see no reason to doubt Zalvidea about \textit{tiʔat} meaning specifically a plank canoe, as opposed to a tule boat or a boat in general. Zalvidea’s father was from Santa Catalina Island, some other ancestors of his were coastal as well (McCawley 1996:17-18), and so he was more likely to know about ocean-going boats than Harrington’s and other linguists’ informants, who were from near the San Gabriel mission, tens of kilometers away from the coast.

8.4 The Chumash Record: \textit{tomol(o)}

By and large, I agree with Klar and Jones’s reconstruction of the Proto Southern Chumash \textit{*tomolo} (B:372-373, 379-381) based on the attested forms in the Chumashan languages. The word is attested in various sources and dialects as \textit{tomol}, \textit{tmolo}, \textit{tomol} and \textit{tomo}; some of the variation is between dialects, and some different forms appear among speakers of the same dialect. \textit{tmolo} appears to be a variant of \textit{tomol}, following a general process of devoicing of \textit{l} word-finally and elsewhere (Klar 1977:21-22). \textit{tmo} appears in Ineseño and Isleño Chumash, a variant of \textit{tomol} with a reduced unstressed syllable. The \textit{m} is glottalized in some Barbareño elicitations. The earliest record of Barbareño \textit{tomol}, ‘lancha o canoa’, is from the Portolá expedition of 1769 (Costansó 1770:40); it appears in various later wordlists and throughout the notes of John P. Harrington.

Based on the forms \textit{tomolo} and \textit{tomol}, Klar and Jones reconstruct the protoform \textit{*tomolo}, with \textit{tomol} the result of final vowel deletion. That is the reconstruction also given by Klar (1977:76). Arguably, one could reconstruct the protoform as \textit{*tomol}, with an echo vowel appearing in some dialects and then further phonologized.

Klar and Jones (B:380-381) further attempt to use internal reconstruction to derive the protoform \textit{*tomolo} from an earlier \textit{*tomoloʔ} (phonetically \textit{[tomolʔo]}), which would better fit their proposed Polynesian source. As far as I can tell, the additional syllable is adduced to explain the glottalization of the Barbareño variant \textit{tomol} and the Ventureño plural \textit{tmtomoʔol}, and which is explained as a result of regressive assimilation to the \textit{ʔ} of the following syllable, since lost. While this scenario is possible, it is not clear to me whether the glottal might not be explained through other routes, which do not require the longer older form. In any event, any of the proposed earlier forms \textit{*tomol}, \textit{*tomolo} or \textit{*tomoloʔ} agrees with my discussion of the word below.

9 The Proposed Polynesian Etymologies

In the following sections, I will discuss the Polynesian etymologies suggested by Klar and Jones for the California words for boat, \textit{tarainxa}, \textit{tiʔat} and \textit{tomol(o)}. The California words would have to match those Polynesian words spoken at the inferred time of contact. As discussed above, the
Polynesian language spoken then would have been something like reconstructed Proto East Polynesian. I find that none of the three proposed Polynesian etymons match the meanings Klar and Jones would assign to them. Moreover, for *tarainxa and *tomol(o), the phonological shapes of the Polynesian words fail to match those of the Californian words. In other words, the Polynesian words did not sound as Klar and Jones claim they did, did not have the right meaning, and so cannot be the sources of the proposed borrowings.

9.1 Gabrielino *tarainxa as Polynesian

The case for Gabrielino *tarainxa or *taraina as a Polynesian loan is first brought up in A (pp. 475-476) and elaborated in B (pp. 390-394, 396). According to Klar and Jones, the stem to be analyzed is *taraina (B:396). They explain the form *tarainxa, suggested to them by Jane and Kenneth Hill, as the Polynesian *taraina, with a Gabrielino “adjectival suffix” -xa, which is given no further explanation or justification.

Leaving -xa for the moment, I next examine Klar and Jones’s claimed etymology, Gabrielino *tarainxa ‘boat’ < Polynesian **taraina ‘carved object’ < PPN *tarai ‘hew, carve’ + **-na ‘nominalizer’, which they compare to the attested Hawaiian kalaina ‘carving’ < kaːlai + -na (Pukui and Elbert 1986; Elbert and Pukui 1979:81). 39

The Hawaiian nominalizer -na is a regular reflex of Proto Polynesian *-ŋa (Biggs and Clark 1993, Krupa 1982:52). Phonologically regular reflexes of *-ŋa are attested in all the major languages of East Polynesia and in many outside it. 40 As noted above, *ŋ changed to n only in Hawaiian, South Marquesan, one Maori dialect, two neighboring Austral dialects, and nowhere else in Polynesia. According to Elbert (1982), the South Marquesan and Hawaiian *ŋ > n shift may have a common origin, in which case the shift occurred independently at most three times; very likely one or more of the occurrences of *ŋ > n are due to some old language contact. In any case, it occurred in the two Marquesian languages, Hawaiian and S. Marquesan, after the differentiation first of Marquesian and then of the Marquesan dialects, in Bay of Plenty Maori after the differentiation of Tahitic and after the settlement of New Zealand, and in Tubuai and Rimatara after the differentiation of the Austral languages. There is thus no evidence of *-ŋa > **-na anywhere in East Polynesia at the time the plank canoe first appeared in California. An East Polynesian word for ‘carving’, cognate with the Hawaiian kalaina, would have been at that time *taraɪa, not *taraina, and would have been borrowed into pre-Gabrielino as *taraɪa.

If the word had been borrowed into Gabrielino as *taraɪa, it would not have changed since then to taraina. The distinction between n and ŋ in Gabrielino goes back at least to Proto Takic, as the following examples show: 41

---

39 I use ** to mark forms which I believe did not exist, in reconstructed or putative languages.
40 In Nuku Hiva Marquesan the nominalizing suffix -ka < *-ŋa is no longer productive, but still appears with a small, closed set of verb stems (Zewen 1987:100).
41 The n/ŋ distinction goes even farther back, to Proto Northern Uto-Aztecan, and possibly to Proto Uto-Aztecan (Campbell 1997:136-137; Dakin 2001). See Kroeber (1907) for examples from elsewhere in Northern Uto-Aztecan.
The Plank Canoe of Southern California

Gabrielino -ŋa ‘locative’ (Munro 2000, ex. 8, 38, 68); Luiseño -ŋa ‘in’, ‘on’, ‘among’, etc. (Elliott 1999).

Gabrielino ni- ‘1SG possessive’ (Merriam 1903a); Luiseño ni- (Elliott 1999); Kitanemuk ni- (Anderton 1988).

Gabrielino noŋ- ‘tongue’ (Hale 1846; Taylor 1860c; Merriam 1903a; Kroeber 1907); Cupeño naŋ- (Hill 2005); Kitanemuk niŋ- (Kroeber 1907; Anderton 1988).

Gabrielino sun- ‘heart’ (Kroeber 1907); Proto Cupan şu-ːni (Munro 1990); Serrano huːn- (Ramón and Elliott 2000:563; *s is often reflected as h in Serrano and Kitanemuk.)

In other words, n and ŋ have existed side by side in Gabrielino and its ancestors from well before the advent of sewn plank boats and up to the historical period. If a Polynesian word *taraiŋa were borrowed into an earlier stage of Gabrielino, it would still appear in the historically attested form as **taraiŋa, not taraina or tarainxa, contrary to what is recorded. On phonological grounds, then, the Polynesian derivation fails.

As mentioned before, Klar and Jones don’t explain the -xa at the end of the form tarainxa, beyond mentioning that -xa appears as a Gabrielino adjectival suffix, and suggesting the decomposition taraina-xa. The ending -xa indeed appears in a few Gabrielino adjectives: kwahoxa ‘red’ (Kroeber 1909:251), jumaxa ‘black’ (Fernandeño; Kroeber 1907), hupaxa ‘sharp’ (Merriam 1903a), and some others. But of the ca. 75 adjectives in Merriam (1903a), only about 10 end with <kah>, <chah> or <hah>. -xa is thus by no means an obligatory or common marker of adjectives. More essentially, there is no explanation why an adjectival suffix would be added to a borrowed noun to create another noun. If anything, one would expect that a noun borrowed into Gabrielino would take an absolutive suffix, resulting in something like **taraiŋa-t, just as Klar and Jones propose for forming Gabrielino tiʔaː-t from Polynesian *ti-a.

Either form — with or without a Uto-Aztecan suffix — could be borrowed from a Polynesian language, with *taraynxa being more fully nativized in Gabrielino than is tarayna. [B:396]

does not hold. There is no morphological path by which to go from Polynesian **taraina (even if that form had existed) to Gabrielino tarainxa.

On semantic grounds there are difficulties as well. The reflexes of the PEP nominalizer *-ŋa have two functions, in which they resemble the English -ing. Most commonly they produce pure nouns associated semantically with the parent verb stem. For example, Hawaiian kalaina ‘carving’ < kalai ‘to carve, hew’. In some languages, they produce verbal nouns, similar to English gerunds, which share some syntactic properties of both nouns and verbs; for example Maori moega ‘sleeping’ (and also ‘bed’) < moe ‘sleep’ (Krupa 1982:50). Klar and Jones are not clear as to which sense their presumed Polynesian word *tarai-na would have. They refer to a “hewn object” (D:89; E:175), but elsewhere to “the process of adzing or hewing — the quintessential technique in maritime construction” (A:476; B:390).

With the first use of *-ŋa, **taraina would mean something like ‘a carved object’. A reflex of *taraiŋa is recorded only in Hawaiian. It is not used anywhere to refer to boats, hulls or planks. For boat planks, the word usually used is papa (Kamakau 1869 and Kamakau 1976:118 for Hawaiian; Handy 1923:157 for Marquesan), from a stem of pan-Malayo-Polynesian distribution

---

42 The Gabrielino -ŋa locative suffix still appears in some present-day Southern California place names: Cahuenga, Topanga, Tujunga, Cucamonga. Many more are recorded in Harrington’s notes.

43 By some estimates, Proto Takic started diverging around 1500 BC (Sutton 2009).
(Pawley and Pawley 1998:186), or else, in Hawaiian, the more generic *laʔau ‘(piece of) wood’.

While *kaːlai refers to boat-making in general in some Hawaiian compounds (e.g., *kahuna kaːlai waʔa ‘boat-building master’, Fornander 1917(5):613), there is no indication of *kaːlaina ever being used to describe a boat or any of its parts. Moreover, this use only appears in Hawaiian, and is therefore a late innovation, well past the permitted time range of Polynesian-California contact. This Hawaiian innovation fits well with the historian Samuel Kamakau’s account, in which the Hawaiian settlers arrived in plank canoes but changed their predominant boat type to the more easily built dugout, taking advantage of the availability of the Hawaiian koa tree (Kamakau 1976:118).

If **taraina is to be a verbal noun, it would be referring to the manufacture of some part of a Polynesian canoe, such as the hull, strakes, or planks. However, as with English gerunds, it would only be used within longer sentences, and would not be easily borrowed into Gabrielino, especially in the scenario of brief and casual interaction Klar and Jones envision for the Polynesian-Chumash contact.

In sum, there seems to be no way by which Proto East Polynesian *taraiŋa ‘carving’ — itself speculated and unsupported — could have become Gabrielino tarainxa ‘tule boat’. The phonological, morphological and semantic evidence all compel rejecting the proposed Polynesian source for this Gabrielino word.

9.2 Gabrielino tiʔaat as Polynesian

Klar and Jones would have Gabrielino tiʔaat derive from a Polynesian word, *tia (A:474-475; B:388-390, 393, 394; C:766; D:89; E:175-176). As shown here, the PEP *tia did exist, and is a good phonetic match to Gabrielino tiʔaat, after the addition of the Gabrielino absolutive suffix. The semantics, however, argue against that borrowing scenario.

Throughout their papers, Klar and Jones suggest several Polynesian words — ‘to sew’, ‘mast’, ‘type of boat’, ‘small sticks used in boat construction’, ‘to pierce’ — all of the form *tia, some cognate, but of different semantics and historical distribution:

...suggesting that the Gabrielino named their sewn-plank boat not after the source material (as did the Chumash) but after some feature of it (short pieces of wood or a mast (cf. the Hawaiian metaphorical extension ‘mast’), or a technique associated with building it (piercing the short pieces of wood to sew them together). [Klar and Jones B:389-390]

For any one of these etymons to be right, all the others have to be wrong. Casting a wide net for etymons in this way depends on semantic imprecision, and makes the argument less convincing than an argument based on a single more certain etymology. In any event, I examine here all of the proposed etymons.

9.2.1 tia ‘to sew’

The following are the documented senses of apparent reflexes of Proto East Polynesian *tia, arranged by language and family, based initially on POLLEX (Biggs and Clark 1993) and supplemented using the standard dictionaries listed in section 7. Translations and slight rephrasings are mine:
The Plank Canoe of Southern California

EAST POLYNESIAN
Rapanui: *tia. To sew (as a cape of tapa cloth).

CENTRAL EAST POLYNESIAN
Hawaiian: *kia. Spike, nail; post; pillar; mast; (?) type of boat (in compounds).
Mangarevan: *tia. To stick in, to drive a wooden piece into the ground, to drive in a nail.
Marquesan: (none)\(^{44}\)

TAHITIC
Tahitian: *tiatia. Small posts.
Tuamotuan: *tia. To stick in, as bunch of flowers; penis.
Maori: *tia. Peg, stake; to stick in a peg or a thatching needle; to adorn by sticking feathers in.
Rarotongan: *tia. Wedge, peg; to drive in peg or stake, to wedge in.\(^{45}\)
Tongareva: *tia. Stake; wedge; stuff, filling (as small stones jammed between large ones).

According to POLLEX, the root *tia ‘pole, stake’ is reconstructed as far back as Proto Malayo-Polynesian. Another sense, ‘to weave a net’ appears in the Tongic languages and in the outliers Tikopia, Takuu and Sikaiana, but not in any East Polynesian languages. The East Polynesian sense can be reconstructed as ‘peg’, ‘stake’, or any such object pushed into yielding matter; or as a verb signifying that action.

Klar and Jones most often associate *tia with the sense ‘to sew’, referring to the lashing together of canoe planks. For this sense they rely on the sense of PPN *tia reconstructed in POLLEX, ‘sew, stick in a peg or a needle, make a net’. But these three senses cannot all be reconstructed at every level of the Polynesian family tree. ‘Stake, peg’ and ‘to push in’ seem to be pervasive and stable throughout Polynesia. The specialized sense of ‘weaving a net’ appears in a number of Polynesian languages but nowhere in East Polynesia. The sense ‘to sew’, however, appears only in Rapanui, in what must be a local innovation in that language. Therefore, pace POLLEX, the sense ‘to sew’ cannot be reconstructed for Proto East Polynesian or Proto Central East Polynesian. In other words, *tia ‘to sew’ did not exist in the language of the Polynesians who would have sailed the eastern Pacific when sewn plank boats appeared in California.

\(^{44}\) Marquesan *tiʔa ‘mast’ is unrelated: it is a regular reflex of Proto Polynesian *tīla ‘mast, boom’, going back to Proto Oceanic. The sound change *r>*ʔ occurs in Marquesan but in no other Polynesian language.

\(^{45}\) POLLEX also lists Rarotongan *tiʔa ‘to close a sack by sewing’, quoting Savage (1962). Savage’s gloss actually reads “to bind or lash, such as in binding or lashing the lugs of the mouth of a sack filled with copra; after sewing up the mouth securely, the sewing twine is wound round the lugs on each side and the final fastening is done and a secure knot made to secure by binding or lashing and knotting”. This refers, then, not to the sewing but to the tying-up of the ends of the sack. The sense of tying or sewing does not appear in the later, comprehensive Rarotongan dictionary of Buse and Taringa (1995) or in any other Polynesian language, and is best left out of this cognate set.

Since Savage’s dictionary omits glottal stops, his <tiʔa> could represent tiʔa, which would be the regular reflex of PCEP *tiʔa ‘to close up, seal, patch, inlay’, cf. especially Tahitian *tiʔa ‘to join things together; to dovetail’ (Davies 1851) and Marquesan *tiʔa ‘to close, seal, plug, cover’ (Dordillon 1904). But neither tia nor tiʔa appear with that sense in any other Cook Islands dictionaries.
There are other reasons to reject *tia ‘sew’ as a Polynesian word describing the lashing together of boat planks. Although much traditional Polynesian boat-making knowledge has been lost, there exist records of several verbs used to describe this action, which I will review here. None of them is a reflex of *tia.

Some of the most extensive published native records of traditional Polynesian boat-building are from Hawaii.\textsuperscript{46} I rely on the accounts of the nineteenth century native historians Samuel Kamakau and Davida Malo, of the ethnographer Abraham Fornander, and on one anonymous account, all published in Hawaiian and English.\textsuperscript{47} I go in some detail here into the existing records of Hawaiian boat construction, not only to document the extent of the relevant Hawaiian vocabulary, but also to give some of the flavor of the various terms in context. The following are citations mentioning the lashing together of planked boats, and the lashing of strakes and end pieces onto dugout hulls.\textsuperscript{48} Corresponding Hawaiian words and phrases and their English translations are marked in boldface:

(1) A laila, kāpili ‘ia ka lā‘au, he ‘ahakea paha, he lā‘au ‘ē a‘e paha; e kāpili mua ‘ia nā mo‘o, a holo ‘ia i ka ‘aha a pa‘a ia; e kāpili ‘ia nā maka ihu a pa‘a ia mau wahi, e kāpili ‘ia nā kupe hope, a holo ‘ia a pa‘a i ka ‘aha; a laila, pau ke kāpili ‘ana o ka wa‘a. [Malo 1987:89]

“After that were attached the carved pieces made of ahakea or some other wood. The rails, which were attached to the gunwales, were the first to be fitted and sewed fast with sinnet. The carved pieces at bow and stern were the next to be fitted and sewed on, and this work completed the putting together of the body of the canoe.” [Malo 1903:171]

(2) Kālai ihola nā kāhuna i ka wa‘a a oki, a kāpili ihola, a kau ka pu‘aki, pā‘ele a maika‘i... [Kamakau 1865]

“The expert canoe builders hewed the canoe hulls, attached the parts, put on the rigging, and painted the canoe black” [Kamakau 1991:3]

\textsuperscript{46} Traditional boat construction is also described in many early Maori texts. Much of the specialized Maori terminology is already available in complete quotations within the dictionary of Williams (1971). At the time of European contact, New Zealand Maori did not have fully planked canoes, though they had boats with washstrakes sewn on.

\textsuperscript{47} There are other Hawaiian language sources, published and unpublished, on traditional canoe construction. I present here all the sources known to me for which both the Hawaiian texts and their English translations have been published. The nineteenth century Hawaiian newspapers quoted here can be accessed online through http://nupepa.org or http://ulukau.org.

\textsuperscript{48} Longer texts in Hawaiian are given in standard orthography: ‘<‘> is the glottal stop, and the macron indicates a long vowel. The published Hawaiian texts of Kamakau and Fornander do not indicate glottal stops or vowel length, and Anonymous (1939) uses them inconsistently; for those, the vowel macrons and glottal stops given here are mine. I am grateful to Puakea Nogelmeier for reviewing my transcriptions and translations and providing corrections. Some of the published translations have been rephrased.
(3) Ua ‘ōlelo ʻia mai nō hoʻi e ka poʻe i ʻike maka i ia mau waʻa, he mau waʻa kuʻi, a he mau waʻa ʻāpana lāʻau i humuhumu ʻia a paʻa i ka ʻaha. Pēlā aku paha ka waʻa o ka poʻe kahiko o Hawaiʻi nei. [Kamakau 1867a] “People who have seen these canoes [war canoes of eighteenth century Oʻahu chief Pele-iʻōlani] have told me that they were ‘joined canoes’, made of pieces of wood sewn securely with coconut husk cords. It would seem that this is the kind of canoe that ka poʻe kahiko [the people of old] had.” [Kamakau 1991:117-118]

(4) Ua ‘ōlelo ʻia, ʻo Kāneaʻiaʻi mā, nā waʻa kaua o Peleʻiʻōlani, he mau waʻa kuʻi, a he mau waʻa ‘āpana i hana akamai loa ʻia. [Kamakau 1867b] “It was said that Kāneaʻiaʻi and such, the war canoes of Peleʻiʻōlani, were joined canoes, boats pieced together and very skillfully crafted.” [translation mine; see Kamakau 1992:240]

(5) ...he kālai ʻāpanapana, a nui ka ʻāpana, a laila kāpilipili, a lilo i waʻa... ‘O kēia mau lāʻau, ʻo ia ka papa o ka waʻa o ka poʻe kahiko i kāpilipili ʻia i kahiko, i ke au o Wākea mā, a mamua aku a mahope mai, ʻo nā waʻa o Wākea, ʻo Kumuʻeli ka inoa, ua kāpilipili ʻāpana ʻia... ...o ka waʻa o ka poʻe o Kahiki mā i holo mai ai i Hawaiʻi nei, he waʻa ʻāpana kāpilipili kō lākou... [Kamakau 1869] “The woods were shaped into pieces — many pieces — then attached together; this became a canoe.” “From these trees were made the worked woods that ka poʻe kahiko fitted together for canoes in the time of Wākea and before and after his time. Wākea’s double canoe, named Kumuʻeli, was of pieces of wood fitted together and so was kaloloamaile [kaloliamaiele], the double canoe of Kuhaʻilima. In the time of Laka mā [ʻLaka and his people’] canoes were hewn out of koa — one large koa tree made one large canoe. This made the work less burdensome and wearisome, and shortened the labor. This was also true in the times of Hakalanileo, Niheu and Kana. But the canoes of the voyagers who sailed from Kahiki to Hawaii were made of joined pieces...” [Kamakau 1976:118-119]

(6) ...a laila, houhou nā puka o ka waʻa, ʻo ka holo nō ia o ka ʻaha, no ka mea, ua pau i ke kauli ʻia mai ka uma o ka ihu a hiki i ka uma o ka lāʻau hope... [Kamakau 1869] “Then he drilled holes in the canoe for the sennit cords, and readied the lashings of the canoe from the curve of the bow to the back curve of the back end piece.” [Kamakau 1976:121]

(7) A ma ka uma o nā lāʻau hope o ka waʻa, a paʻa ia i ka holo ʻaha ʻia, a laila, hahau ka pola i waena konu o nā waʻa... [Kamakau 1870] “After the clamping down of the rear pieces of the canoe and the fastening with running sennit-cord, the platform midway between the canoes was lashed on.” [Kamakau 1992:42]

49 kuʻi is spelled <ku-i> in Kamakau’s published text, presumably being explicit about the presence of a glottal stop.

50 It is not clear from the context whether the text refers only one boat, Kāneaʻiaʻi, or many. The plural is sometimes used to refer to singular double-hulled boats.

51 Kahiki is not only Tahiti in the Societies, but distant lands in general, in this case Central East Polynesia: Kahiki place names earlier in the text include locations in the Societies and the Marquesas.
(8) ‘O ka mea e pa’a ai kēia mau lā‘au he kaula ‘aha.
“These pieces are tied on to the canoe [rims] with the sennit.” [Fornander 1917(5):612-613]

(9) ‘O ka ho‘opa’a ‘ana, me ke kaula ‘aha e humuhumu ai a pa’a i ka wa’a.
“In fastening, the sennit is used to tie these [braces] on to the canoe.” [Fornander 1917(5):636-637]

(10) Houhou nō ho‘i ka puka ‘aha ma ia mau lā‘au mai mua a hope, ho‘okomo ka iwi lā‘i. Pau kēia, ho‘opilipili mai mua a hope; kau koa‘ekea, kauali mai mua a hope. Pau kēia, ho‘okomokomoko ka ‘aha ma nā puka a pau mai mua a hope...
“...Holes were also bored into those pieces of wood from front to back (and) ti leaf stems inserted. After this was done, they were fitted from prow to stern; adjusted and perfectly fitted from front to back. When this was done the sennit braid was threaded into all the holes from prow to stern...” [Anonymous 1939:158-159]

“...When the piecing of the canoe parts was done, the thwarts were put in [and] tied firmly in place.” [Anonymous 1939:158-159]

The verb roots in these texts referring to sewing together of boat parts are (see also Pukui and Elbert 1986): humu ‘sew’, kauali ‘lash’ (perhaps < li: ‘lace, as shoe’; also used as a noun), and kuʔi ‘join; sew’; more specifically holo ‘thread through’, literally ‘run’ (cf. English ‘running stitch’); and more generally paʔa ‘join’ (also used adverbially, a paʔa ‘tightly’), hoa ‘tie, lash’, kapili and hoʔopilipili ‘attach’, koaʔekea ‘adjust and fit parts to the canoe body’, and hoʔokomokomoko ‘insert’ (< hoʔo ‘causative’ + komo ‘enter’, reduplicated for repeated action), referring to sennit cords pushed through plank holes.

Of the roots specifically referring to plank sewing, humu has cognates in Marquesan (Handy 1923:157) and some non-East Polynesian languages (< Proto Nuclear Polynesian *sumu ‘tie, lash’). li: has cognates throughout Polynesia and beyond. The etymology of the Hawaiian verb kuʔi, however, is unclear. It is glossed ‘join, stitch, sew, splice, unite’ (Pukui and Elbert 1986). Conceivably, it could be a reflex of Proto Polynesian *tui ‘to thread pierced objects on a string; to sew’ (Biggs and Clark 1993), a very widespread root, reconstructed back to Proto Austronesian and with reflexes in nearly every Polynesian language.53 *tui is in fact reflected in Hawaiian as kui, ‘to string pierced objects; to thread as beads’. However, *tui > kuʔi would require an epenthetic glottal stop, an irregular and unusual change of form.

kia, the Hawaiian reflex of Klar and Jones’s *tia, does not appear in these texts in any context, and as noted above, it is not documented elsewhere in Hawaiian with any sense resembling ‘to sew’.

Outside Hawaiian, I found one mention of a Tahitian verb referring to lacing canoe planks. In an account of the great canoe of the legendary Hiro, by tradition the inventor of the Tahitian plank canoe, the pahī (see also Handy 1932:46), there appears the Tahitian text of a song sung by the boat builders as they lashed together the boat planks, working on opposite sides and passing the cords to each other through the drilled holes:

---

52 Koakanu, Fornander’s source here, was a professional canoe builder.

53 *tui ‘thread pierced objects on a string’ is reflected in all the major East Polynesian languages. It also has the sense ‘to sew’ in all except Rapanui, Hawaiian and perhaps Tahitian (Handy 1932:5 has Tahitian tui ‘to sew on thatch [of house]’).
The Plank Canoe of Southern California

...E tui i roto, e puputa i vaho,
E tui i vaho, e puputa i roto.
Nati hua, nati mau...

“...Thread it from inside, it comes outside,
Thread it from outside, it goes inside.
Tie it fully, tie it fast...” [Henry 1928:550]

Here is a reflex of the ubiquitous Polynesian *tui ‘sew, thread’. It is also used in this context in Maori, another Tahitic language: “tui, Lace, fasten by passing a cord through holes. Used of fitting a canoe, lashing the rauawa [washstrake], etc.” (Williams 1971).

Other verb roots referring to plank lashing and the like are Tongareva hau ‘lash together’ (Buck 1932:193) and Hawaiian hau ‘to lash’, from PPN *faʔu, used in other contexts of boat building and fitting; Maori aukaha ‘lash the bulwark to the body of a canoe’ (Williams 1971), probably cognate with the Tongan haukafa ‘to lash (a boat or canoe) with sennit’ (Churchward 1959) and Samoan sauʔafa ‘to tie with sennit’ (Pratt 1893), ‘lash canoe plank to keel’ (Krämer 1994, 2:291 after Krämer 1902, 2:253);54 Maori mimira ‘to fasten an end piece to the hull of a boat’ < mira ‘lash’ (Best 1925:73; Williams 1971); Mangaian tamnounou < mou ‘hold together’ (Mauriaiti et al. 2006, under kiri); Aitutaki (Cook Islands) tutaki ‘join’ referring to boats built from two dugout bases joined end to end (Hiroa 1927:259); and Tuamotuan faRO: ‘tighten the final lashing of a canoe’ (Stimson 1964) and Rarotongan ?ARo: ‘to lash or lace, as the sides of the canoe with sinnet’ (Savage 1964, as <aro>). The verb *fono ‘to join pieces together’ has reflexes throughout Polynesia. It is used in the specific sense ‘to attach canoe planks’ in West Polynesia and the outliers, but not in East Polynesia (Biggs and Clark 1993). Some of these verbs also refer to the lashing of booms and other parts to the body of a boat.

To sum the Polynesian evidence regarding *tia ‘to sew’: there is a respectable amount and variety of available materials in several East Polynesian languages describing the lashing on of canoe planks and end pieces. There exist several verbal roots of varying ubiquity referring to the sewing together of boat planks. *tia is not used in any of these texts, nor does it refer to sewing in general anywhere except Rapanui. The only words meaning ‘to sew’ associated with plank canoe construction are the Tahitian verb tui and the Hawaiian kuʔi in the expression waʔa kuʔi ‘sewn boat’.

When a language (attested or reconstructed) is imperfectly known, guesswork is sometimes the only available option when trying to find the form of an unknown word, in this case ‘to lash together boat planks’. But here, a great deal of vocabulary specific to that meaning has already been recorded or reconstructed, and it is less likely that yet another common synonym has somehow slipped through undetected.

As shown, all East Polynesian verb roots reflecting *tia have the primary sense of forcing a sharp implement into a material. The shift from ‘spike’ to ‘sew’ makes sense only for sewing with a needle, where the hole for the thread is created as the thread is pushed through it. That is not the case in the so-called sewing of planked boats, where a hole is first drilled, and a cord is then passed through through the existing hole, in what would be more accurately called in

54 The Maori sound correspondences are irregular. Assuming a PPN form *saukaʔa (Biggs and Clark 1993), from *sau ‘?’ + *kaʔa ‘sennit’, its regular reflexes would be Tongan haukaʔa and Samoan sauʔafa, as observed, but Maori shows aukaha, not the expected haukafa (orthographically <*haukawha>). ō-h occurs sporadically and dialectally elsewhere in Maori (Marck 2000:43-44), and in fact *kaʔa ‘sennit’ is reflected as Maori kaha, but the initial ō is irregular. Note that this putative sau is unrelated to the Tahitian, Hawaiian and Tongareva hau ‘to lash’ < PPN *faʔu.
English ‘lacing’ or ‘lashing’; that meaning is accurately reflected in the root *tui ‘to string through’, which also happens to serve as the commonest metaphor for sewing cloth and such in the Polynesian languages, and for plank lashing in a few. In other words, *tia is unattested as the term for ‘sewing’ boat planks not by accident, but because its meaning is not the right one.

9.2.2 *titia/tiatia ‘short sticks’

In their earlier papers (A, B) Klar and Jones suggested other possible meanings for *tia to fit their hypothesis. Tahitian titia is glossed in one dictionary as ‘short sticks used for fastening together the pieces of a canoe when building it’ (Davies 1851);\(^{55}\) that may be the titiʔa ‘cross pieces in a canoe’ of Andrews and Andrews (1944), from tiʔa ‘straight; across’ (<PEP *tika). In other words, the reference is not to details of hull construction, but to the assembly of the boat, cf. tiʔatiʔa ‘uprights attaching float to outrigger’ (Handy 1932:39). The form tiatia ‘small posts’, is given in POLLEX, following one vocabulary (by Frank Stimson), with no further context.

It could also be that Davies’s titia parallels the recorded Tahitian titi ‘peg’, from ti ‘peg’ (Andrews and Andrews 1944), itself a shortened form of tia, discussed above. In the context of planked boat construction, titi is a short wooden peg which is forced into the hole in the planks of a sewn canoe, in order to hold the tension in the cord passing through it (Bataille-Benguigui et al. 2008:81); this technique is practiced in the Tuamotus, Tahiti and elsewhere in East Polynesia (see e.g., Haddon and Hornell 1936:89, 107, 142); the reduplicated form titi probably signifies a diminutive (e.g., Krupa 1982:49-50). As vague as the given gloss for titia is, it clearly does not refer to short planks sewn together, as Klar and Jones suggest it is (A:475).

Even supposing titia (recorded only from Tahitian) is related to titi ‘pegs for securing plank lashings’, it would make a poor candidate for borrowing with the meaning ‘boat’. These pegs are perhaps the smallest and least conspicuous parts of a sewn plank boat. To use them to refer a planked boat would be akin to using ‘spark plug’ or ‘lug nut’ to refer to a car. Moreover, such tightening pegs were never used in the plank boats of California. The Gabrielino would not have named their new boat after a part of the boat which they did not use.

9.2.3 *tia ‘mast’

The Hawaiian kia ‘mast’ (*tia in most pre-1800 dialects) reflects two unique semantic developments, ‘stake, post’ > ‘vertical pole’ > ‘mast’, neither of which is recorded from any other East Polynesian language. It is used in metonymic compound expressions referring to particular kinds of boats, e.g., kia lua ‘brig, two-masted schooner’ (lit. ‘two masts’), but never refers to ‘boat’ by itself. There are two problems with using this word as a source for the Gabrielino loan. First, the word is a Hawaiian innovation, and involved two semantic shifts, from pole (stuck in the ground) to pole (any tall stick-like object), and then to mast. Between Proto East Polynesian and tia ‘mast’ there are the splits of Proto Central East Polynesian, Proto Marquesis, and Hawaiian, and then these two semantic shifts. These five events must have been separated from each other by some time, and each stage removes the word further from the putative time of Polynesian-Gabrielino contact.

Secondly, the boats of Southern California never had masts, since they never had sails, as was discussed and emphasized by Jones and Klar (A:469; C:766). As with titia above, it is hard

---

\(^{55}\) Davies (1851) does not mark vowel length or glottal stops. The dictionary of the Académie Tahitienne (1999) attempts to clarify the pronunciation for every item in Davies, but marks titia as having an unknown pronunciation.
to imagine why the Gabrielino would use for their new style of boat the name of an item which their boats did not have.

9.2.4 *tia- ‘type of boat’

Another Hawaiian innovation is the bound morpheme kia- (*tia in pre-nineteenth century Hawaiian) evidenced in the words for several types of boats (Pukui and Elbert 1986; Andrews and Parker 1922): kialoa–kioloa, ‘long, light and swift canoe’ (< loa ‘long’); kiapoko ‘short canoe with rounded hull, used for fishing near the shore’ (< poko ‘short’); kiapoho ‘a canoe with deep, curving hull’ (< poho ‘depression, hollow of a canoe [etc.]’); kiapa: ‘swift-sailing canoe’, and in nineteenth century Hawaiian ‘any vessel equipped with cross spars, bark’. Not much is known about these boats. Some additional details are available on the kialoa, described as a small, fast boat with low sides, seating one or two people, used for racing but especially in leading a fishing fleet (Holmes 1993:70, 123); such a small boat would have been a dugout, not a plank boat. Since even ‘a long kia’ is a small boat, it seems that kia is not a class of boat which includes large voyaging vessels and such. The morpheme kia might have come about by metonymy from the sense ‘mast’, or may have some other etymology.

Thus, Gabrielino *tiʔact is not a loanword from a Polynesian root related to Hawaiian *tia-, which describes a small dugout, entirely different from the type of boat which would have reached the Americas. And as above, this sense appears in Hawaiian only, a late development and past the appearance of the plank canoe in California.

9.2.5 *tia ‘to pierce’

Finally, Klar and Jones mention a Mangarevan word, tia ‘to pierce, bore’, which they would like to relate to the drilling of holes in boat planks (A:475; B:389). In general, as I have shown above, *tia refers to pushing through with a pointed instrument, not to drilling. The gloss ‘to pierce, bore’ is inaccurate. The only source for it, as mentioned by Klar and Jones, is Tregear’s Polynesian comparative dictionary (Tregear 1891). Tregear’s own dictionary of Mangarevan (Tregear 1899) has ‘to pierce, to stick in’, with no mention of boring or drilling. Tregear’s source for Mangarevan was a manuscript prepared decades before by French missionaries resident in Mangareva (Rensch 1991a:11). I presume the French gloss in his source was percer, which can refer to either piercing with a sharp tool or to drilling, unlike the narrower Mangarevan sense of ‘piercing’ only. Tregear’s earlier English translation appears to have mistakenly reflected the broader semantics of a French gloss.

As with the words for plank sewing, there is direct evidence for a word referring to drilling holes in canoe planks. For Hawaiian, we have houhou (Kamakau 1869 and Anonymous 1939:158-159 quoted above). hou and its reduplicated form houhou, ‘to drill’, have cognates throughout the Polynesian Languages (Biggs and Clark 1993).

9.2.6 *tia: Conclusion

Klar and Jones propose various Polynesian words, all reconstructed as *tia, each with a separate semantic route toward being borrowed as Gabrielino *tiʔact. Their most prominent one, ‘to sew’,
cannot be reconstructed to Proto East Polynesian and was certainly never used to refer to boat construction. The Mangarevan ‘to pierce’ suffers from the same flaws. The Hawaiian root for one particular boat type is a late innovation which does not refer to voyaging boats. The words for ‘short sticks’ (in Tahitian) and ‘mast’ (in Hawaiian) are late innovations which refer to features not used in Californian plank canoes. Their explanation,

...the Gabrielino named their sewn-plank boat not after the source material (as did the Chumash) but after some feature of it (short pieces of wood or a mast, cf. the Hawaiian metaphorical extension ‘mast’), or a technique associated with building it (piercing the short pieces of wood to sew them together). [B:390]

is a haphazard reach for a range of would-be etymons, each of which excludes the others, in the hope one of them would fit the desired result. As it turns out, none of them does.

9.3 Chumash tomol(o) as Polynesian

Klar and Jones claim that the reconstructed Chumashan tomol(o) can be explained as a loan from a Polynesian form, *tumuraːʔau, meaning something like ‘useful wood’, and designating the material used in constructing sewn-plank canoes. I show here that this etymology fails on semantic and phonological grounds.

Klar and Jones’s proposed etymon *tumuraːʔau is a reflex of Proto East Polynesian *tumu raʔakau, a compound of PEP *tumu and *raʔakau. *tumu is a stem reflected in many East Polynesian languages with the meanings ‘origin, source’ (in the abstract), ‘base, foundation’ (of concrete objects), ‘trunk (of tree)’, and in some languages ‘tree’, either as a stand-alone noun or in compounds denoting specific kinds of trees.

EAST POLYNESIAN

Rapanui: tumu tree trunk; origin, source of an idea.

CENTRAL EAST POLYNESIAN

MARQUESIC

Hawaiian: kumu. bottom, base, foundation; main stalk of a tree.

Mangarevan: tumu. tree trunk, stump; cause, origin, source.

Marquesan: tumu. tree trunk, stem; tree; bottom, base, foundation; the stumps of a beard [Crook et al. (1998)].

TAHITIC

Tahitian: tumu. base or trunk of a tree or plant (below the first leaves), stump, stem; root metaphorical); base, foundation, reason, cause [additional sources: Wahlroos (2002), Lemaître (1973)].

Tuamotuan: tumu. source, root, cause, origin; trunk, stump.

Maori: tumu foundation; stump, post.

Rarotongan: tumu. cause, reason; source (e.g., of a river), place of origin; foundation of a house, base of a mountain; tree stump, tree trunk, tap-root.

Tongareva: tumu. buttocks; base; cause, reason.

Reflexes of *raʔakau occur throughout Polynesian, with the meanings ‘tree’ or ‘wood’. In some languages reflexes of *raʔakau mean ‘piece of wood’, ‘stick’; in some they refer to specific
wooden implements. In several East Polynesian languages the word means ‘medication’ or ‘medicine’, through metonymy akin to the English ‘herb’. ‘Wood’ or ‘tree’ are reasonable metonyms for ‘medicine’ in the context of Polynesian medicine; of the two dozen principal plant species of the traditional Tahitian pharmacopeia, about half are trees or woody shrubs (Petard 1972). Klar and Jones attempt to somehow use this secondary meaning to interpret the semantics of the protoform *tumu raʔakau. However the two have clearly separate histories, and no such relationship is apparent.

EAST POLYNESIAN

Rapanui: raʔakau. castor bean; goods, property; medicine.

CENTRAL EAST POLYNESIAN

MARQUESIC

Hawaiian: laʔau. tree, plant, wood, timber, forest, stick, pole, rod, splinter, thicket, club; medicine; canoe end-piece (in compounds).

Mangarevan: rakau. wood; tree; medicine.

Marquesan: ʔakau, kaʔau. wood, tree, plant; type of club.

TAHITIC

Tahitian: raʔau. plant, tree, wood, timber; medicine.

Tuamotuan: rakau. medicine; tree, plant; stick, twig, piece of wood in general; log; spear, weapon.

Maori: rakau. tree; wood, timber; stick, spar, mast; weapon.

Rarotongan: ʔakau. tree, bush, plant; timber, piece of wood, stick, pole, plank, board, bat, racquet; medicine.

Mangaia: rakau. tree, bush, plant; piece of wood, stick, wooden object; medicine.

Tongareva: rakau. stick, plank, timber, wood.

The compound of the reflexes of *tumu and *raʔakau appears in Tahitian, Rarotongan, and the Marquesic languages, meaning either ‘tree trunk’ or ‘tree’. In none of these languages has this compound undergone any other semantic extension. From its distribution, the form can be reconstructed to Proto Central East Polynesian.

EAST POLYNESIAN

Rapanui: —

CENTRAL EAST POLYNESIAN

MARQUESIC

Hawaiian: kumulaʔau. tree.

Mangarevan: tumu rakau. tree trunk.

Marquesan: tumu ʔakau, tumu kaʔau (with metathesis). tree.

TAHITIC

——

57 In Hawaiian boat construction, the projecting bow and the stern pieces are sometimes made of separate pieces sewn on to a dugout base, and respectively called laʔau ihu ‘bow wood’ and laʔau hope ‘stern wood’. I have not seen either of those called just laʔau.

58 The meaning ‘medicine’ is attested in the first substantial Rapanui dictionary (Roussel 1908; Churchill 1912), but not in later wordlists. Modern Rapanui uses raʔakau only in the sense of ‘castor bean’ (Fuentes 1960; Arredondo 1988; Blixen 1972). The Tahitian form raʔau appears in loanwords (Arredondo 1988).
Tahitian: *tumu raʔakau. tree.
Tuamotuan: —
Maori: —
Rarotongan, Mangaia: *tumu rəkau. tree; trunk, log.
Tongareva: —

Reflexes of *tumu raʔakau never mean ‘wood’, only ‘tree trunk’ and hence ‘tree’. It appears that the compound is used specifically to avoid the ambiguity between ‘wood’ and ‘tree’, and in some languages it has largely replaced the reflex of *raʔakau as the word for ‘tree’. That argues that the word never meant ‘wood’, as Klar and Jones would have it, even if it was present in Proto East Polynesian.

Klar and Jones are not precise in describing the semantic path from Polynesian *tumu raʔakau to Chumashan *tomolo. By their original proposal (A:474, 476; B:384), a reflex of *tumu raʔakau meant ‘wood’, ‘source of wood’, or ‘economically useful tree’ in the Polynesian donor language, and was borrowed into Chumash with the sense of ‘tree for making boat planks’, while at the same time expanding its Chumashan meaning through metonymy to signify ‘planked boat’; elsewhere, they suggest the Polynesian word for ‘wood’ was misunderstood by the Chumash to have meant ‘boat’ (A:476). In later papers (B:397; D:89; E:175), they advocate a more complex scenario within Chumash from ‘wood’ to ‘planked boat’. In all these variations, the Polynesian *tumu raʔakau is taken to have meant ‘wood’ or ‘wood source’.

But as shown here, no reflex of *tumu raʔakau ever meant ‘wood’. Nor would the Chumash borrow a word for ‘wood’, to them a familiar and much-utilized material, for which the word has been reconstructed back to Proto-Chumashan *pono’, and is reflected in all its daughter languages (Klar 1977:115-116).

Semantically, the closest one can come to rescuing Polynesian *tumu raʔakau as a source for Chumashan tomolo is by positing that the Polynesian word ‘tree’ was borrowed into Chumashan with the sense ‘yellow pine’, a tree favored for boat construction, which then became the word for plank boat (this semantic path has not been suggested by Klar and Jones.) The metonymic extension from ‘yellow pine’ to ‘boat’, which I discuss at length below, is certain. However, the Chumash would not have been likely to borrow a Polynesian word to replace the name of a tree long familiar to them, after a very brief and superficial encounter.

Phonologically, Jones and Klar’s suggested Polynesian source for the borrowing, *tumu raʔau, does not fit what is known of Polynesian historical phonology. To get from PPN *tumu raʔakau to *tumu raʔau, two shifts are required, namely *ʔ>∅ followed by *k>ʔ, as discussed in section 7. The change *ʔ>∅ can be reconstructed as having affected PCE, since no Central East Polynesian language has preserved PEP *ʔ. But the change *k>ʔ occurred separately and later in several Central East Polynesian languages and cannot be reconstructed to any older subgroup.

As I have shown above, the Chumash tomol came to exist, at the latest, just as people were entering East Polynesia, and therefore speaking Proto East Polynesian (not PCEP, as Klar and Jones suggest at times). The split of PEP into Rapanui and PCEP would have yet to occur, the PCE shift to *ʔ>∅ would be later still, and the shift from *k to ? would occur in some daughter languages even later. Even if some unattested dialect of PEP existed in which *ʔ had already been lost, the subsequent *k>ʔ shift would still be unexplained. Klar and Jones refer to this phonological mismatch (B:384) and propose, ad hoc, such an unattested Polynesian dialect. But to acknowledge the earlier sound change, they propose that this would be a dialect of PCEP, which would have to be considerably later than the date of contact.
Put another way, if Polynesians had reached the California coast at the end of the first millennium AD, the only time at which the chronologies of both the earliest *tomol* and the settlement of East Polynesia might overlap, their phonetic form of the word in question, if they had it, would then be *tumu ra?akau*. I won’t attempt to predict the exact form this word would evolve to in the Chumashan languages by AD 1800, but it is certain that the *k* would persist, since Proto Chumashan *k* is reflected as *k* in all Southern Chumashan languages (Klar 1977:11-13, 29). Since the word *tomol(o)* and its reconstructed protoforms contain no *k*, it cannot have come by that route.

To summarize, *tomol(o)* is not a Polynesian loan. There is no evidence that a reflex of the Polynesian protoform which Klar and Jones propose existed when the *tomol* was invented. If it did, it did not sound as they claim it did, did not mean what they claim it did, and would not have been borrowed as they suggest.  

9.4 Discussion: The Linguistic Case for Polynesian Contact

Klar and Jones’s linguistic argument for Polynesian-Chumash contact is not supported. The three Polynesian sources proposed would be reconstructed as *tara?ia* ‘carving’, *tia* ‘spike’ and related meanings, and *tumu ra?akau* ‘tree trunk’ in the Polynesian language spoken at the time of the proposed Californian contact. None of them are semantically plausible as sources for a term for boat in a borrowing language, as detailed above, and none of them are reflected in boat-building terminology anywhere in East Polynesia. Two of the words, *tara?ia* and *tumu ra?akau*, are also excluded for phonological reasons as the sources of the Gabrielino *tarainxa* ‘tule boat’ and Chumash *tomol* ~ *tomolo* ‘plank boat’, and neither of the two can be securely reconstructed to proto East Polynesian.

While each of these claimed etymologies fails to stand on each its own, the combination of all three is even less plausible. This combination would require a scenario in which a brief encounter between the Polynesian visitors and the indigenous Americans would result in no less than three different borrowed terms, one into Chumash and two separate ones into Gabrielino, with each requiring a separate set of implausible semantic and phonological changes. As Anderson (2006:759) points out, the Chumash would be more likely to have borrowed directly a reflex of the near-universal Proto Polynesian *waka*, ‘boat’.  

On closer look, the case for Polynesian-Chumash contact turns out to have no foundation in archaeology, linguistics or boat technology. While it remains possible that prehistoric Polynesians reached North America, there remains no evidence to prefer placing this contact in Southern California over any other location between the Aleutian Islands and Colombia. Any

59 Terry Hunt (quoted in Lawler 2010:1347) proposes that *tomol(o)* might have been a late borrowing into Chumashan, through one of the many Polynesians recruited as sailors aboard European ships in the late 18th and early 19th centuries. Aside from the same semantic objections which apply to an old borrowing, tomol ‘boat’ was recorded in 1769 (as mentioned above), only two years after the first European contact with Tahiti and a decade before Cook reached Hawaii, and long before any East Polynesian sailors were so recruited.

60 Klar and Jones (A:476; B:392-393; D:766) argue that the Chumash would not have borrowed the generic word for ‘boat’, as they doubtlessly already had terms for boats in general and for specific types of boats. However, American English has borrowed from Mexican Spanish the generic terms *sombrero* ‘hat’ and *salsa* ‘sauce’ to refer to more specific meanings associated with Mexican culture. The Chumash could likewise have borrowed a generic Polynesian term to refer to a Polynesian type of boat.
such contact remains in the realm of pure speculation.

10 Local Etymologies

I have shown above that the proposed Polynesian etymologies do not hold, which reopens the question of the linguistic history of the Gabrielino and Chumashan words under discussion. In the etymological study of any language, one encounters words resistant to historical analysis. This is true even for languages such as English, with its wealth of documentation, historical materials from the language itself and from related languages, and centuries of intensive study. Certainly many opaque etymologies will be expected for a poorly documented language such as Gabrielino, or in members of an isolated language family such as Chumashan. Therefore, the lack of an internal etymology for a Chumashan word does not strengthen the case for an external origin, as Klar and Jones (A:473, B:381) argue for \textit{tomol(o)}. As it turns out, I will show here that all three words in question — Gabrielino \textit{tarainxa} and \textit{tiʔat}, and Chumashan \textit{tomol(o)} — have discernible linguistic histories, which do not involve borrowings from other languages.

10.1 A Local Etymology for \textit{tarainxa}

Besides the early wordlists, nearly all the extant data on the Gabrielino language comes from the field notes of John Peabody Harrington, who worked with some of the last speakers of the language and collected a great deal of phonetically precise and linguistically sophisticated lexical and grammatical data. Harrington made an effort to re-elicit the earlier wordlists of Hale and Taylor. After several failed attempts to re-elicit Hale’s or Taylor’s forms for ‘boat’ from his informants (Harrington 1986, 3:103:632, 747), one informant (perhaps Jesús Jauro) produced the following:

\begin{itemize}
\item \textit{g. [Gabrielino] taráaynxa’} está horqueteado, tiene horqueta, e.g., está horqueteado el palo, tiene horqueta el palo. [is forked, has a fork, e.g., the stick is forked, the stick has a fork]
\item \textit{g. taráayn’exáa}, abre las piernas, spread your legs.
\item \textit{Tr[anslate]s. una horqueta: pokúu’ taráaynxa’},
\item e.g., the palo de sauco horqueteado [forked stick of elder] used by game players.
\item \textit{N[escit] “tarayna” [Harrington 1986, 3:103:112. Comments in brackets are mine. Gabrielino forms are bolded for clarity]}
\end{itemize}

There seem to be two forms here, the adjectival \textit{ta-rajnxa’} ‘forked’ and the imperative verb \textit{ta-rajn?exa}: ‘be forked (imperative)’. Both forms are based on a stem \textit{ta-ra}: ‘fork’. In the neighboring Takic language Kitanemuk there exists a nominal root \textit{tara-} ‘fork’ (Anderton 1988:516), as in \textit{tara-hu’t}, ‘cradleboard’, lit. ‘forked willow’, and \textit{tara-kaʔj} ‘cloven (said of deer’s hoof)’, ‘forked pole’. In the closely related Serrano, the word for ‘forked stick’ is \textit{tara’qa’}, also meaning ‘cross’ (Ramón and Elliott 2000:58, 59, 149, 553). In Kitanemuk, the derivational suffix \textit{-kaʔj} produces from a nominal stem \textit{X} another nominal or adjectival stem meaning ‘characterized by \textit{X}’ (Anderton 1988:143), in this case ‘fork’ > ‘forked, forked object’. The Serrano example appears to exactly parallel the Kitanemuk one. Possibly the Luiseño placename \textit{taráxa}, originally referring to Saddleback Mountain (Elliott 1999), is related. I thus tentatively parse the Gabrielino form \textit{ta-rajnxa’} as \textit{ta-ra-i-n-xa’}, \{fork-?-?-CHARACTERISTIC\}, with the stem and final suffix corresponding exactly to the Kitanemuk forms. The medial morpheme or
morphemes are obscure to me at present. Merriam’s Gabrielino wordlist also includes the following pairs:

- <O-hā’> ‘sand’ : <O-hā’ting-ah> ‘sandy’
- <Kwe-nar> ‘mud’ : <Kwe-nah-ting-ah> ‘muddy’
- <To-tah> ‘rock’ : <To-to-ting-ah> ‘rocky’ [Merriam 1903a]

All three pairs show a pattern of deriving an adjective from a corresponding noun by adding a suffix, transcribed as -<ngah> or -<ingah>, to the stem. If these can be read as -<(i)-nxa/-nxa, they would provide further examples of the adjectivizing morpheme in tarainxa. In Merriam’s orthography these morphemes would normally be read as -ŋa/-ŋa, but the locative suffix -ŋa does not easily fit here semantically, unless, perhaps, ‘sandy’, actually refers to ‘a sandy place’, etc. I cannot at this point easily choose one of these interpretations over the other.

As mentioned in section 8.2, Merriam (1903a) records the Gabrielino form <Hoo-pā’-kah (a point) tar-ri’n-hah> ‘A kind of pointed instrument’, following a list of other tools (but separate from the list of musical instruments), and across the page from the term for ‘tule boat’. <Hoo-pā’-kah> is also glossed elsewhere as ‘sharp’ and as ‘stone-pointed arrow’. I read this as hu’pakta ta’rainxaxa, lit. ‘split point’. This might be a a tool like a fork, or it might refer to a musical instrument, the split-stick clapper (Wallace 1978b:644). Harrington’s po’ku’ta’rainxaxa (po’ku:7‘one’, translating Spanish una) may refer to a game piece, or again to a clapper, as used for musical accompaniment during gambling games.

Harrington’s verb ta’rainxaxaʔ, the imperative ‘be split’, is distinct from ta’rainxaxaʔ; its final morpheme derives from xa: ‘be, have’. It does not relate directly to the issues here.

Although the full parsing of the form is not fully clear, the meaning ‘forked’, is unambiguous. In Harrington’s examples the word is used as a noun as well as an adjective, i.e., ‘forked’, ‘forked thing’, as with the Serrano and Kitanemuk cognates. It would be very unusual for such a long phonetic form, ta’rainxaxaʔ, to represent two unrelated homonyms, ‘forked thing’ and ‘boat’. I propose here that the forms given by Hale, Taylor and Merriam for ‘boat’ are in fact the same word as Harrington’s form for ‘forked’. In the three-bundle and five-bundle tule canoes of Southern California and elsewhere, each side was formed of a long bundle of tule reeds tied together, with the bundles pushed apart over the keel bundle in the middle; the ends of two side bundles are tied to each other at either end of the boat, giving the appearance at each end of a single bundle forking into two thinner side bundles; hence ta’rainxaxaʔ ‘forked thing’ refers to a tule-bundle boat, as glossed by Merriam. When the early vocabularies were collected, the old plank boats and their name had been mostly forgotten, but Gabrielino people, even those living near the San Gabriel mission, far away from the coast, were still familiar with tule boats and some knew their name.

With a native etymology for this word, an explanation based on borrowing from another language is no longer necessary. Certainly the word cannot be considered ‘anomalous’, as Klar and Jones put it (A:475; B:390, 396; D:90; E:176), at least not in the usual sense of a odd phonotactics or opaque morphology, which often characterize borrowings.

10.2 A Local Etymology for tiʔat

Here I propose an etymology for the Gabrielino word for tiʔat, ‘planked boat’, though I consider it less certain than that of tarainxa. I propose that it comes from an otherwise unattested
Gabrielino verb stem, \(ti\?'a\)- ‘to stack up’, and that the word literally means ‘stacked up thing’.

Luiseño has a verb stem, \(te\?'a\)- (intransitive form), glossed ‘be supported, be braced; get stuck; be stacked up’ (Elliott 1999), and ‘exert opposing force, parry a blow, prop up a structure’ (Bright 1968). It is used to refer to wood stacked up in a fire. The transitive form, \(te\?'i\)-, glossed ‘to support, brace’, also refers to a bird building its nest (Bright 1968); this verb underlies the noun \(te\?'a\)-iʃ ‘bird’s nest’.

Hill and Nolasquez (1973) and Hill (2005:199) list the Cupeño place name \(a\ʂwət pə\?ti\?'a\ ‘Eagle’s nest, place where baby eagles where captured’, which they gloss \(a\ʂwə-t pə\?'ti\?'a, {eagle-NPN 3SG-roost-PSD}, and a corresponding item, \(ti\?'a- ‘to roost, of birds (class VI verb)’. I propose that Cupeño verb \(ti\?'a- is cognate with the Luiseño one, and means not ‘to roost’ (as in Hill) but ‘to build by stacking up’ (as in Luiseño). I reconstruct the Proto-Cupan form *\(ti\?'a-, since Luiseño e often corresponds to Proto Cupan *i (Bright and Hill 1967:115).

This sense, of a structure made of smaller things stacked holding each other up, fits a boat whose sides are made of rows of attached small planks set atop each other like rows of bricks. I do note that a plank boat is different from pile of firewood or a bird’s nest, in that the pieces do not cross each other, and the structure is not held together by gravity alone.

Based on the limited records for the language, Gabrielino always lengthens the second vowel of CVCV verb stems. Other, similar lengthening processes occur in the language for other verb stem shapes and for non-verbal stems, under more complex conditions. Consider the following examples from Munro (2000), based on Harrington’s notes, with some Takic cognates (K. = Kitanemuk, Anderton 1988; L. = Luiseño, Elliott 1999):

- kwaʔa: ‘eat’ K. kwaʔ, L. qwaʔ
- maxa: ‘give’ K. mak, L. maˈxani (transitive)
- pefə: ‘leave’
- moka: ‘kill’ K. mik, L. ˈmokna ‘kill’
- jake: ‘dance’ L. ˈjaki ‘[kind of dance]’
- čeʔe: ‘sing’
- jari: ‘remain’
- kovi: ‘be hungry’
- pako: ‘enter’
- hoho: ‘fart’ K. hahuʔ?
- koko: ‘kill’

And so, a Gabrielino cognate of Proto Cupan *\(ti\?'a- ‘stack up’ would surface as \(ti\?'a-. It remains to be shown that a nominal form can be constructed from this verb stem by adding the absolutive suffix -t, with the derived noun taking on the sense of the patient or object of the verb.

The most common way of deriving an unpossessed patient noun from a verb stem in the Takic languages is by adding the absolutive suffix -iʃ~iʃ~iʃ, for example, Luiseño \(te\?'-iʃ ‘bird’s nest’ mentioned above, Kitanemuk mak-iʃ ‘gift’ < mak ‘give’ (Anderton 1988), or Gabrielino kwaʔ-iʃ ‘food’ < kwaʔa: ‘eat’ (Merriam 1903a). However, there exists in some Takic

---

61 The gloss {eagle-NPN 3SG-roost-PSD} is following Hill (2005:199). NPN is a non-possessed noun suffix, the so-called Uto-Aztecan absolutive suffix. PSD is a suffix marking a noun as possessed. An alternate analysis would be \(ašwa-t pə-\?ti\?'a, with \(ti\?'a a stem which does not take a possessive suffix (Hill 2005:170-171).
languages a more marginal derivational pattern involves adding the absolutive suffix -t. In Kitanemuk, what Anderton (1988:150, 692-697) calls the ‘general nominal’ is formed by suffixing -at~ats to an unpossessed verb stem, e.g., *mayhaʔ ‘give birth’ > *mayha-t ‘child’; *punitaʔ ‘play game’ > *punita-t ‘game’. In Cupeño (Hill 2005:298-299), an unpossessed noun can be formed from a transitive verb stem by adding the absolutive suffix -t~ət, this derived noun may the sense of a patient of the verb.

The Cupeño and Kitanemuk evidence makes it plausible that the -t nominalizer existed throughout Takic, including Gabriélinó. The Cupeño semantics do not exactly match those expected for Gabriélinó tiʔac-t, but the semantics of nominalization in Gabriélinó may differ. There may be enough material in the notes of Merriam and Harrington to clarify this issue in future work. At present, the possibility remains that Gabriélinó tiʔac has the etymology ‘stacked-up thing’, from an otherwise unrecorded root but with recorded cognates in Luiseño and Cupeño. In any case, as with *taʔrainxeaʔ above, the word cannot be considered ‘odd’ as Klar and Jones suggest (B:389), at least in the sense of morphological shape or phonotactics.

10.3 A Local Etymology for *tomol(o)

As Klar and Jones themselves have mentioned (A:476; B:376-378, 392, 397), in some Chumashan languages the word for ‘planked boat’ is formally identical to that meaning ‘yellow pine’. This has been noted in print at least as early as Kroeber (1910:268 n.3), and later by Heizer (1941a:60-61), Heizer and Massey (1953:298), and Hudson et al. (1978:23). Klar and Jones (B:376-378) present in detail the recorded Chumash forms. In their interpretation, *tomol ‘plank boat’ is a later semantic development from *tomol ‘yellow pine’ (for which they apparently advocate a Polynesian origin), with the sense of ‘wood suitable for plank boat making’. They contrast *tomol ‘yellow pine’ with Chumash terms for other types of pines.

I argue here that *tomol ‘yellow pine’ is an old Chumash word, of unknown etymology, and is the source for the term for plank boat, as was already suggested by Heizer (1941a), Hudson et al. (1978: 22-23) and perhaps others.

The following sources attest the Chumashan *tomol ‘pine’. The list is not exhaustive, and

---

62 Hill sees an aspectual distinction between the Cupeño -its nominalizer and the -ot nominalizer, following the earlier analysis of Jacobs (1975:71). In Hill’s analysis, the -ot deverbal nouns refer to the objects of verbs describing actions in the immediate past, or ones still relevant within the current discourse context. This does not seem to apply in the Kitanemuk examples, nor, I speculate, in Gabriélinó.

63 Klar and Jones’s comment about tiʔat being ‘odd’ is attributed to a personal communication with Pamela Munro. They never clarify whether that refers in fact to the word itself. Elsewhere (A:475, B:390) they quote Munro as saying that “tiʔat is somewhat odd in having no stress or length marked”, which refers to the transcription, not the word itself. This comment of Munro’s may be based on the secondary source (Hudson et al. 1978) used by Klar and Jones, which quotes Harrington but omits his stress and length marks.

64 In western U.S. English, the term ‘yellow pine’ encompasses ponderosa Pine (*Pinus ponderosa*) and Jeffrey Pine (*Pinus jeffreyi*), two species of closely similar appearance, lumber quality and geographic distribution within California.

65 Klar and Jones (B:376) mention Kroeber (1910) as a source for *tomol ‘pine’; Kroeber’s sources are Taylor and Gatschet, as given here. They also mention Harrington as another source for Island Chumash *tmolo ‘pine*. Harrington only recorded *tmolo ‘boat* when reeliciting Taylor’s wordlist. His informant, Fernando Librado, speculates that “since make cayucos [canoes] out of pine, might call the latter also *tmolo*” [Harrington 1986,
more instances probably exist in Harrington’s notes:

\[
\begin{align*}
'tomol' & \quad 'pino’ & \quad \text{Purisimeño [<Tómol>, Arroyo de la Cuesta 1837:8].} \\
tomol & \quad 'pine-tree’ & \quad \text{Barbareño or Cruzeño. The informant was born on Santa Cruz island ca. 1781 but was baptized and lived at the Santa Barbara mission [<tomol>, Taylor 1860b].} \\
'tomol’ & \quad 'Pine’ & \quad \text{Barbareño (Kaswa, near the mission) [<to'-molgh>, Gatschet 1879:444].} \\
tomol & \quad 'pine’ & \quad \text{Barbareño. “The kind of pine we see on top of the ridge here are called tomol — that is what the people called it.” [<tomol>}, \text{ Luisa Ignacio, informant; Harrington 1986, 3:20:283, recorded 1913-14].} \\
tomto’mol & \quad 'pines’ & \quad \text{Barbareño [Juan de Jesús Justo, informant; Blackburn 1975:209].}
\end{align*}
\]

The form given by Juan de Jesús Justo displays a glottalized \( m \), as do some of the Barbareño variants for the homonym signifying ‘boat’. This further reinforces the identity of the two forms, especially in light of Klar and Jones’s reconstruction of the earlier form. based on the presence of this glottalization, discussed in section 8.4.

The identification of \( tomol \) specifically as ‘yellow pine’ comes from Harrington, who distinguishes it from \( tak \), a class of trees including all other pines of the area (Hudson et al. 1978:48 n. 3), and less suitable for plank boats. This seems reasonable, but remains to be verified.

Of the sources here, Arroyo de la Cuesta’s form is in Purisimeño, which was spoken well outside the range of yellow pines, but the word could have been borrowed from the neighboring Barbareño or Ineseño. Taylor’s informant gives the forms for both ‘pine’, \( <\text{tomol}> \), and ‘canoe, boat’, \( <\text{tomolo}> \). The different forms in this doublet might be due to free phonological variation, but more likely, the informant used the Cruzeño form for ‘boat’, but had borrowed a Barbareño term for ‘yellow pine’, since no pines grow in the Channel Islands except a stunted form of the Torrey Pine (\( \text{Pinus torreyana} \)), now restricted to parts of Santa Rosa Island (Griffin and Critchfield 1972). Luisa Ignacio’s comment, if recorded near Santa Barbara, would refer to the top of the Santa Ynez mountains, which are too low to support yellow pines; perhaps her identification was inaccurate.

It appears that \( tomol \) can only be demonstrated to mean ‘pine’ in Barbareño. Although terms for various types of pines have been recorded from Ventureño and Ineseño speakers (Klar and

---

3:68:31] and “since cayucos made of pine, might perhaps call a pine tree \( \text{t'molo} \)” [ibid., 3:68:196]. These appear to be ad hoc etymologies, fashioned in response to Harrington’s inquiry. Jones and Klar (2012:221) read this as saying that “the primary meaning of \( tomol \) was not pine,” but Librado is clearly non-committal, and would not have known the early history of the word.

66 Taylor gives the informant’s name as Joseph Camuluyaset. In mission records he appears as José Crespin Camuluyatset, born in Liam [Santa Cruz Island], and baptized in Santa Barbara in 1819 at age 38 (Huntington Library Early California Population Project, http://www.huntington.org/Information/ECPPmain.htm ).

67 In Gatschet’s orthography \( <\text{gh}> \) is ‘a sonant guttural aspirate (Arabic ghain)’ [1879:423]. Oscar Loew, the collector, probably heard \( to’mol \) or perhaps \( tomol \), compare his \( <\text{ulgh}> \) ‘foot’ to \( \text{ʔiʔ} \) (Whistler 1980).
Jones, B:376), I know of no record of tomol as a term for pine in these languages, nor any other term specifically for ‘yellow pine’. Whether a cognate existed in Ventureño or Ineseño is unknown.

Drift logs of redwood were the preferred material for canoe planks, but tomol, or yellow pine, was considered a workable material as well (Hudson et al. 1978: 46-50). Could tomol pine have taken its name from the word for the plank canoe?

In English, for example, there are many artifacts named for the material from which they were manufactured: ‘glass’, ‘wood’, ‘box’, ‘broom’, ‘iron’, ‘silver’, ‘cloth’, ‘clothes’, ‘marble’, ‘horn’, and many others. On the other hand, I know of no clear examples of materials named after items made of them. A simple explanation is that in general a material is known, and thus has a name, long before any particular use is made of it: glass (the material) was known and named before glass (the vessel), and so on.68

In the specific case of trees and woods named after artifacts, a search in a large English dictionary for ‘-boat’ and ‘-wood’ compounds yields only ‘bowwood’, ‘buttonwood’, ‘greasewood’, ‘coachwood’, ‘lancewood’, ‘leatherwood’ and ‘bottle tree’. These are all species exotic to England, and became known to English speakers together with the use to which they were put. In no case was an older name for a native tree replaced with the name of an artifact or a compound based on an artifact.

In the languages of California, I know of three examples of trees named for their use as sources of boat lumber; all are from far northern California. In Karuk, the word for ‘redwood’ is ʔuθkanpahip, lit. ‘oceanward canoe tree’ (ʔuθ-kan-pah-ʔiːp {ocean-LOC-canoe-tree}, Bright 1957). In nearby Chimariko, the word for ‘redwood’ is mutumana, ‘canoe plant’ (mutu-mu {canoe-PLANT}), Dixon 1910:314). These examples parallel the English ones: redwood trees do not grow in Karuk or Chimariko territories, and dugout canoes, when used by these peoples, were obtained from the Yurok who lived in the redwood belt to the west (Davis 1961). As in the English examples, the words are compounds, not straight metonyms, and refer to an exotic wood known primarily for one use. In Klamath, the suffix -ʔm is used to derive wonđʔm ‘fir species used for canoe making’ from wonđ ‘canoe’. Here the species is not exotic to Klamath territory, but the word relies on a derivational process which is productive in Klamath (cf. ʔamda ‘digging stick’, ʔamdalam ‘mountain mahogany’), but very rare elsewhere.

Yellow pines were undoubtedly familiar to the Chumash since well before the advent of the plank canoe. Yellow pines grow at high elevations in the mountains north of Santa Barbara and Ventura, within Chumash territory. Their range overlaps much of that of the piñon pine (Pinus monophylla), whose seeds, i.e., pine nuts, were a significant food source of the Chumash and other California Indians (Grant 1978:516; Timbrook 2007:142-146; Barrows 1900:63); pine pitch, too, was collected in the mountains on pine-nut collecting trips (Hudson et al. 1978:52). Yellow pines must have had a Chumash name before plank canoes came to be, and this name would not have been replaced by the word for ‘plank canoe’, especially as pine was secondary to salvaged redwood in plank canoe construction (Hudson et al. 1978:46-50). This parallels the earlier argument against the replacement of the word for ‘yellow pine’ by a Polynesian borrowing.

Klar and Jones attempt to bolster their case for a Polynesian source of the word tomol(o) by

68 In Chumashan languages one finds the sets po’n ‘tree’, ‘wood’ and ‘board’, ‘plank’, ‘stick’ (Whistler 1980); pox ‘agave’ and ‘agave fiber cord’ (Hudson & Blackburn 1982:90); ʔaxpiʔil ‘nerve, sinew, muscle’ and ‘bowstring’ (Whistler 1980). I believe that here as well the materials preceded the objects.
arguing that the word (or its suggested protoform *tomolo) is unusual in being so long without being a transparent compound (B:381, 385). I will not attempt to offer an etymology of tomol(o) ‘pine’, but a counterexample is Barbareño toq’olo ‘armpit’, with the same phonotactic shape as tomodolo or its variant to’molo. toq’olo is morphologically opaque, and as a term for a body part is not likely to have been borrowed from another language. Perhaps toq’olo or tomodolo ‘pine’ will some day find full etymological explanations.

10.4 Pine Trees and Wooden Boats in Southern California

Chumash is not the only language in the area where words for ‘pine’ and ‘boat’ are formally identical. Pine/boat homonymy exists in Luiseño, some 200 km south of Chumash territory, and in Kitanemuk, adjacent to Chumash territory in the interior. These two Takic languages are the only languages known to me, in California or elsewhere, in which ‘boat’ and ‘pine’ (or any wood for that matter) are homonyms.69 The rarity of this semantic equation and the close proximity of the languages which exhibit it imply a historical connection.

The following are the lexical items under discussion:

<table>
<thead>
<tr>
<th>Language</th>
<th>Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luiseño</td>
<td>wi’xe-t</td>
<td>‘tree sp. (pine?)’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘tule boat’</td>
</tr>
<tr>
<td>Luiseño</td>
<td>‘pawxi-t</td>
<td>‘yellow pine’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘wooden boat’</td>
</tr>
<tr>
<td>Serrano</td>
<td>wiakt</td>
<td>‘sugar pine’</td>
</tr>
<tr>
<td>Kitanemuk</td>
<td>kwi’akt</td>
<td>‘tule boat’</td>
</tr>
<tr>
<td>Proto Cupan</td>
<td>*weket ~ *wexet</td>
<td>‘pine sp.’</td>
</tr>
<tr>
<td>Kitanemuk</td>
<td>kwekt</td>
<td>‘tule boat’</td>
</tr>
</tbody>
</table>

I derive all of the above, ultimately, from Proto Uto-Aztecan *wokon- ‘pine’ or ‘ponderosa pine’ (Fowler 1983:248), and relate the parallel meanings to the Chumashan ones, through contact.70

10.4.1 Takic Terms for ‘pine’

The following lists all the Takic reflexes of Proto Uto-Aztecan *wokon- ‘pine’, which vary

69 I have done a cursory survey of words for ‘boat’ or ‘canoe’ in standard dictionaries of several California languages. Often the word is a short morpheme with no obvious etymology (Klamath, Shasta, Yurok, Karuk, Achomawi, Yokuts [several varieties]). In a few languages the word has a clear morphology based on a verbal stem: Wiyot, lit. ‘it comes’; Hupa, ‘they travel in it’; Wintu, ‘travels by water’; Maidu, ‘bridge which floats things’; Salinan, ‘travels by water’ (Harrington 1986, 2:84:233); Mesa Grande Diegueño, ‘that which floats’. In a number of others the word recorded is a Spanish or English loan word (Yuki, Nisenan, N. Sierra Miwok, S. Sierra Miwok, Plains Miwok, Tümpisa Shoshone, Chemehuevi), though other words may turn up in older materials. The Gabrielino etymologies suggested in this paper, ‘forked thing’ and ‘stacked-up thing’, are in a class of their own, and so are the Chumash, Luiseño and Kitanemuk ones based on ‘pine’.

70 Manaster Ramer (1993) reconstructs PUA *wokon- ‘pine’; earlier authors reconstruct *woko-.
formally and semantically within the family. These almost never refer to the piñon pines (*P. monophylla, *P. edulis*), PNUA *tiba-* (Fowler 1983:237), which form an unambiguously separate lexical category in the NUA languages.  

Direct reflexes of Proto Takic *woko-:
- Cahuilla 'wexe-t ‘pine’ (desert dialect), ‘ponderosa pine’ (mountain dialect)
- Luiseño *wi’xe-t ‘tree sp.’, *we’xe’-ta ‘pine sp.’
- Gabrielino *wexa-t ‘pine’ (a loanword?)
- Kitanemuk *wokoh-t ‘gray pine (*P. sabiniana*)’
- Serrano *wo’xo’h-t (?) ‘pine sp.’

A reflex of Proto Takic *woko-, of unclear provenance, and its loans:
- Serrano *wiako-t~ ’wiak-t~wiak-t ‘sugar pine, Coulter pine (*P. coulteri*)’
- Cupeño *wiçaka-t ‘piñon pine’
- Gabrielino *wi’faxar~ *we’faxar ‘pine’
- Kawaiisu (Numic, not Takic) *wijahaka-ti-bi ‘sugar pine’

Loans within Takic:
- Gabrielino *woxo-t ‘gray pine’ < Kitanemuk *wokoh-t ‘gray pine’
- Luiseño compounds based on *wi’xe-t:
  - Luiseño *pa-wxi-t ‘yellow pine’

---

71 The following are the sources for this list; the forms are transcribed into standard orthography, except where they are phonetically imprecise: Desert Cahuilla — *wexet ‘pine’, Seiler and Hioki (1979); <Wë-ah-tut> ‘ponderosa pine’ (Merriam 1907a); Mountain Cahuilla — <Wë-eh->, <Wë-ah-> ‘ponderosa pine’ (Merriam 1910); Cupeño — *wo’xiti-t ‘pine’, Hill (2005:472), Hill and Nolasquez (1973); <We’-chuk’-ket po-wel-’lah> ‘[the base of] piñon pine’, Merriam (1933), analyzed as in Hill (2005:191); Luiseño — *pawxit and *wi’xe’ut, Elliott (1999); *wixet and *pawxit are identified as unspecified kinds of tree in Hyde and Elliott (1994:90); *wi’xenivfla, Bright (1968); *wexe’mefl (Soboba dialect), Harrington (3:103:650); Gabrielino — *wach-o’t ‘gray pine’, Merriam (1903b); *ush-’a-gar (Gatschet 1879), *wish-ye-arker (Taylor 1860a), *we’faxar (Harrington 1986, 3:103:650), *wi’faxar (Harrington 1986, 3:102:626); <wëxat ‘pine’ (Galloway 1978, probably after Harrington); Kitanemuk — Anderton (1988), specifically *wo’koh (after Harrington) or *wo’koh (after Zigmond) ‘gray pine’; *Waw’-kot ‘gray pine’ (Merriam 1903e); Serrano — *Wi’-yahkt ‘sugar pine’, *We’-ahkt ‘sugar pine’, *We’-ahkt ‘Coulter pine’, Merriam (1907b,c); <wo’sa’t ‘Has edible seeds. a smaller pine sp., lots at Big Bear’ (Harrington 1986, 3:101:23); Munro (1977:312) quotes Donald Crook as saying Serrano stress tends to fall on initial and long vowels, but is generally lexically determined; but Ramón and Elliott (2000:xxxiv) state that Serrano has no word-level stress. Kawaiisu — *wijahaka-ti-bi ‘sugar pine’, Zigmond (1981:50); Proto Cupan — Bright and Hill *we’xet (1967:183); the justification for the form *we’ket is given here in the text.

The word *we’xen ‘another kind [of pine]’ appears in Harrington (1986, 3:103:650), among several Luiseño (*’Reyano*) words, which themselves appear within the Gabrielino notes. From the context *we’xen appears to be Luiseño, not Gabrielino. In addition, the glottalization and the absolutive *-ta are characteristic of Luiseño, and the stressed syllable xe is not long, as would be typical for Gabrielino.

72 Harrington’s <wo’sa’t> is irregular either in the language or in its transcription. Harrington usually uses <o> to mark the sound [o] (Anderton 1991) but a root *wixi- or such cannot be related to the protoform *woko- by any known Takic sound changes. On the other hand wo’xo’h would be a straightforward cognate for Kitanemuk wokoht, the recorded regular reflex, but that would have Harrington transcribing Serrano o’ as <o> instead of his customary <or>.
Luiseño wi’xe-nivifla, we’xe-mevefla ‘pine sp.’
Derived form of Proto Cupan *we’ke-* *we’xe-:
  Cahuilla ‘wexatu-t ‘ponderosa pine’ (desert dialect)
  Luiseño wi’xe’tu-t ‘sugar pine, Coulter pine’
  Cupeño wə’xiti-t ‘pine’

Some of these etymologies need further comment. Cahuilla ‘wexe-t, a regular reflex of the Proto Takic, is semantically broad, and possibly reflects an earlier situation in the Takic languages, where the mountain flora were not yet familiar in detail; this may be the situation in Luiseño as well, though details on the semantics of Luiseño wixe-t are lacking. In Serrano, the word was specialized to refer to sugar pines, and to the similar Coulter pines; and in Luiseño, this root was used to form a new stem to refer to the yellow pines.

Forms deriving from *wiakət and its loans appear in Serrano, Gabrielino, Cupeño and Kawaiisu. The Gabrielino and Cupeño forms show strengthening of the glide ia, to j and ç respectively. Serrano shows variable syncope of the last vowel, which regularly occurs following a stressed syllable (Hill 1967:261). In Gabrielino, the -t absolutive has changed to -r, as occasionally occurs elsewhere (Kroeber 1909:269). The spirantization of the velar in Serrano indicates that it is a back k, since k does not spirantize (Hill 1967:256).

The root wiakə- is not a regular reflex of PUA *wokon- in any of the Takic languages or reconstructed protolanguages. I tentatively assume that it ultimately derives from that proto-root. A possible path would be through a Cupan *wike-, borrowed into Serrano or Proto Serran, and undergoing an irregular change *i > ia, as for example Serrano po’niava-t ‘skunk’ (Fowler 1983:237) corresponding to Kitanemuk poniva-č ‘id.’ (Anderton 1988) and PNUA *poni- ‘id.’ (Fowler 1983:237). The Gabrielino and Kawaiisu forms could then be loans from Serrano. The shifts *i > *ia and the strengthening of *ia in the Cupeño form argue for a Gabrielino loan.73

For Luiseño ‘pawxi-t ‘yellow pine’, Bright (1968) proposes the etymology pa:-wi’xe-t, {water-pine}. Phonologically, this etymology is a good match, since it explains the initial stress of the compound as a consequence of the long vowel of pa: While it is tempting to relate ‘water’ in the proposed compound to the Luiseño use of the tree as boat material (discussed below), I suggest a different etymology. pa- occurs sporadically in several Northern Uto-Aztecan languages, as an augmentative morpheme.74

Luiseño  ‘nawqut ‘sumac, *Malosma laurina’
        ‘pa:nawqut ‘sumac, Rhus ovata’

Luiseño  ʂukat ‘deer’
        pa:ʂukat ‘elk’
Cahuilla pu:l ‘curing doctor’
        pa’vu?ul ‘bear shaman’75
Kitanemuk haɡa?atʃ ‘bee’
        pahaɡa?atʃ ‘yellowjacket’
Kawaiisu  tibija ‘deer’
        parihija ‘elk’
Tübatulabal  ?a:nint ‘yellow ant’
        pa?:a:nint ‘red ant’

73 Merriam’s Cupeño vocabulary, the source of wiçakə-t, has other apparent Gabrielino loans.
75 Intervocalic p>v, as in *pu?ul > pa-vu?ul, is common in Takic and elsewhere in Uto-Aztecan.
The Plank Canoe of Southern California

Hopi kunja ‘fringed sagebrush, 
Artemisia frigida’

This list is not exhaustive, and several other examples exist in the languages listed and in others. However, there are not enough examples in any one language to judge whether this morpheme is productive in it, and so to judge whether words containing it have been borrowed from somewhere else. The pa- morpheme is likely cognate with Southern Paiute paʔa ‘high’ (Sapir 1931) and Kawaiisu paʔa ‘high, long, tall’ (Zigmond et al. 1991). In Luiseño, at least, the aː is long and attracts stress, just as paː- ‘water’ does. The existence of the pa- augmentative has been briefly noted by Hill (2005:201) for Cupeño and Takic in general.

Some dictionaries of various Uto-Aztecan languages etymologize the pa- morpheme in various compounds as ‘water’ without any semantic justification, where the augmentative clearly makes sense. I suggest that in this case reading ’pa-wxi-t as ‘water pine’ is not well-supported, and that it is better interpreted as ‘big pine’, distinguished from smaller, economically unimportant pines.76

Whatever the details of the etymologies for the various reflexes of Proto Takic *woko-, that simple stem is the only one for ‘pine’ that can be reconstructed for Proto Takic, with various branches of the family innovating terms for specific pine varieties. This fits with a scenario where speakers of Proto Takic originated in an area with little diversity of economically useful pines. The speakers of the daughter languages would then have separately developed terms for the varieties of pine in the mountain terrains which they came to know or occupy.

10.4.2 Luiseño Boat Words: pawxit and wixet

The identity between the Luiseño forms for ‘wooden boat’ and ‘yellow pine’, ’pawxi-t, had been noticed before by Kroeber (1925:654). Kroeber (1910:268) had previously noted the ‘pine’=‘boat’ semantic equation in Chumash; but the connection between the Chumash and Luiseño examples was noted before only by Heizer and Massey (1953:298).77 Jones and Klar, who refer to Heizer and Massey’s paper in other matters (A:460, 461) do not mention this point.

Luiseño ’pawxi-t ‘wooden boat’ is recorded by Sparkman (quoted in Elliott 1999), ‘a canoe formerly used by Luiseño fishermen’; DuBois (1908:131), ‘canoe (also a box carved out of a log to keep things in)’; Bright (1968), ‘dugout canoe’; and most extensively Harrington, in his notes to Boscana (1978:112-113). Harrington shows clearly that pawxit applies to wooden boats in general.78 The earliest record of the word may be Crespi’s 1769 Juaneño form, <paut> ‘canoa o

76 Serrano has the pair ju’haf’s ‘pine, ponderosa pine’ and ’pa-juhařaf’s ‘bigcone spruce (Pseudotsuga macrocarpa)’ (Ramón and Elliott 2000:210; Merriam 1907b, 1909). In contrast to the Luiseño pawxit, here pa- probably does mean ‘water’, referring to the sap-rich heartwood of this tree.

77 “This peculiar canoe-pine linguistic parallel can hardly be fortuitous, and leads one to suspect some specific connection between the Luiseño dugout and the Chumash plank canoe. This possibility is enhanced by the fact that for a long distance north of the Chumash and south of the Luiseño wooden canoes of any kind are unknown. How the development of these types occurred is impossible to say, nor is it easy to imagine what relationships the two boat forms have, since their occurrences are geographically exclusive and they are technologically distinct.” (Heizer and Massey 1953:298)

78 Harrington (1986, 3:115:141) translates pawxit as ‘board boat’; and in his notes to Boscana (Boscana 1978:112-113) he has the Juaneño forms ’pawxi-t ku’awtal lo’xa-ʃ ‘plank canoe’, lit. ‘canoe made of pieces of wood’, {wooden.boat-ABS wood.piece.INST make-NMLZ} (see Elliott 1999:408, 463); and ’pawxi-t ha’ku-l’i-ʃ’

149
balza’ (Crespi 2001:306-307), though the elision of the xi is hard to explain.\footnote{Compare Crespi’s <piut> ‘tobacco’ to the later recorded form ’pivat (Elliott 1999).}

The Luiseño wi’xe-t ‘tule boat’, formally identical to the tree word, is recorded as early as Hale (1846), as wa’xe-t; by Sparkman, “tule canoe formerly used by Luiseño fishermen” (Elliott 1999); by Harrington, “wixét, made of pevéega, round tule” (Harrington 1986, 3:115:141), also with the form wi’xe:ʔet; and by Bright (1968).\footnote{Hale’s list is in the Acjachemem (Juaneño) variety, which usually shows *i>a in unstressed syllables (Lobo et al. 2005:45; Woodward 2007:90).}

The data are summarized by Elliott (1999).

\subsection*{10.4.3 Kitanemuk k\textipa{kiwakt} and k\textipa{kwekt} ‘boat’}

Two forms in Kitanemuk are recorded for ‘canoe’: k\textipa{kiwakt~kiwaxt} and k\textipa{kwekt}, used to describe the tule boats of the neighboring Yokuts of Buena Vista Lake (Anderton 1988).\footnote{The second k of k\textipa{kiwakt} is marked as retracted by Harrington, with k and k possibly distinct in Kitanemuk, as in Serrano. Zigmond and Merriam, the other transcribers of the language, do not mark this distinction.} The forms resemble no Yokuts words, but do resemble the Serrano wiak\textipa{x~kiwakt~kixat} ‘sugar pine, Coulter pine’ and the Proto Cupan *wexet ‘pine’. Kitanemuk territory bordered Chumash territory up to historical times, and there are Chumash loans in Kitanemuk. This would make a Kitanemuk ‘boat’=’pine’ equation plausibly related to that in Chumash, if one could explain the Kitanemuk form.

To establish a connection between Serrano wi\textipa{kt} and Kitanemuk k\textipa{kiwakt}, the shift of the initial consonant from \textipa{w} to \textipa{k} has to be explained. Both consonants are present in all Takic languages, normally as reflexes of the same consonants in the proto-language, e.g., in the reflexes of PUA *wokon- ‘pine’, discussed above, and *kwa ‘eat’. However, at least one other example occurs in Kitanemuk showing \textipa{w}>\textipa{k}:

\begin{verbatim}
  Kitanemuk k\textipa{kwakt\textipa{skaveyk\textipa{t}}: Serrano wa\textipa{f\textipa{t\textipa{utik, wa\textipa{f\textipa{kuvik ‘seven’}}}}[Anderton (1988); Merriam (1907d); Hill (1967:27)]
\end{verbatim}

Other than ‘seven’, the numerals 1-10 correspond nearly exactly between Kitanemuk and Serrano, e.g., Kitanemuk wa\textipa{f\textipa{sa}, Serrano wa\textipa{f\textipa{jah, ‘four’, from which the word for ‘seven’ is derived. A likely explanation for the initial consonant of Kitanemuk k\textipa{kwak\textipa{txaveyk\textipa{t}} is that the w assimilated to k under the influence of the subsequent k.

With this, I posit w>k as a sporadic sound change in Kitanemuk, witnessed by the word for ‘seven’. I suggest it operated on an older Kitanemuk form *wi\textipa{akt}, ‘pine’, either cognate with the identical Serrano form or borrowed from it. Thus at some point Kitanemuk *wi\textipa{akt became kw\textipa{kt}, either before or after taking on the secondary meaning ‘boat’, following the semantics of the neighboring Chumash.

This scenario rests on the existence of a sporadic sound change, attested by only one other word, which adds uncertainty to the explanation. Further support to its existence comes indirectly from a well-established similar change in another language: in Italic, word-initial labial *p may assimilate to labiovelar *k, conditioned by k in the following syllable. This change is sporadic, and is witnessed in Latin by only three instances: k\textipa{inx\textipa{e ‘five’ < PIE

\footnote{‘dugout canoe’, lit. ‘hollowed canoe’ \{wooden.boat-ABS be.hollow-CAUS-NMLZ\} (see Malécot 1963 for morphology). The mention of plank boats could refer to Chumash or Gabrielino boats, or it could have referred to European wooden boats, built of nailed planks.}
such a change is therefore possible, though it may operate only sporadically. Two of these three Latin words happen to be a numeral and a tree, as in Kitanemuk.

As mentioned above, the identity of the velar consonant in the Serrano form wiakt is supported by its occasional spirantization to x, though that phonetic detail is not indicated by Merriam, the source for the form. The presence of the k in corresponding place in the Kitanemuk form further supports the historical connection between the two forms.

Anderton (1988:380) notes that kwiakt is a likely borrowing, because of the final stress, which is atypical for Kitanemuk. It is not clear where the stress falls in the Serrano source proposed here, but stress on the a, which would fit with Anderton’s observation, is not excluded by the data.

The other Kitanemuk form, kwekt, follows similar reasoning. I start with a loan from a Cupan source, *weket; the Cupan languages are the only ones which front Proto Takic *o to e or i. Although all the Cupan languages show spirantization of k to x, this change may have occurred in Proto Cupan after the vowel change, which justifies this protoform. Next come *w>kʷ and lenition of the unstressed vowel, producing kwekt, exactly as with wiakt > kwiakt above. The Cupan source is necessitated by the vowel change, but the location and sociolinguistic situation of this borrowing are unclear.

Serrano and Kitanemuk are closely related and geographically adjacent languages, and the Serrano word could plausibly be borrowed into Kitanemuk for a species of pine distinct from wokoh-t ‘gray pine’, the directly inherited word. Alternatively, the doublet *wokoh-t/*wiak-t could have existed in Proto Serran and inherited by its daughter languages.

There is nothing to indicate at what stage of the formal development of this word it took on the secondary meaning ‘boat’.

10.4.4 Roseño Chumash tak ‘pine’, ‘dugout canoe’

Woodward (1934:121) quotes an earlier publication where a Santa Rosa Chumash of ca. 1820 described the use of dugouts and of plank canoes, and gave two words for ‘boat’, recorded as <toak> and <comow>. As discussed by Klar and Jones (B:378), these words resemble the Central Chumashan words for two types of pine, tak and tomol. Klar and Jones are uneasy with the odd sound changes which the quoted Roseño forms show. I concur, though I believe that these apparent sound changes could be the result of copying errors. I think it is unlikely that <toak> represents tok~toq, the word for the milkweed cordage used for lashing boat planks.

If these forms are correct, then Roseño tak provides another example of a ‘pine’=‘boat’ semantic loan. If <comow> refers exclusively to a plank boat, then perhaps <toak>, by contrast, refers to the dugout canoe; both are described in the same source and in the same order.

10.4.5 The Areal Spread of ‘pine’ = ‘boat’

---

82 I have not seen the original form of this text, published in the Santa Barbara Gazette. Woodward accurately quotes a version published in the San Francisco Herald of Dec. 11, 1859.
83 Barbareño and Ineseño tak ‘Pinus sp.’ (Timbrook 2007); Barbareño taq ‘white pine’ (Whistler 1980). Hudson et al. (1978:48 n.3) are not certain about the identity of tak, but believe it refers to all pines other than piñon pines and yellow pines.
84 tak as an example of the boat/pine equivalence has been suggested before by Heizer (1941a).
The distribution of ‘pine’/‘boat’ words — clustered in Southern California and nonexistent elsewhere in the region — indicates a historical connection. As in the Chumash case, the original sense of Luiseño wixet and pawxit was ‘pine’ (going back to Proto Uto-Aztecan), with the sense ‘boat’ coming later. The appearance of this unusual semantic equation in Chumash and the geographically close but unrelated Luiseño can be explained as a semantic loan from one language to the other. Semantic borrowing (Durkin 2009:136; Hock 1986:398) is the process by which a language adopts a meaning for a word on the model of another language. As an example, the English *star*, originally referring to the celestial object, later came to take on the additional meaning ‘performing celebrity’; on that model, the Spanish *estrella* ‘star, celestial object’ took on the same secondary meaning, as did the equivalent words in Russian, Turkish, Hebrew and other languages.

In the scenario I suggest here, both the Luiseño and the coastal Chumash had dugout canoes constructed from pine logs; whichever of the two was the first to develop dugout canoes named the boat after its source material, yellow pine; and this type of boat construction and the semantics of its name passed together from one people to the other. At some later time, the Chumash developed the plank canoe, which mostly replaced the dugout but retained its name.

A similar scenario holds in Kitanemuk, where the dual meaning of *kwi’akt ~ kwekt* is explained as a semantic loan from the neighboring Chumash. No wooden boats existed in inland Southern California, and so the shift ‘pine’ > ‘wooden boat’ > ‘boat’ could not be explained by internal development alone. The semantics of the Kitanemuk word are comfortably explained as a result of contact with the dual meaning of *tomol(o)* in a Chumashan language.

As a working hypothesis, I will assume that the Chumash built wooden boats before the inhabitants of Luiseño country, and named them after pine trees. Chumash *tomol(o)* ‘yellow pine; wooden boat’ would then be the source of the semantic development of Luiseño *pawxit* ‘ponderosa pine; wooden boat’ and *wi’xet* ‘pine, tule boat’. But it is possible, in principle, that wooden boats were first built in the south, where Luiseño is now spoken, and that the technology and the term were later taken up in the north, where the planked canoe was later elaborated. It is even possible that the ultimate source was some coastal Yuman language (Hinton 1991:152), which was spoken where Luiseño is now.

The Luiseño doublet *pawxit ~ wi’xet* could have developed in several ways. By one scenario, *wi’xet* first took on the general meaning ‘boat’, as in Kitanemuk. *pa-wxit* would derive separately for the two senses of *wi’xet*, to mean ‘large pine’ and ‘large boat’, the latter referring to wooden boats. The semantic range of *wi’xet* ‘boat’ would then shrink to refer only to the lesser, tule-made boats.

In another scenario, Luiseño *pa-wxit* ‘ponderosa pine’ would derive from *wi’xet* ‘pine’. Next, *pawxit* would have taken the secondary meaning ‘wooden boat’, after the Chumash model. And finally, *wi’xet* would have taken on the meaning ‘lesser boat (i.e., tule boat)’ by back-formation based on *pawxit*, i.e., ‘big pine’ : ‘big boat’ → ‘(smaller) pine’ : ‘(smaller) boat’. This scenario is more complex, but, as in Chumashan, associates the wooden boat with the particular type of pine from which it is made.

Notably, Gabrielino did not adopt ‘pine’ as a metaphor for either the tule boat or the plank boat. Semantic shifts are in general not predictable, and so any historical conclusions based on this are uncertain; however, it may be that the ancestors of the Gabrielino were not initially bilingual in Chumash, and so used neologisms unmotivated by Chumash semantics to describe the local boats.
10.5 Other Boat Terms in Southern California

Some other boat terms occur in Southern California, which are of interest to the discussion here.

10.5.1 Kawaiisu kwijakata ‘tule boat’

Zigmond et al. (1991) give the Kawaiisu form kwijakata, ‘tule balsa said to be used by the Yokuts’. This is clearly a loan from kwiaqt of the neighboring Kitanemuk, showing its irregular \textit{w} > \textit{kw}. This implies that the Kawaiisu came to be in the vicinity of Lake Buena Vista after the Kitanemuk, and after the Chumash had developed the word \textit{tomol(o)} ‘boat’ and passed it as a semantic loan to the Kitanemuk. This is consistent with the chronology of Sutton (2010a).

10.5.2 Kitanemuk tomoɬ ‘large boat’

Anderton (1988:380) gives Harrington’s form tomoɬ, ‘big ship, canoe, steamship’, and notes it as a subset of the more general kwekt/kwiakt ‘boat’ discussed above. This Chumashan loan refers to European ships, but it is not certain whether it refers to the Chumash plank boat as well. ɬ is not a native Kitanemuk phoneme but occurs in Chumashan as an allophone of \textit{l}. This appears to be a late loan, perhaps even post-mission.

10.5.3 Ventureño Chumash ʔaxipeneʃ ‘dugout canoe’

Klar and Jones (A:472, B:374) discuss the Ventureño form ʔaxipeneʃ, ‘dugout canoe’, analyzed as ʔaxi-\textit{pen-ef} \{work.wood-strip.off-RESULTATIVE\} ‘worked timber’ (see Hudson and Blackburn 1982:338). The verbal stem ʔaxi-\textit{pen} ‘to work wood’ is recorded in Barbareño and Ineseño as well (Whistler 1980; Applegate 2007) and the etymology is transparent. But Klar and Jones’s statement that “the word is probably from an old stratum of Chumashan development” is not necessarily true. The word is as easily explained as a later replacement for an earlier term, \textit{tomol(o)}, which went from signifying ‘dugout boat’ to ‘plank boat’, as I propose, with ʔaxipeneʃ filling the semantic gap left behind.

10.5.4 Purisimeño Chumash <šwašwax> ‘canoe’

Pinart’s 1878 vocabulary of Purisimeño gives the form <šuašuax>, ‘canoe’ (Heizer 1952:44-45). This is the only known record of this form. Klar and Jones (B:395-397, D:89, E:174-175) discuss it as a possible older word for ‘canoe’ in general, one predating \textit{tomol(o)}.\footnote{Klar and Jones suggest that the name swaxil, a village on Santa Cruz Island, derives from this root and means something like ‘boat place’ (B:395, 397). They do not give any other examples of a Chumash suffix \textit{-il} or a similar compound-forming root, in placenames or elsewhere, and I haven’t found any either.}

The exact phonetic form heard by Pinart is uncertain. Some of the words he recorded bear a final <x> which corresponds to a phonemic /x/ in other, more phonetically reliable sources. For example, his Ventureño <tsitsalsax> ‘thumb’, corresponds to Barbareño ɨ\textit{s}’\textit{a}lxax (Whistler 1980), and his Ineseño <suačax> and Ventureño <čuačax>, ‘arm’, match Barbareño \textit{waf}’\textit{ax}. In other cases the final orthographic <x> occurs where the word actually ends with a phonemic vowel or semivowel, e.g., Ineseño <mohox> ‘beach’ for muhuw (Applegate 2007), or Roseño <huimax> ‘Santa Rosa Island’ for wi’\textit{ma} or wima\textsuperscript{9} (Whistler 1980; Applegate 1974:194). In the
latter cases I read Pinart’s <x> as a mishearing of a devoiced echo vowel, e.g., [wimaʔa] in the last example.

With the latter interpretation, <šwaʔwaΧ> could be [fwawʔa], a reduplicated form of fwaʔa. If we allow for f~s alternation (Applegate 1972:60-61; Klar and Jones E:182; but see Klar 1977:127), then this root could be read as swa or swaʔa, ‘tule, Scirpus sp.’ (Heizer 1952:55; Timbrook 2007:203, 206). With that, fwaʔa could have meant ‘tules’ and referred to a tule boat.

Other than the f~s alternation, this etymology suffers from uncertainty in the identification of the particular species of Scirpus which swa refers to. Of the several species of tule which grow in the area, only round tule, Scirpus acutus, is suitable for building tule boats. Timbrook (2007) has Ineseño swa ‘S. americanus, S. pungens’, but also has Ineseño swow ‘S. acutus’. Pinart (in Heizer 1952:54-55) records <sua> as ‘tule’ in Purisimeño but <sua> ‘round tule’ in Barbareño. It is not clear if this variation reflects true dialectal differences or if it reflected imperfect knowledge of plant names by the speakers who supplied these words.86

## 10.5.5 Miscellaneous Luiseño Words for ‘boat’

Harrington (Boscana 1978:113) records several additional Luiseño boat words. wotilaf ‘rowboat’ (also Harrington 1986, 3:115:141) derives from woti ‘to row’. yatalaf ‘floating thing, boat, raft’ (see also Elliott 1999) derives from yata ‘to float’, possibly a loan translation from ‘lipay ‘Aa (Mesa Grande Diegueño) kutuyalaf ‘boat’, lit. ‘that which floats, is carried by water’ (Couro and Hutcheson 1973). The root yata itself is a Yuman borrowing with cognates elsewhere in the family (‘lipay ‘Aa tuyalp ‘carried away by water’; Cocopa yalyal ‘float, as paper on water’, Crawford 1989).

Harrington also records the Spanish loans votti (< bote ‘dinghy’), vuksi (< buque ‘ship’) and vapoz (< vapór ‘steamship’).

Gatschet (1879) has the Luiseño (‘Kechi’) word <ë-val> for ‘canoe’, recorded by Eric Bergland. This appears to be ñiva-1 ‘large wooden spoon, trowel, stirring paddle; species of wood used in making earthenware’ (Elliott 1999). It could be a semantic extension of the word referring to the implement, or refer to a dugout made from that type of wood.

### 11 Reconstructing the Prehistory of Wooden Boats in Southern California

In historical times, Chumash canoe planks were usually fashioned from driftwood, of which coastal redwood (Sequoia sempervirens) was the preferred species (Hudson et al. 1978:47-49). Redwood is resistant to weathering, strong, light, straight-grained, shrinks little, and is easy to

---

86 The Chumashan swa~swax or fwa~fwax resembles words for ‘boat’ or ‘canoe’ in several California languages: Northern Paiute saki (Hale 1846); Coast Miwok šaka (Callaghan 1970); Plains Miwok soka (Merriam 1903c); Northern Valley Yokuts fua (Kroeber 1959:10). Some of these might be related genetically or through contact; that remains to be investigated. Callaghan (2001:322) suggests linking Coast Miwok šaka and Proto Maiduan *dʒak ‘bridge, boat’ through contact.

The Chumashan swa ‘tule, Scirpus sp.’ and Gabrielino swa-r ‘Mission tule (Juncus textilis)’ (Merriam 1903b) may be related to each other, and perhaps to some of the words for ‘boat’ above. Munro (1983:290) derives Gabrielino swa- from a protoform *siya-. Cf. also the Ventureño syit ‘base of stems of Juncus textilis’ (Timbrook 2007) and the Gabrielino <Se’-e> ‘Round tule, Scirpus lacustris’ (Merriam 1903b).
work (Anonymous 1999:1-16 and passim), which makes it a superior material for planked boat construction. Abundant driftwood is carried south to the Santa Barbara Channel by the California current. Why, then, was \textit{tomol(o)} ‘pine’ used as the Chumash metaphor for the planked boat, rather than \textit{wi\textsuperscript{2}ma} ‘driftwood, redwood’?\footnote{Barbareño \textit{wi\textsuperscript{2}ma} ‘plant sp.: red pine’ (Whistler 1980); Ineseño \textit{wima\textsuperscript{2}} ‘redwood’ (Applegate 2007). In Hudson et al. (1978) the word refers to driftwood in general and redwood in particular, and is also the name for Santa Rosa Island, presumably because much driftwood is found there.}

According to Hudson et al. (ibid.), pine and other wood types were considered potential boat timber. \textit{tomol} (yellow pine) was the next best thing to redwood, and it appears that even inferior woods such as Torrey pine could be used when nothing else was available. Blackburn (1975:209) quotes a Chumash tale in which Coyote travels to a location called \textit{tomto\textsuperscript{2}mol} or \textit{hultomto\textsuperscript{2}mol} (‘the pines’) to buy “\textit{tomol} pine boards” from an old man there.\footnote{\textit{tom~to\textsuperscript{2}mol} \{REDUP-pine\} ‘pines’, \textit{hu=l-tom~to\textsuperscript{2}mol} \{REMOTE=DEF-REDUP~pine\} ‘the pines’ in Barbareño, the native language of the storyteller, Juan de Jesús Justo; see Wash (2001:59, 61).} He then carries them home and goes on to build some planked boats with them. While this story lends weight to accounts of the occasional use of pine for boat planks, it also highlights the difficult requirement of transporting a boat’s worth of wooden planks, some 100-200 kg, from deep in the mountains to the seashore.

I propose here that the Chumash plank canoe evolved from dugouts, similar to those attested in historical times among the Luiseño, and that it was this type of boat that was first named after yellow pines, the material used in their construction.

11.1 Dugout Boats in Southern California

11.1.1 Channel Chumash Dugouts

The Chumash built dugout canoes in addition to the \textit{tomol}, though these are less well documented (Woodward 1934:120; Heizer 1955:151, after Henshaw; Hudson et al. 1978:31-37; Hudson and Blackburn 1982:338-340, mostly after Harrington; Cunningham 1989:61-63). Woodward’s ultimate source, Father Antonio Ripoll, lived in Santa Barbara around 1820. He describes dugout canoes, symmetrical in shape, 10m long by 1m deep and wide, carved out by stone tools. There is no mention of the use of fire to hollow the logs. Hudson et al. describe dugout construction as recounted to Harrington by Fernando Librado and perhaps others. These boats were made of willow (\textit{Salix} sp.) or cottonwood (\textit{Populus trichocarpa}). These are large, fast-growing trees which grow near streams at low elevations, but their wood is very heavy when unseasoned, and therefore makes boats of low freeboard which are not suitable for the open ocean; such heavy boats also carry less weight. In general, dugouts were hollowed by repeated burning and gouging. The boats were not stable, and were not used in the open ocean. Henshaw appears to describe the canoes of Santa Rosa island as dugouts made with stone tools alone, without fire. Some post-missionary dugouts were hollowed by mechanical means alone, and were outfitted with benches and oarlocks in European fashion.

Several studies (Heizer 1940; Robinson 1942:208-209; Lee 1981:51; Cunningham...}
1989:62-63) describe miniature boat effigies which have been found in Chumash areas on the mainland and the channel islands. Some of these effigies can be confidently recognized as depicting Chumash-style plank canoes, but recognizing other as depicting dugouts is more ambiguous. For example, the boat effigy of Sequit Canyon described by Cunningham has nearly symmetrical pointed ends, consistent with the descriptions of Harrington and Henshaw, but a strongly curved keel, which does not fit easily with the description of the Chumash dugout.

Heizer and Massey (1953:298) argue that the Chumash dugout was a post-missionary introduction, probably from the Luiseño. They base this on the observation that many planked canoes were seen among the Chumash by early travelers, but no dugouts were described until Ripoll’s account of the 1820s. This argument is weak, since the lack of early observations indicates only that planked canoes were predominant, and does not exclude dugouts as a minor type of boat. The Chumash probably had the knowledge of producing dugouts at the time, but not utilized it very often, as was the case with tule boats, which are missing from the early records as well. In any case, the existence of the pine=boat equation in Luiseño, Chumash and Kitanemuk indicates earlier sharing of boat-building knowledge among the people of the area, and certainly predating Costansó’s record (Costansó 1770:40) of the Barbareño Chumash word *tomol*.

### 11.1.2 Northern Chumash and Salinan Dugouts

Further north, some record exists of dugout boats at the northernmost corner of coastal Chumash territory, where it meets Salinan territory. On his voyage south from British Columbia in 1793, Vancouver spotted a few kilometers off the coast between San Simeon and Morro Bay a boat, “neatly formed of wood, much after the Nootka fashion” (Vancouver 1984:1087, Menzies and Eastwood 1924:314), and paddled by four people. He got close enough to recognize the shapes of the paddles, which suggests that he would have recognized the boats as built of planks if they were so, but instead recognizing them as dugouts, as are those of the Nootka (the Nuuchahnulth of Vancouver Island). Alternatively, Heizer and Massey (1953:301) propose that the canoe was one of two plank boats purchased by the mission at San Luis Obispo from the Santa Barbara Chumash, some twenty years earlier.

Early in the twentieth century, Harrington’s Migueleño Salinan consultants Pacifico Archuleta and Juan Solano described dugout canoes made by burning the interior of a log of oak or live oak (*Quercus* sp.), or of sycamore (*Platanus racemosa*). Archuleta had seen them on the beach in what is now Cayucos (Harrington 1986, 2:84:233, 127, 128; Immel 2007). The Antoniaño Salinan David Mora told Harrington only of tule boats, which he considered superior to plank boats by being lighter and harder to sink (Harrington 1986, 2:87:461). It appears from these fragmentary data that the Antoniaño, whose territory reached the coast (Gibson 1982), did not use dugouts. The testimony of the Migueleño Salinans, who lived inland, refers to dugout usage by their neighbors of Estero Bay, but no further north.\(^{89}\) As in the description of the dugouts of the channel Chumash, the use of heavy timber, here oak and sycamore, precludes the use of these boats on the open ocean. Notably, there is no historical record of the use of ponderosa pine, even though stands of it grow near the coast, north of San Simeon. As with the

---

\(^{89}\) Cayucos, the point in question, is on what may have been the boundary between the northernmost Chumash and the poorly described Playanos to the north. Whether the Playanos were a distinct ethnic group, and if so whether they were Chumashan, Salinan, or something else, is not known (Milliken and Johnson 2003:128-134).
Chumash dugout, the use of fire rather than just mechanical means argues against a European source for this technology. Despite this ethnological record of the use of dugout canoes, the archaeological record shows no significant evidence for offshore fishing in the area at any time in the past (Joslin 2010).

11.1.3 Luiseño Dugouts

The Luiseño are the only group in the area which used dugouts as a significant means of water transport in historical times:

Some wooden canoes were also made from the trunks of trees. It is stated that voyages were formerly made with these as far as San Clemente Island. [Sparkman 1908:200]

Pauhut, canoe (also a box hollowed out of a log to keep things in)...pauhit, yellow pine, also dug-out canoe [of cottonwood?]. [DuBois 1908:131]

[Marcus Golsh] reported stories that his grandfather had told of making canoes on the forest-mantled slopes of Paauw [paʔauʔaw, Mount Palomar], at the Pine Camp of Uuszkun [ʔuʔʃuʔkun, Doane Valley]. These fire-and-abrasion hulls of yellow pine were finished late in the fall, so that after abatement of winter storms they could be paddled down tributary streams to the River San Luis Rey and out to the coast, where they were sold to shore-side villages for use at sea. [Cunningham 1989:61-62]

Cunningham’s account highlights the significance of log transportation methods to the feasibility of constructing dugout canoes. Ponderosa and Jeffrey pines grow in Southern California at elevations above 1500m (Burns and Honkala 1990). Mt. Palomar, some 50 km from the coast, is the nearest location where these pines grow. Doane Valley, according to Harrington (Boscana 1978:113), is the only place on Mt. Palomar where yellow pine grows. According to Golsh, the Luiseño also traded logs burned into charcoal to the islanders; these were brought down from the mountains, presumably in the same way, then tied together into rafts and floated across the channel (Cox 1968). Water transport through swollen rivers, as described, is the only practical way to get logs or dugouts from these mountains to the coast.

11.1.4 Explaining the Distribution of Dugout Boats

This constraint on log and canoe transportation severely limits the locations where dugout canoes can be built and launched. The only large rivers in Southern California which drain an area where yellow pines grow are the San Luis Rey, discussed above; the Santa Ana, which travels

---

90 Boscana (1978:24) wrote in his 1822 Luiseño ethnography, “they constructed out of logs very swift and excellent canoes for fishing.” From the context, however, this appears to refer to the Chumash, not the Luiseño. This quote is known only from a translated version (Robinson 1846:240). It does not appear in the only known version of the Spanish original (Reichlen and Reichlen 1971).

91 Marcus Golsh (1890-1988) was a Rincón Luiseño tribal leader. His grandfather was Santiago Duro. This passage is to be read with some caution, since Cunningham may have woven together Golsh’s report with some of Harrington’s notes to Boscana (1978). However, the passage is consistent enough to be acceptable in its entirety. Cunningham say he elicited this account from Golsh on several different occasions.
from the San Bernardino mountains to the shore in Orange County; the San Gabriel, which
drains the San Gabriel mountains and ends near Long Beach; the Santa Clara, which meets the
Pacific in Ventureño Chumash territory; and the Santa Ynez, which reaches the coast in
Purisimeño Chumash territory, near Lompoc. Stands of ponderosa pine also grow very close to
the coast to the north, around San Simeon and south of Big Sur. As with the San Luis River,
heavy flow in these rivers is limited to the late winter and early spring (see Horne 1981:20 for
the Santa Clara River).

This partly explains the distribution of dugout canoes in Southern California: dugouts existed
only where they could be built out of suitable material (yellow pine) and transported to the coast.
The Gabrielino, by this argument, did not have dugouts because the long, shallow and
intermittent San Gabriel River was not adequate to carry pine dugouts from the San Gabriel
mountains to the coast. Other rivers, such as the Santa Ana and Santa Ynez river, may have
been capable of transporting logs and boats but required a long travel from source to coast.

In all three areas — Luiseño, Santa Barbara Channel Chumash, and Northern Chumash —
heavier woods such as cottonwood, sycamore or oak were used in historical times for building
dugout boats for coastal use. This appears to be a late development, probably later than the
arrival of Europeans, the establishment of the missions, and the cessation of travel to the islands.
For traveling and fishing near the shore and in esteros such boats would have sufficed, and the
woods from which they were built were from easily accessible coastal trees. It is very possible
that the transportation of dugouts from inland always required a large number of people, even
assisted by rivers, and that the depopulation of native communities after European contact made
such projects harder to carry out.

There is no record of dugout canoes used on the California coast from north of Salinan
territory until reaching Wiyot territory some 500 km to the north (e.g., Kroeber 1922:269),
though tule boats were used in Monterey and San Francisco bays. This cannot be be
explained entirely by the lack of appropriate wood. For example, redwood grows abundantly in
the Santa Cruz mountains south of San Francisco, and reaches close to the coast near Santa Cruz,
and could have been fashioned into dugouts. Even in far northern California, ocean-going dugout
boats are not designed for open ocean navigation (Hudson 1981b).

By one argument, oceangoing boats, including dugouts, were scarce along the central
California coast because of the difficulty of navigation in that exposed area (Arnold and Bernard
2005:110). There may be some truth to that, but sailing in that area is not always excluded, and
even in the supposedly sheltered Santa Barbara Channel safe sea and weather conditions are not
guaranteed (Fagan 2004:7-8).

It appears, then, that the distribution of wooden boats in coastal California was not
conditioned solely on availability of wood, on ocean conditions, or on access to offshore

---

92 There is one mention of the transportation of a large log from the mountains among the Gabrielino (McCawley
1996:161-164). A certain mourning ceremony, held every few years, required the erection of a 10-15m pole, which
was cut from a pine tree and brought back to the ceremony area in the valley. There are no more details about the
method of transportation, but it cannot have been easy, and transporting logs or dugouts all the way to the coast
would have been more difficult still.

93 Pinart’s vocabulary of Esselen records the item <ualkošex> ‘canoe (dugout)’ (Heizer 1952:76). There is no other
indication that the Esselen ever used any boat other than a tule boat. Fish remains from Esselen territory (Breschini
and Haversat 2004:119) are limited to near-shore species, suggesting minimal use of boats. Esselen territory is
mountainous down to the coast, and lacks large coastal plains suitable for settlements.
The Plank Canoe of Southern California

fisheries. The best predictor of the their presence is the need for transportation to the islands off the coast. As Sparkman has noted for the Luiseño, dugouts were used for travel to the islands (pace Fagan 2004). Dugouts were built in Southern California because they provided a good means for reaching the Channel Islands, and were built wherever appropriate wood was available.

The dugouts of the northernmost Chumash, far from the Channel Islands, are an exception to this. Apparently they were used for fishing, in preference to tule boats, but at present I do not know if that was due to absence of tule, the availability of ponderosa pines near the coast, or some other reason.

11.1.5 Dugouts: Summary

I have argued that dugout boat construction was traditionally known throughout Southern California, and that dugouts were built for the purpose of travel to the Channel Islands. The Luiseño dugout is known to have been constructed of yellow pine transported from the mountains to the coast, and took its name, pawxit, from the name of the tree. The boat of the neighboring Chumash likewise took its name from the Chumash word for the yellow pine, tomol, and must have likewise been a pine dugout, with one language borrowing the secondary meaning from the other. This is easily consistent with the Chumash plank canoe developing after the dugout but retaining the earlier name, but not with scenarios involving an introduced plank canoe accompanied by a borrowed name.

11.2 From Dugouts to Plank Boats

Dugout construction required the transport of a heavy and unwieldy boat from the mountains over rough terrain or by the lucky placement of a river. The Chumash plank canoe did not have this disadvantage. Logs can be split into planks, which can be transported to the coast in several trips and even uphill. Easier yet, planks could be fashioned from driftwood found in coastal areas, which often was of the more durable redwood.

This provides clues to the transition from dugout to plank boat among the Chumash. Increasing contact with the islands and exploitation of ocean resources required more numerous and larger boats, while the number of dugouts which could be transported to coast through the Santa Clara river at times of high flow (Horne 1981:20) was limited. The development of planked canoe technology overcame this limitation and allowed for the production of a great number of voluminous and seaworthy boats, as was already suggested by Heizer (1938:221).

This scenario of transitioning from dugout to plank canoes finds a parallel in East Polynesia. At the time of European contact, the most elaborately developed fully-planked boats in East Polynesia were found in the Tuamotus, a chain of low-lying atolls, poor in large, high-quality trees. Tuamotuan planked canoes were seaworthy, and often very large; they were so highly valued that they were exported to the richer and larger Society Islands, and sometimes Tuamotuan boat-builders would be brought there as well (Haddon and Hornell 1936:79).

94 It is also conceivable that sometime in the past climatic conditions in Chumash territory were favorable to the growth of yellow pine at lower elevations, including the slopes of the Santa Ynez mountains above Santa Barbara, and that changing conditions moved the range of yellow pine further inland and forced the Chumash to find more accessible wood and techniques for its use. Direct evidence for the past abundance of pine, based on pollen records from the region, is ambiguous (Heusser 1978; Heusser 1995; Davis 1992), and provides no species-level detail; this scenario remains a speculation.
Evidently the lack of large timber and the necessity of frequent travel among many small, widely-scattered islands led to the development of these boats, which were more labor-intensive, but ultimately more versatile than the dugouts fashioned from the larger trees of the high islands. I believe that this same path was followed in the development of the Chumash plank canoe, especially on the timber-poor but driftwood-rich northern Channel islands.

There is no remaining evidence of the intermediate stages between the simple dugout and the tomol, but dugouts may well have been more complex in the past. For example, they may have been made more seaworthy by widening and flattening the hull, or by the addition of one or more layers of strakes. If dugout canoes were used for deep sea fishing before the transition to fully formed plank boats, the timing of the invention of the tomol based on early fish remains, as proposed by Bernard (2001, 2004) and Arnold and Bernard (2005), may have to be reconsidered.

11.3 Plank Boats Outside the Santa Barbara Channel

Outside the Santa Barbara channel, plank boats were used by the Gabrielino to the south as far as San Pedro. Plank boats were encountered in the northern Channel Islands, and on San Clemente and Santa Catalina islands (Wagner 1928:47; Vizcaíno 1959).

Details on the Gabrielino plank boat are lacking, but it was apparently similar to the tomol. The only known difference is the shape of the prow and stern, inferred from boat effigies (Hudson et al. 1978:96-97); this detail might reflect different uses of the boat in these two areas. While there exist abundant early observations of Chumash boat construction, no such evidence exists for plank boats manufacturing in Gabrielino territory. Triangular stone drills, with which the Chumash drilled holes in canoe planks, do not appear in the Gabrielino archaeological record. These two observations have led Cunningham (1989:76) to propose that the Gabrielino obtained plank boats from the Chumash, rather than manufacturing them themselves. And indeed the Gabrielino would not have been able to obtain boat construction material easily, whether redwood drift logs, probably limited to the eastern part of the Santa Barbara Channel, or pine planks, which would require transportation by land over a great distance and mountainous terrain. The same argument applies to the island Gabrielino of Santa Catalina, and possibly to the people of San Clemente as well; more information about driftwood abundance there would clarify the issue.

Plank boats were expensive to produce, and the Gabrielino would have needed material wealth to trade for them. Such wealth was generated through the natural resources of Santa Catalina, including shells for ornaments and steatite (McCawley 1996:112). As the inhabitants of the southern Channel Islands were Gabrielino (Hudson 1981a; Sutton 2010b) or anyway Takic speakers (Munro 2002), their boats may have ultimately been supplied through the mainland Gabrielino. In any event, plank boats were not as ubiquitous among the Gabrielino as among the Chumash, as attested by the use of tule boats for ocean travel among them (McCawley 1996:125).

The Luiseño did not use plank canoes. Their dugouts, as mentioned before, have gone as far as San Clemente, and would have reached Catalina even more easily. But in both places the more distant Luiseño and their boats must have had a lesser presence than the Gabrielino. Accordingly, their share of the wealth of these islands would be small, and they would not have had the means to trade for Chumash plank boats. This lack of plank canoes may explain why the Luiseño continued to use dugouts up to the early historical period.
12 Explaining the Distribution of Planked Boats in the Americas

Plank boats are rare in the Americas, especially compared with the rest of the world. As I have shown, plank canoes developed in Southern California from dugout canoes, out of the necessity for seaworthy boats capable of crossing from the mainland to the Channel Islands, over distances of tens of kilometers. On the west coast of the Americas, no other such islands or wide bights exist between the Olympic Peninsula on the Canadian border and Chiloé Island in Chile, with the exception of the Coiba and the Perlas Islands off the coast of Panama.\textsuperscript{95} In the Pacific Northwest, very large trees were available for building beamy dugouts, and even those required sewn-on strakes for handling the waves through long crossings, as was discussed above in section 3. Other offshore islands throughout that area were either too small to be useful (e.g., the Farallon Islands, off San Francisco), too remote to be familiar, or close enough to shore to be reachable by boats of modest capabilities (e.g., Cedros Island off Baja California). In polar and sub-polar areas of the Americas, where workable lumber is rare, bark boats and animal skin boats are seaworthy substitutes for wooden boats. The sewn-plank boat of the Patagonian coast, the \textit{dalca}, was developed by adapting the sewn-bark boats of the Chono from the south for use with the wooden planks of the Huilliche from the north (Lothrop 1932:249, 251). The \textit{dalca} was developed not only for seaworthiness, but also to be easily disassembled and transported over land (Lothrop 1932:247). Southern California is the only locality on the west coast of the Americas where large islands off the coast required open crossings, and where arid climate limited the availability and accessibility of suitable trees for dugout construction.

On the east coast of the Americas, a similar situation holds. The only place where offshore islands required open crossings was in the Caribbean, and there large logs suitable for dugout construction were available. As in the Pacific Northwest, the technique of sewing planks on to a dugout base was developed there to increase the seaworthiness of the boats.

In contrast, the coasts of Asia, the Mediterranean and Europe are surrounded by abundant targets for seaworthy boats. And, of course, ocean navigation was at the heart of the settlement and daily life of Oceania. In all of these areas, dugouts with sewn-on strakes and fully planked boats were known until the advent of metal nails and metal tools.

The east coast of Africa, while free of islands or large bays, was frequented until recently by planked boats, part of the large trade network which stretched across the Indian Ocean. The area most analogous to the smooth coasts of the Americas is the west coast of tropical Africa south of Senegal, where ocean navigation was mainly along the coast, and where planked boats and sewn-on strakes were unknown (Smith 1970), as in most of the Americas.

13 Conclusion

Jones and Klar have presented what they consider archaeological, ethnological and linguistic evidence for a Polynesian origin of the plank canoe of Southern California. I have shown here that none of that evidence is valid. There is nothing to show that the Chumash \textit{tomol} and the Gabrielino \textit{ti’at} were inspired by external contact.

Linguistic and ethnographic evidence from Southern California suggests a long history of pine-built dugout canoes, which would be the ancestors of the Chumash plank canoe. The homonymy of the words for ‘pine’ and ‘boat’ in Barbareño and Roseño Chumash, in Luiseño,

\textsuperscript{95} The natives of Coiba and of the Perlas archipelago were exterminated soon after European contact. I know of no information regarding their boats.
and in Kitanemuk supports this model, and points out some new historical detail.

The equivalence of Chumashan *tomol* ‘yellow pine; plank boat’ and Luiseño *pawxit* ‘yellow pine; dugout’ and *wixet* ‘pine; tule boat’ indicates that both communities shared the technology of dugouts built of yellow pine. The Chumash later elaborated the dugout into the familiar plank boat, with the dugout remaining a marginal form. It would be difficult to reconcile an external introduction of the plank canoe into Chumash territory with the usage of the parallel term by the Luiseño to describe their dugouts.

The semantics of the Chumashan *tomol* guide those of the Kitanemuk *kwekt* and *kwiakt*, originally meaning ‘pine’, but later referring to the tule boats of the Yokuts of Buena Vista Lake. Buena Vista Lake has existed since ca. 2000 BC (Kennett et al. 2007:537), and was presumably navigated soon thereafter, by whatever people lived by its shores. According to Sutton (2010a), the ancestors of the Kitanemuk arrived at their historical homeland at about that time. By the simplest linguistic scenario, these Uto-Aztecan settlers then came in contact with Chumash speakers, and fashioned their word for ‘boat’ after the Chumash model. Coastal Chumashan speakers therefore already had dugouts made of yellow pine and called something like *tomolo* by ca. 2000 BC. That places the origin of the word thousands of years before the arrival of humans in Polynesia, and before the earliest evidence of planked boats in California. Since dugouts appeared millennia before the first evidence of deep-sea fishing in the area, their creation was motivated by some other needs, such as safety or increased cargo weight, in which they were superior to tule boats. Lastly, when the ancestors of the modern Kawaiisu arrived at their present location, perhaps about AD 1000 (Sutton 2010a), they borrowed the Kitanemuk word for the Buena Vista tule boats, *kwiakt*, as *kwiajakata*.

In the model given here, the plank canoe was innovated in southern California because of the increased need for large, seaworthy boats which could frequently travel to the offshore islands of the channel, coupled with the lack of accessible trees suitable for building large dugouts. This model explains the rarity of sewn-plank canoes in the Americas, and helps explain their distribution in the rest of the world.

Acknowledgments

I thank Justin Spence for his help at the Survey of California and Other Indian Languages archive in Berkeley. For answering questions and for other discussions, thanks to Ken Hill, Andrew Garrett, Jeanne Arnold, Andy Pawley, Jane Hill, Juliette Blevins, Matthew Spriggs, Pamela Munro, Paul Geraghty, Theresa Donham, Gary Breschini, Adrienne Adams, Chester King, and Irina Nikolaeva. I particularly thank Puakea Nogelmeier for reviewing and offering corrections to my transcriptions of the Hawai’ian materials. I have benefited from the use of a pre-publication version of Golla (2011). Special thanks to Jaina Davis.

References


---

96 It is possible in principle that they had originally another word for the tule boat and replaced it with *kwekt–kwiakt* later, but there is no clear reason or motivation for such a replacement.


Arroyo de la Cuesta, Felipe. 1837. Lecciones de Yndios [...] Quaderno de Lengua de San Luis Obispo, de la Purisima, y de Santa Ynes, con las notas gramaticales correspondientes a este idioma prepositivo y preverbal, que no tiene analogia alguna con el Mutsun ni con el Juncaleño. Manuscript. BANC MS C-C 63a. Berkeley, CA: Bancroft Library.


The Plank Canoe of Southern California


Honolulu.


Costansó, Miguel. 1770. *Diario histórico de los viages de mar, y tierra, hechos al norte de la California...* Mexico City: Imprenta del Superior Gobierno.


The Plank Canoe of Southern California


Yoram Meroz


The Plank Canoe of Southern California


Joslin, Terry L. 2010. Middle and Late Holocene Hunter-Gatherer Adaptations to Coastal Ecosystems Along the Southern San Simeon Reef, California. Ph.D. dissertation (anthropology), University of California, Santa Barbara.


The Plank Canoe of Southern California


The Plank Canoe of Southern California


The Plank Canoe of Southern California


Merriam, C. Hart. 1903c. *[Mo-kal’-um-ne Vocabulary, 1903]*. Manuscript. C. Hart Merriam Papers V211/V89 (microfilm reel 43), Bancroft Library, Berkeley.


The Plank Canoe of Southern California


The Plank Canoe of Southern California


yoram.meroz@gmail.com
1 Introduction

The purpose of this study is to describe and account for non-templatic and variable affix order in Kuna, a Chibchan language spoken in Panama. The majority of Kuna verbal affixes can be ordered transitively, meaning that a fixed ordering or template can be constructed such that each affix is positioned at a certain morphological distance from the verb. The affixes fall into two major groupings: inner affixes, which appear nearer the verb root, and outer affixes, which appear farther away. However, some complications are introduced by the future affix -oe which straddles the boundary between the two groupings. An exceptional group of affixes (-oe ‘FUT’, -suli ‘NEG’, -mala ‘PL’, -moga ‘also’, -bali ‘again’) is able to produce adjacency pairs with not only the orders -a-b and -b-c but also -c-a. This property will be referred to as non-transitive ordering; it is what makes a template model untenable. Furthermore, three-affix combinations from the exceptional group exhibit variable affix ordering, such that there are two possible grammatical orderings of a set of three affixes (-a-b-c and -c-a-b). This is not attributable to dialect differences, since a single speaker has been recorded using both orderings at different points in the same conversation. The question of which two orderings are grammatical is not completely predictable from the data on two-affix combinations. This unpredictability is referred to as non-cumulativity and makes an analysis based purely on adjacency pairs untenable.

In both the unexceptional and the exceptional groups, the ordering of adjacent affixes is very important to determining grammaticality. The proposed analysis incorporates the importance of adjacency by using constraints on ordered pairs of adjacent affixes, or bigrams, following Fabb (1998) and Ryan (2008). A grammatical word must be constructed of a series of grammatical ordered pairs, or bigrams. A set of inviolable constraints against the ungrammatical bigrams accounts for the ordering of all unexceptional affix combinations, as well as all of the two-affix combinations of exceptional affixes (including those exhibiting non-transitive ordering). The remaining data, which consists of three-affix combinations exhibiting non-cumulativity, is only partially explained by the inviolable bigram constraints. In order to explain non-cumulativity, two hypotheses are proposed. The first, based on ideas from Hay and Baayen (2002), is that frequency of bigrams in a text corpus could correlate with grammaticalization of the bigrams into units, such that three-morpheme combinations will preferably contain frequent (or grammaticalized) bigrams. However, a corpus study fails to support this hypothesis. Non-cumulativity is better accounted for using a low-ranked alignment constraint at the boundary between inner and outer affixes.

Finally, the question remains as to how a speaker in a given conversational context chooses between two different orderings of the same string of three affixes. It is proposed that this decision is based on presence or absence of focus on the negative. Focus in this sense refers to a contrast between alternatives, as proposed by Krifka (2007). Data from conversations suggests
that when there is an overt contrast between the negative statement and a positive counterpart, the speaker is more likely to use the \textit{NEG}-final ordering. This could be because the \textit{NEG}-final ordering allows the stress to fall on the negative morpheme, because of priming effects from the positive version of the verb form, or because the word-final position makes the negative morpheme more salient.

Section 2 presents data on affix ordering in Kuna, including background information on the inner and outer affix zones and the templatic behavior of the majority of affixes. Section 3 presents data on the exceptional behavior of a certain group of affixes, which would require a template or level-ordering model to allow looping. Section 4 discusses morpheme-specific selectional restrictions and formalizes the selectional restrictions account using OT bigram constraints, showing that most of the data can be accounted for using bigram constraints on ungrammatical bigrams, while the remainder can be accounted for with low-ranked bigram constraints on grammatical bigrams. Section 5 presents data from the corpus study on bigram frequencies, which shows that there is little evidence for the low-ranked bigram constraints. Section 6 revises the OT analysis by replacing the low-ranked bigram constraints with an alignment constraint justified by data from Section 2. Section 7 is the conclusion. The appendix presents conversational data on how affix order is affected by focus on the negative.

2 \textbf{Data on Templatic and Non-templatic Affix Ordering in Kuna}

2.1 \textbf{Inner and Outer Affixes in Kuna}

Kuna affixes can be divided into two zones or groups; the inner affixes appear closer to the verb root and attach only to a verb stem or other inner affix, while the outer affixes appear farther from the verb root and attach to a wider variety of hosts. In examples (1)-(4), all affixes appearing before the ‘=’ sign are inner affixes, while all affixes appearing after it are outer affixes.\footnote{All data are from naturally-occurring conversations recorded on the date given, unless otherwise noted.}

\begin{enumerate}
\item \texttt{gun-bi=sur-moga-d}\footnote{The rhetorical suffix \textit{-de}, \textit{-d}, glossed RHET, has very little semantic content; it is used for rhetorical effect.} \texttt{eat-want=NEG-also-RHET}
\begin{flushright}
‘I don’t want to eat any.’ (June 26, 2008, rec. 11 5:29)
\end{flushright}
\item \texttt{ebu-sa=mo-ye}\footnote{The Kuna orthography used here is mostly similar to the IPA, but /y/ is pronounced [j] and /ch/ is pronounced [ʧ]. The symbols /b,d,g/ are used for single (voiced/voiceless) stops, while /p,t,k/ are used for double (voiceless) stops.} \texttt{touch-PFV=also-QUOT}
\begin{flushright}
‘“[God] touched [your heart] also” (I said).’ (June 29, 2008, rec. 1, 2:03)
\end{flushright}
\item \texttt{immak-nai=mo-soge}\footnote{Kuna /g/ or /k/ before a vowel is pronounced /y/ ([j]). Kuna /ss/ is pronounced /ch/ ([ʧ]).} \texttt{do-PROG=also-say}
\begin{flushright}
‘“[You] are doing [it for free]” (I said).’ (June 29, 2008, rec. 1, 2:01)
\end{flushright}
\end{enumerate}
However, because the exceptional affixes do not always respect the inner/outer boundary, the two groups shall be defined morphologically, as follows:

a. inner affixes: attach to verb stem or other inner affixes
b. outer affixes: attach to any lexical root (verb root, verb stem, noun, adjective), inner affixes, or other outer affixes

For most Kuna verbs, the verb stem is identical to the verb root. However, there is a class of verbs in which the verb root needs to be combined with another morpheme to form a verb stem; these are called verbs with defective roots by Erice (1980:125) and are also noted as “regular verbs composed using gue” (the copula) by Puig (1946:95). Inner affixes may attach only to a verb stem, while outer affixes can attach directly to the root. In example (5), the outer affix -moga attaches directly to a defective root. In example (6a), the inner affix -sa cannot attach to a defective root; rather, it must attach to a verb stem, as shown in (6b). The inner/outer boundary is again marked using the symbol ‘=’.

(5) a. dob=moga
    fear=also
    ‘is also afraid’

(6) a. *dobsa
    b. dob-gu-sa
    fear-COP-PFV
    ‘got scared’

The defective roots therefore can be used as a test to classify affixes as either inner affixes (such as -sa) or outer affixes (such as -moga).

Outer affixes may also attach to different categories of words, such as adjectives or nouns. Sherzer (2003:17) notes that -mala ‘PL’ and -moga ‘also’ can attach to both verbs and nouns. In example (7), the outer affix -mala, ‘PL,’ attaches to a noun. In example (8), it attaches to an adjective.

(7) santos=mala
    saints=PL
    ‘saints’
    (June 26, 2008, rec. 11, 14:15)

(8) iglesia ollo=mala-d-de
    church empty=PL-RHET-RHET

---

5 Examples (5) and (6) have no recording; they are constructed as representative examples of defective verbs.
6 Cf. daksa (non-defective), morphologically dake-sa ‘see-PFV’.
7 This is the standard way of pluralizing borrowed nouns, adding the Kuna plural marker after the Spanish plural marker.
‘The churches [were] empty.’ (June 26, 2008, rec. 11, 15:41)

Likewise, in example (9), the outer affix -suli ‘NEG’ attaches to an adjective, and in (10) it attaches to a possessive adjective.¹⁸

(9) saila war-gwena mullu=suli
hair CL-one short=NEG
‘One had long hair.’ (June 26, 2008, rec. 11, 17:59)

(10) we bebida ani-ga-d=suli-n-dake-n-ye
that drink 1-DAT-NMLZ=NEG-IMPV-see-IMPV-QUOT⁹
‘“That drink wasn’t for me,” (she said).’ (June 26, 2008, rec. 11, 9:29)

Some outer affixes may also appear as independent words, including suli ‘no’, bali ‘again, back’, and sunna ‘truly, able to,’ but not including -moga ‘also’ or -mala ‘PL.’

The inner and outer affix groups might be roughly defined as affixes and clitics respectively. The inner affixes generally conform to Zwicky and Pullum’s (1983) definition of affixes, being more selective regarding their bases, only attaching to other affixes, and showing some ‘arbitrary gaps’ in which affixes they can attach to.¹⁰ They also exhibit some semantic idiosyncrasies.¹¹ The outer affixes are less selective about their bases and appear to show no arbitrary gaps or semantic idiosyncrasies. To avoid confusion between the cross-linguistic characteristics of affixes and clitics and the language-specific characteristics of the two groups of affixes, they will continue to be called inner and outer affixes.

The existence of the affix -oe ‘FUT’, which has some characteristics of both the inner and the outer affixes, suggests the two groups of affixes occupy different areas on a continuous scale from more selective to less selective affixes. There appears to be a cross-linguistic tendency for affixes that appear closer to the verb root to be more selective regarding their hosts, and for affixes appearing further away to be less selective. This tendency in English morphology has been discussed by Hay and Baayen (2002) and Hay and Plag (2004). The distinction between inner and outer affixes in Kuna follows this cross-linguistic tendency of ordering more selective affixes inside less selective affixes.

2.2 Templatic Behavior in Kuna Affixes

The distinction between inner and outer affixes gives us some insight into the structure of the Kuna verb. However, it gives no information about ordering within each group. This section

---

¹⁸ The possessive adjective is formed from the dative and nominalizer, forming the meaning ‘that which is for (me).’

⁹ The three-affix sequence -dakenye has been grammaticalized as a quotative.

¹⁰ There are some affixes that only appear with certain roots, and some roots that cannot appear with certain affixes. The affix -dapi ‘continuing on’ rarely or never appears with any root other than nae ‘go,’ although the entire string -nadapi may appear after another root x, to express the meaning ‘going along x-ing.’ The affix -ali, ‘PFV, towards the speaker’ obligatorily replaces the perfective -sa in the verb barmie, ‘send,’ but cannot be used with other verbs. An affix -ali on other verbs has the meaning ‘for the first time.’

¹¹ For example, returning to the affix -dapi, although the affix usually means ‘going along,’ when combined with the future affix -oe, it gives us the meaning ‘move towards 2nd person;’ the standard way of telling someone that you will visit them in their home is an bese nadapoe ‘I will come/go visit you.’ Here, the meaning ‘going along’ is no longer operative.
shows how a fixed partial ordering model such as a template or level-ordering system can explain a large portion of Kuna affix ordering. However, section 3 will show why this type of model is inappropriate for the system as a whole.

Because many Kuna affixes appear in a fixed order, it seems logical to attempt to fit them into a position-class or template model (Kari 1989, Hyman 2003). In a fixed system of affix ordering, each affix is permanently located in a morphological position that remains constant relative to other affixes in the system. Fixed ordering systems are transitive in that if one affix is ordered before another, then it will always come earlier in the word, regardless of any intervening material. Positions are sometimes grouped into zones defined semantically or phonologically (Kari 1989). Words are constructed such that affixes located in earlier positions always come before the affixes in later positions, regardless of any intervening material. In most template models, affixes assigned to the same position class are in complementary distribution; however, there are some models in which affixes in the same position class may appear adjacently with variable ordering.

Kiparsky’s (1982) theory of level ordering is in a sense another type of fixed-ordering system. It is similar to template morphology in that it assigns each affix to a level (somewhat similar to a position class). The derivation of the word proceeds linearly through the levels. Even if no affix is attached at a given level, the word continues to proceed to the next level, where another affix may be added. However, a word may not add more than one affix in a given level or proceed backwards through the levels. Because of this, level ordering is transitive in the same way that templates are. However, unlike the position classes in a template, each level is associated with different phonological rules, which apply to the entire word in the same order as the levels. Furthermore, bracket erasure allows more affixes to be added after the category of the word has been changed by a category-changing affix; a phenomenon that occurs in Kuna.

Sherzer (1989, 2003) proposes a seven-position template for Kuna (Sherzer 2003:12). It appears that in Sherzer’s model, affixes from a single position class can co-occur. In (11), an abridged version of Sherzer’s (1989:266) template is shown, including only affixes that appear in this paper. Each position contains a list of affixes; there is no significance to the vertical location of the affix in the list.

<table>
<thead>
<tr>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>mai, PROG</td>
</tr>
<tr>
<td>gwichi, PROG</td>
</tr>
<tr>
<td>sii, PROG</td>
</tr>
<tr>
<td>nai, PROG</td>
</tr>
<tr>
<td>bukwa, PROG</td>
</tr>
<tr>
<td>dii, PROG</td>
</tr>
<tr>
<td>lege, PASS</td>
</tr>
<tr>
<td>sa, PFV</td>
</tr>
<tr>
<td>o, FUT</td>
</tr>
<tr>
<td>mala, PL</td>
</tr>
<tr>
<td>moga, ‘also’</td>
</tr>
<tr>
<td>bali, ‘again’</td>
</tr>
<tr>
<td>suli, NEG</td>
</tr>
<tr>
<td>le, COND</td>
</tr>
<tr>
<td>diba, COND</td>
</tr>
<tr>
<td>sunto, ‘truly’</td>
</tr>
<tr>
<td>ye, QUOT</td>
</tr>
</tbody>
</table>

12 The orthography and some of the glosses have been changed to match the conventions of this paper. In the original, -le is labeled ‘clause connector’ and -diba ‘modality.’ In the original chart, Sherzer has -sunto instead of -sunna for ‘truly.’ I consider the sequence -sundo a combination of -sunna ‘truly’ and -do ‘then.’

13 The choice of progressive in position 1 indicates the posture of the subject and other attributes; from the top, the progressives are lying, standing, sitting, hanging/ floating/default, collective, moving/default.
emphasis on affix order. In order to make progress towards the goal of accurately predicting affix order, this section proposes a fixed ordering system that expands Sherzer’s positions 2 and 3 into five distinct positions. The position-1 affixes listed above will be located in two separate positions; if all of Sherzer’s position-1 affixes were included, they would generate even more positions.

In (12), a revised ordering of a group of common inner affixes is shown. These inner affixes exhibit fixed ordering, with the passive -lege appearing closest to the root, followed by the perfective -sa, and then the progressive affixes. The progressive affixes are considered to be at the same position or level in the fixed ordering, and they are in complementary distribution.

(12)  IN₁  IN₂  IN₃
      -lege -sa  -nai
      PASS  PFV  -gwichi
      -sii
      -mai
      -dii
      -bukwa
      PROG

Example (13) shows the ordering -lege-sa, and example (14) shows the ordering -sa-bukwa.

(13)  immak-le-sa
      make-PASS-PFV
      ‘[It] was made.’     (June 26, 2008, rec. 11, 10:33)

(14)  desayuno immak-s-bukwa⁽¹⁴⁾
      breakfast make-PFV-PROG.COLL
      ‘[They all] had breakfast.’     (June 26, 2008, rec. 11, 5:44)

Similarly, a subset of the outer affixes can be organized into a fixed ordering, as shown in (15). However, for this group of affixes, the transitive ordering pattern is only reliable as long as certain other affixes do not intervene; this issue is discussed in section 3.

(15)  OUT₁  OUT₂  OUT₃  OUT₄
      -suli  -mala -bali  -sunna
      -moga
      NEG  PL  ‘again’  ‘truly’
      ‘also’

Examples (16)-(17) show the ordered pairs mala-bali, suli-moga, mala-moga, and suli-bali.

(16)  sii=mar-bali-d
      sit=PL-again-RHET
      ‘[Those other ones] are sitting there too.’     (June 29, 2008, rec. 1, 3:50)

⁽¹⁴⁾The progressive morpheme -bukwa has the additional meaning ‘in a group’ which is abbreviated ‘COLL’ for collective.
Variable Affix Ordering in Kuna

(17) gun-bi=sur-moga-d
eat-want=NEG-also-RHET
‘I don’t want to eat any.’ (June 26, 2008, rec. 11, 5:29)

(18) gu-dii-dii=mar-mo-ye
COP-prog-prog=PL-also-RHET
‘They are also [just like me].’ (June 26, 2008, rec.11, 3:06)

(19) gun-sa=sur-ba
eat-PFV=NEG-again
‘[We] didn’t eat [that chicken] either.’ (July 2, 2008, rec. 1, 5:20)

Example (20) demonstrates the ordering of -sunna after -moga.

(20) bato  sunmak-de=mo-sun-do
already speak-INCP=also-truly-RHET
‘[They] are already starting to talk too.’ (June 26, 2008, rec. 11, 19:40)

For this particular group of outer affixes, the ordering is fixed, as shown in (15). As long as the word contains only the outer affixes in (15), the negative -suli will always come before -bali or -moga, regardless of whether -mala is present or not.

The two partial orderings are combined in (21). (Please see (12) and (15) for glosses).

(21) inner affixes = outer affixes
IN1  IN2  IN3 = OUT1  OUT2  OUT3  OUT4
-lege -sa -nai = -suli  -mala -bali -sunna
-gwichi -moga
-sii
-mai
-dii
-bukwa

This combined ordering dependably produces grammatical words: if one begins with a stem and adds at most one affix from each position, proceeding from left to right, all of the words constructed will be grammatical. Furthermore, the inner/outer distinction suggests morphological zones similar to those found in templates (Kari 1989). In this sense, the fixed ordering functions as a template. However, this system will only function as long as the negative affix -oe, which participates in the exceptional ordering group, is not included.

3 Exceptional Ordering Involving the Future -oe

The future affix -oe disrupts any possible template model of Kuna because it is impossible to assign the affix to any single position in the proposed template. The presence of this affix

---

15 gun-sa is pronounced /gucha/.
16 The pair bali-moga or moga-bali appears to be unattested; the two affixes are in complementary distribution. It is possible that this is the result of the semantic similarity of -bali ‘again’ and -moga ‘also.’
triggers non-transitivity and non-cumulativity in affix ordering, properties that cannot be described using any type of fixed-ordering analysis.

3.1 The Future -oe and Non-transitivity

The affix -oe is an inner affix as defined by its inability to combine with defective roots. However, unlike most inner affixes, it can combine with roots that are not verbs, as shown in (22).\(^\text{17}\)

\[
(22) \quad \text{war-gwen nap-oe} \\
\quad \text{CLF.long-one more-FUT} \\
\quad \text{‘[There] will [be] one more.’} \quad \text{(unrecorded conversation, Jan. 6, 2009)}
\]

Usually, -oe appears after other inner affixes and before outer affixes. Example (23) illustrates the ordering -lege-oe, where -oe appears immediately before the outer affixes.

\[
(23) \quad \text{uk-leg-o=dake-n-ye} \\
\quad \text{give-PASS-FUT=see-IMPV-QUOT} \\
\quad \text{‘“It will be given out, distributed,” (they say).’} \quad \text{(June 29, 2008, rec. 1, 1:47)}
\]

Examples (24) and (25) show -oe appearing before the negative -suli, an outer affix of the group OUT\(_1\). As expected, -oe appears right before the inner-outer boundary.

\[
(24) \quad \text{nu-gu-o=sur-iye} \\
\quad \text{good-COP-FUT=NEG-QUOT} \\
\quad \text{‘“[He] won’t get better,” (people told me).’} \quad \text{(June 29, 2008, rec. 1, 26:40)}
\]

\[
(25) \quad \text{endarb-oe=sur-sun-dake-n-ye} \\
\quad \text{wait-FUT=NEG-truly-see-IMPV-QUOT} \\
\quad \text{‘“[I] really won’t wait [for you],” (she said).’} \quad \text{(July 2, 2008, rec. 1, 1:00)}
\]

As expected, the affix -oe also appears before outer affixes that come later in the word, such as -dibe, as shown in (26).

\[
(26) \quad \text{sog-dag-o=dibe-ye} \\
\quad \text{say-come-FUT=COND-QUOT} \\
\quad \text{‘“[What] will [she] say [to me]?” (I wondered).’} \quad \text{(July 2, 2008, rec. 1, 0:03)}
\]

However, the future affix -oe also appears after certain outer affixes, even though it is an inner affix. The unusual behavior of -oe has been noted by Sherzer (1989:265): “when the future affix oe [position 2] occurs along with pali, mala, or moka [position 3], it follows them.” Examples (27) and (28) show how -oe can come either before or immediately after the plural -mala in different contexts (it can never appear immediately before -mala). The symbol ‘\(=\)’ is

\[\text{17} \quad \text{The suffix -oe is also unusual in other ways, being one of very few vowel-initial suffixes, and the only suffix comprised entirely of vowels. While stress in verbs usually falls on the penultimate syllable, it never falls on -oe, and may shift to the antepenultimate syllable in order to avoid falling on -oe.}\]
still placed before the first outer affix, even if -oe appears after an outer affix.

(27)  dak-o=sur-ma-riye
      see-FUT=NEG-PL-QUOT
      ‘“[You]-PL won’t see [him] anymore,” (he said).’    (June 29, 2008, rec. 1, 23:16)

(28)  oyo-na=ma-o-ge
      show-go=PL-FUT
      ‘[They] will go show [the place to you].’    (June 29, 2008, rec. 1, 37:08)

This effectively destroys the fixed-ordering system described in section 3, because there is no way to fit the affix -oe into the fixed partial ordering shown in (26). The affix -oe is occupying two different non-adjacent positions in the template, making the ordering non-transitive, which contradicts the basic structure of template morphology.

The affixes -moga ‘also’-, -oe, and -bali ‘again’ behave similarly to -mala in terms of their ordering with respect to -oe. Examples (29) and (30) illustrate the behavior of -moga and -oe:

(29)  na=mog-oe-na
      go=also-FUT-IMPV
      ‘They were also going to go.’    (June 29, 2008, rec. 1, 12:48)

(30)  gu-o=sur-mo-ye
      become-FUT=NEG-also-QUOT
      ‘“[You] will not [get yourself tied up] either,” (he said).’    (Kungiler 1997:35)

Examples (31) and (32) illustrate the pair -bali, -oe. Note that the sequence of morphemes in example (32) is rare in conversation because the same meaning is often expressed using bali as a separate word preceding the verb, especially when -suli is present (see examples 54, 59, 60, and 64 in the appendix). Example (32) is repeated with additional context as example (65) in the appendix.

(31)  dani=mar-bal-o-ye
      come=PL-again-FUT-QUOT
      ‘“[We] will come back,” (they said).’    (June 26, 2008, rec. 11, 8:41)

(32)  na-o=sur-bali-ye
      go-FUT=NEG-again-QUOT
      ‘“I won’t go either,” (I said).’    (Aug 15, 2010, rec. 4, 4:35)

Example (31) also illustrates how -oe can appear after the combination -mala-bali. In general, the affixes -mala, -bali, and -moga have the same ordering properties with respect to -oe, and when two of them appear together, they appear adjacent to each other in the order shown in (15)/(21), regardless of the position of -oe.

The affix -sunna may appear either immediately before or immediately after -oe, and it always appears last when combining with -mala, -moga, or -bali. In example (33), it appears after -mala and before -oe. Note that the outer affix -mala is attached to a non-verbal root.
(33)  
\[
\text{bane}=\text{mar-sun-oe} \\
\text{tomorrow}=\text{PL-truly-FUT}
\]

‘[See you] tomorrow, for sure.’  
(unrecorded conversation, Jan. 6, 2009)

In example (34), -sunna appears after -oe:

(34)  
\[
\text{iki an totok-o=\text{sun}}^{18} \\
\text{how 1ST play-FUT=\text{truly}}
\]

‘How should I dance?’  
(Sherzer 2004:34)

The ordering of -sunna with respect to -oe is unusual for Kuna, in that the two affixes can appear adjacently in either order. However, cross-linguistically, variable ordering of adjacent affixes is relatively common, and it is easier to account for than variable ordering between non-adjacent positions. The ordering of -sunna and -oe can be explained with a fixed partial ordering simply by locating the two affixes in the same position class, although -oe will still need to appear in two positions for other reasons. For this reason, the variable ordering of -oe and -sunna does not pose as much of a problem as does -oe’s interaction with -mala, -moga, and -bali.

While some template models tolerate variable ordering, it is only possible within one zone or between adjacent zones. In the Kuna situation, variable ordering does not even respect the boundaries between the two major zones of the proposed template. Aside from the non-transitive ordering, there are also some other reasons why a fixed ordering may not be the most appropriate model. The principal attraction of describing certain languages in terms of templates is that templates model blocking within position classes. Blocking in this context refers to situations in which semantically compatible affixes cannot co-occur. This phenomenon can be explained by assigning the two affixes to a single position class, such that only one affix in the class can appear in any given word. It effectively explains complimentary distribution of affixes, as long as they are not separated by intervening template positions. Kuna has little or none of this type of blocking. Affixes that “block” each other in Kuna generally do so for semantic reasons; for example, only one of the six progressive affixes may appear in any given word. The template hypothesis offers us little advantage in this respect.

A level ordering analysis would not be the best fit either. The principal attraction of a level ordering analysis (Kiparsky 1982) is that it can be used to correlate affix order with independent features such as phonology. In the case of Kuna, there are no phonological reasons for assigning affixes to a given level; affixes must be assigned to levels based on their ordering properties alone.\(^{19}\) Furthermore, the non-transitive ordering poses the same problem that it does for the template model, in that it would require looping between levels. Another potential problem with

---

\(^{18}\) Sherzer’s orthography uses single /p,t,k/ for single (voiced/voiceless) stops.

\(^{19}\) Kuna affixes are vowel-final, and the final vowel is deleted before the addition of another affix. If the process results in a three-consonant cluster, then the consonant preceding the deleted final vowel is deleted as well. Some affixes have irregular short forms, where the entire final syllable is deleted even though no consonant cluster is being avoided. Four affixes never undergo truncation: the perfective inner affix -sa, the outer affix ‘possibly,’ -dibe or -diba, the conditional outer affix -le and the quotative outer affix -ye. In the case of -dibe/-diba, the lack of truncation may be in order to avoid homophony or near-homophony with the nominalizer -di. In the other cases, the lack of truncation appears to be due to the affixes’ unusually reduced phonological form. The last two of the group, -le and -ye, not only resist truncation but also fail to trigger truncation of immediately preceding affixes. The imperfective -na also fails to trigger truncation.
using level ordering for the Kuna data is bracket erasure. According to bracket erasure, an affix attaches to the word as a whole, without reference to specific features of affixes within the word. Since Kuna affixes appear to be very sensitive to adjacency, this may pose some problems. Some affixes, especially within the inner affix group, may attach to one affix in a given position class, but not another. For example, an alternative perfective marker -de only appears on the verb nae, ‘go,’ and not on other verbs. Kuna morphology has too many morpheme-specific conditions to be modeled using rules that can only refer to classes of words or affixes.

While some aspects of Kuna morphology are describable using a template or level-ordering analysis, these theories cannot adequately explain its non-transitive ordering properties. At the same time, the Kuna data does not contain the types of complications that these theories are best suited to explain.

3.2 The Exceptional Affixes and Non-cumulativity

The variable ordering triggered by -oe is also non-cumulative. Here, cumulativity means that a word made up of grammatical pairs of affixes (bigrams) is a grammatical word. In other words, if a word stem-a-b is grammatical and a word stem-b-c is grammatical, then the word stem-a-b-c is also grammatical. Furthermore, if the word stem-b-c is grammatical and the word stem-c-a is grammatical, then the word stem-b-c-a is also grammatical.

For many groups of affixes in Kuna, such as those shown in the partial ordering in (21), cumulativity holds. However, cumulativity breaks down when the affix -oe is involved. In (35), a summary of the grammatical adjacency pairs is shown:

(35)  -oe-suli
     -suli-mala, -suli-bali, -suli-moga
     -mala-oe, -bali-oe, -moga-oe

If Kuna were a cumulative system, then all of the following orders would be grammatical:

(36)  a. stem-oe=suli-mala    stem-oe=suli-bali    stem-oe=suli-moga
     b. stem=mala-oe-suli    stem=bali-oe-suli    stem=moga-oe-suli
     c. *stem=suli-mala-oe   *stem=suli-bali-oe  *stem=suli-moga-oe

However, the orders in (c) are ungrammatical. Orders (a) and (b) are illustrated in (37), repeated from (27), and (38).

(37)  dak-o=sur-mar-ye (order a)
     see-FUT=NEG-PL-QUOT
     ‘“You-PL won’t see [him] anymore,” (he said).’ (June 29, 2008, rec. 1, 23:16)

(38)  na=mal-o-suli (order b)
     go=PL-FUT-NEG
     ‘[The children] won’t go [to school tomorrow].’ (July 3, 2008, rec. 2, 7:08)

Although the corpus is limited, it appears from both the corpus and the elicitations that order (a) is used more often than order (b). While neither form is frequent in the corpus, order (b) is not used at all in the published texts. In the translation elicitations, some consultants used only order
(a), while all of those who used order (b) also used order (a). In the conversational elicitations, order (a) was used more frequently (see appendix). Order (c) is so far unattested in the corpus and was not produced spontaneously in elicitation sessions. However, a minority of consultants in elicitation sessions accepted order (c) when I produced a word using this order. This means that order (c) is more grammatical than, for example, order (d) \*stem-oe-mala-suli, which contains no grammatical ordered pairs. There seems to be a continuum of grammaticality a > b > c > d, where (a) is at the top, and (c) and (d) seem to be below a threshold of utterability. According to Coetzee (2008), the grammatical words of a language may differ in their degree of grammaticality, where the more highly grammatical words violate fewer constraints than the marginally grammatical words. Using a response-time experiment, Coetzee shows that the less grammatical words have longer processing times. Although there is no data on processing in Kuna, the frequency of bigrams may be correlated with relative grammaticality and frequency of bigrams.

The fact that (c) is ungrammatical makes the ordering system non-cumulative. This non-cumulativity is demonstrated using finite-state diagrams constructed from the bigrams listed in (35). In order to form a grammatical word, begin at the stem and follow any arrow. Continue following any of the arrows from one node to the next, stopping at any point, but never going back or jumping between nodes. If the Kuna system were cumulative, the diagram would show a closed loop of affixes, as in Figure 1.

![Figure 1: Hypothetical language with cumulative ordering](image)

However, figure 1 does not describe Kuna. It can account for all of the data presented in section 2 (fixed partial ordering) and section 3.1 (non-transitive ordering), but it cannot account for non-cumulativity, incorrectly generating forms such as \*stem-suli-mala-oe. In Figure 2, this problem is solved by using more than one copy of a single affix.

![Figure 2: Hypothetical language with non-cumulative ordering](image)
Variable Affix Ordering in Kuna

Figure 2: Non-cumulative ordering in Kuna

Figure 2 is an effective way of determining which combinations of affixes are grammatical. However, the system is too powerful, in that it is essentially a way of dictating which strings are grammatical. Also, figure 2 does not capture the similarities in the data; instead, it uses redundant structures to model the same phenomenon in two different contexts.

To eliminate the copies and use figure 1, the system must somehow remember what path was taken to get to the present node, so that the word can be ended to avoid creating an ungrammatical string such as *suli-mala-oe. In other words, constraints must be non-local. This is not possible in a finite-state model.

4  Local Analyses of Kuna Affix Order

4.1  Selectional Restrictions

Setting aside non-cumulativity for the time being, non-transitivity can be accounted for using a local analysis that constrains only adjacent combinations of affixes rather than the structure of the verb as a whole. The analysis draws on research on English by Fabb (1998). Fabb (1998) argues that selectional restrictions are more effective at predicting English suffix ordering than is level ordering, which incorrectly predicts the grammaticality of a number of ungrammatical suffix strings. In Fabb’s analysis, the selectional restrictions are inviolable constraints on the behavior of individual suffixes. Some of the constraints determine whether or not a suffix can attach to an already-suffixed word, and some determine which major category a suffix can attach to. These two types of selectional restrictions explain the majority of the suffix-ordering data in English. The remaining data is explained by a group of selectional restrictions that refer to the identity of the individual suffix at the end of the base form. Fabb (1998:538) argues that his selectional restrictions make level-ordering “redundant,” since it does “no extra work” that is not accomplished by the selectional restrictions. Of course, a selectional-restrictions account is only preferable to the extent that it makes correct predictions about a larger portion of the data; if level ordering were equally effective, then the additional predictive power of the selectional restrictions would be unnecessary.

The bigram constraints proposed for Kuna are similar to the subset of Fabb’s selectional restrictions that are morpheme-specific, in that they determine whether or not a given affix may attach to a base that ends in a certain morpheme. Unlike Fabb’s analysis, the analysis of Kuna does not include broader selectional restrictions based on major category, although the low-ranked alignment constraint does refer to membership in the categories of inner and outer affixes. The selectional restrictions that are sensitive to affixes in the base are fundamentally local constraints, in that they only apply to the last affix in the base and do not make reference to any features of the base word as a whole.

Morpheme-specific selectional restrictions such as those in Fabb (1998) and the constraints proposed for Kuna are theoretically very different from the rules that govern affixation in a level-ordered system. In level ordering (Kiparsky 1982), bracket erasure effectively means that it is impossible for an affix to be sensitive to the presence of individual affixes in the base. Although the data from English suffixation shows that certain affixes attach to words ending in affix -a but not words ending in affix -b, this behavior is explained in terms of the base’s location at a given level as it proceeds through the level-ordering derivation. In order to introduce constraints that
are sensitive to specific affixes in the base, a serious modification to Lexical Phonology is necessary. Hammond (1992) introduces this change to the theory of affixation, naming it *morphemic circumscription*. According to Hammond, morphemic circumscription “allows affixes to be attached to a peripheral morphological constituent of the word” (1992:195). In other words, rather than selecting for an entire base, affixes can select for individual affixes within a base, provided that they are at the edge. Hammond justifies adding this power to current affixation theories by presenting various types of examples, including bracketing paradoxes and affixation that is contingent on prosodic characteristics of affixes within the base. Although Hammond does not propose using morphemic circumscription for non-transitive ordering effects, it is important for describing the Kuna data, because it allows for affixation that is sensitive only to peripheral affixes of the base. In this sense, morphemic circumscription is a local method of word-formation, similar to the morpheme-specific selectional restrictions of Fabb (1998).

In some ways, local selectional restrictions are more restrictive than global selectional restrictions such as those based on major category or derivational level. Each selectional restriction determines the order of a particular pair of affixes rather than simply providing a general guideline for which types of bases an affix can attach to. For a language that exhibits transitivity, any account based on global selectional restrictions can also be described using only morpheme-specific selectional restrictions. However, there are also ways in which a local account is less powerful than a global account, giving the language more room for variation. In particular, a local account allows for non-transitivity in affix ordering, because it does not make any statements about the ordering of affixes that are non-adjacent: if the string -a-b is grammatical and -b-c is grammatical, then a local account makes no statements about the grammaticality of -a-x-c. In terms of the Kuna data, this is an argument in favor of morpheme-specific selectional restrictions.

The data in section 3 has shown that it is impossible to build a fixed, transitive global ordering of the Kuna affixes. However, we can characterize the data in terms of the grammaticality or ungrammaticality of adjacent pairs of affixes. An ordered pair of affixes is considered grammatical if it is attested in any context, although it may be unacceptable in some contexts due to non-cumulativity. This may seem to weaken the claim that every pair is either grammatical or ungrammatical. However, as discussed in section 3.1, an ungrammatical word constructed completely of grammatical pairs is much more acceptable than any word in which an ungrammatical pair appears. This is an important point in favor of the importance of adjacency in Kuna affix ordering. For the purposes of this chart, a pair is considered grammatical if it appears in any grammatical word.

In table 1, the affixes listed on the vertical axis represent the first member of a pair, and the affixes listed on the horizontal axis represent the second member of a pair; the inner-outer affix boundary is indicated with a thicker line. The cell at the intersection of a column and a row represents the relevant ordered pair. Cells are marked with a /√/ to indicate that the ordered pair in question is attested. Cells left empty indicate that the pair is unattested; pairs may be unattested either for semantic or morphological reasons. Note that for a hypothetical language with perfectly transitive morphology, all of the cells above the diagonal would be checked, and all of the other cells would be empty, assuming that each affix is at a different level in the transitive ordering (in fact, -moga and -bali would be at the same level). To the extent that table 1 does not follow this pattern, a transitive or template model is inappropriate for describing Kuna

---

20 -nai and other progressive affixes are reduplicated when the subject is plural; -mala ‘pl.’ is optional in these forms.
In order to use the table to evaluate the grammaticality of a word, break the word down into ordered pairs of affixes and locate the cell corresponding to each of the pairs. If every cell is checked, then word is evaluated as grammatical. Using this method, the table correctly evaluates the grammaticality of all of the examples presented in sections 2 through 3.1. The table makes no statements about overall ordering, only about adjacent affixes. This means that words that exhibit non-transitivity are evaluated as grammatical, since they consistently contain the same ordered pairs. However, it incorrectly evaluates words that exhibit non-cumulativity as grammatical, namely words with orders of the type shown in (36c).

<table>
<thead>
<tr>
<th></th>
<th>lege</th>
<th>sa</th>
<th>nai</th>
<th>oe</th>
<th>suli</th>
<th>mala</th>
<th>moga</th>
<th>bali</th>
<th>sunna</th>
</tr>
</thead>
<tbody>
<tr>
<td>stem</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>lege PASS</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>sa PFV</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>nai PROG</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>oe FUT</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>suli NEG</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>mala PL</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>moga ‘also’</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>bali ‘again’</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>sunna ‘truly’</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 1

4.2 Bigram Constraints and the Unexceptional Affixes

In order to formalize the information in table 1 as a generative system, this section proposes a group of local constraints on adjacency. Markedness constraints on ordered pairs of adjacent suffixes have been proposed by Ryan (2008), who calls them bigram constraints. Following Fabb’s (1998) morpheme-specific selectional restrictions and Hammond’s (1992) morphemic circumscription, the bigram constraints determine whether affixation is possible based on the identity of the affix and of the final morpheme of the base. The constraint *x,y penalizes the appearance of a pair of adjacent suffixes, or bigram. The constraints are evaluated solely on the basis of adjacency without reference to positioning in the word. Given a constraint *x,y, a pair of suffixes -x-y located at the beginning of the word will receive exactly the same penalty as the

---

21 It makes no statements about some affixes that appear in the conversational data but are not important to our analysis.
pair \(-x-y\) located at the end of the word; while the sequences \(-x-a-y\) or \(-y-a-x\) will not be penalized. In Fabb’s (1998) analysis, the morpheme-specific constraints are used to explain residual data that is not covered by broader selectional restrictions. In contrast, for Kuna, morpheme-specific constraints account for the majority of the data. In this sense, this analysis is similar to that of Ryan (2008) who models suffix order in Tagalog using constraints on ungrammatical bigrams.

The bigram constraint analysis’s local property allows it to account for various types of non-transitivity in affix ordering. Ryan presents data from Tagalog on a suffix which can appear in either of two non-adjacent positions; this can be modeled by bigram constraints (Ryan 2008:18). Bigrams can also account for a “context-sensitive reorderability,” in which two affixes with the order \(a-b\) can also appear in the order \(b-a\), but only in certain morphological contexts (p. 19).

Bigrams can be used to account for “free variation within a fixed span,” in which a certain affix may appear in various different positions in an otherwise transitive ordering structure. This presents an improvement over alignment constraints, which can account for variable ordering only for adjacent affixes, and without context sensitivity.

Of these various ordering phenomena that can be modeled using bigram constraints, the one most similar to the Kuna data comes from suffix ordering in Quechua. Using the Quechua data, Ryan demonstrates how bigram constraints can model a three-suffix “loop,” similar to the Kuna loop shown in (35). The possible adjacent orderings in Quechua, taken from Muysken (1981), are given in example (39):

\[
\begin{align*}
(39) & \quad \text{a. -schi-naya (also -naya-schi)} \\
& \quad \quad \text{b. -naya-ru} \\
& \quad \quad \text{c. -ru-schi} & \text{(Ryan 2008:21)}
\end{align*}
\]

Ryan successfully models this situation using the constraint ranking in (45):

\[
\begin{align*}
(40) & \quad *-ru-naya, *-schi-ru >> *-naya-schi, *schi-naya & \text{(Ryan 2008:21)}
\end{align*}
\]

All of the ungrammatical forms can be ruled out using the two high-ranked constraints. The low-ranked constraints, which penalize grammatical orderings, are included in order to generate the variable results shown in (39a). Ryan’s bigram analysis effectively accounts for the non-transitivity found in Quechua.

### 4.3 Using Bigram Constraints to Model Non-transitivity

Bigram constraints are ideal for modeling the non-transitivity exhibited in the Kuna data, which is similar to the Quechua example. If the data is limited to words containing single bigrams, the analysis is very simple, using only inviolable constraints. Examples (41)-(43) are representative of the situation of non-transitivity:

\[
\begin{align*}
(41) & \quad \text{dak-o=suli} \\
& \quad \quad \text{see-FUT=NEG} \\
& \quad \quad \text{‘will not see’}
\end{align*}
\]

\[
\begin{align*}
(42) & \quad \text{dak=sur-mala} \\
& \quad \quad \text{see=NEG-PL}
\end{align*}
\]
Variable Affix Ordering in Kuna

‘do not see’

(43) dak=mal-oe  
    see=PL-FUT  
    ‘will see (PL)’

In order to account for these examples, the following constraints on adjacent affixes are necessary:

* suli-oe  
* oe-mala  
* mala-suli

The constraints are top-ranked, inviolable constraints. The very simple tableau 1, where all constraints are top-ranked, shows how these constraints account for example (41); examples (42) and (43) can be accounted for similarly. Other basic constraints are assumed to be active and top-ranked, including MAX, DEP, and ALIGNROOT-L.

<table>
<thead>
<tr>
<th>see (dake)</th>
<th>* suli-oe</th>
<th>* oe-mala</th>
<th>* mala-suli</th>
</tr>
</thead>
<tbody>
<tr>
<td>FUT (oe)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEG (suli)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) dak-o-suli</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) dak-sul-oe</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tableau 1

This shows that it is simple to account for the single-bigram (root-a-b) words using bigram constraints.

A system of bigram constraints can be built up to account for all other words with unexceptional orderings, such as all of the examples in section 2. In order to do this, the data shown in table 1 is used to generate inviolable constraints against all unattested adjacency pairs. This is a cumbersome system, but it can describe both the unexceptional pairings and the non-transitive pairings.

It is a little more complicated for bigram constraints to account for the three-affix words (root-a-b-c) that exhibit non-transitivity. The forms (44a-c) are representative of these words:

(44)  
  a. dak-o=sur-mala  
      see=FUT=NEG-PL  
      ‘will not see (PL)’
  
  b. dak=mal-o-suli  
      see= PL-FUT-NEG

22 Note that the input form for our OT tableau is semantic. The winning candidate must have some exponent of each of the semantic elements that appear in the input form (in other words, the constraint MAX-MEANING is top-ranked). If there is more than one affix that can represent the meaning, the input form does not specify which should be used. As it turns out, for the data in tableaux 1-3, there is only one possible exponent for each semantic element.
‘will not see (PL)’

c. *dak=sur-mal-oe
   see=NEG-PL-FUT
   intended: ‘will not see (PL)’

Because bigram constraints only apply to adjacent affixes, a word of the form root-a-b-c is evaluated as two separate bigrams, -a-b and -b-c. If all three trigrams shown in (44) were grammatical, it would be easy to account for the data using the bigrams introduced so far. However, (44c) is in fact ungrammatical in Kuna. This non-cumulativity distinguishes the Kuna data from Ryan’s Quechua data, and requires a different type of analysis.

4.4 Using Bigrams to Model Non-cumulativity

According to Ryan (2008), the bigram constraint system does not account for the type of non-cumulativity we have identified in the Kuna data in section 3.2: “bigrams cannot model a scenario in which a trigram is illformed but both its constituent bigrams are independently wellformed” (Ryan 2008:24). This statement is true for Ryan’s analysis, which uses Maximum Entropy Grammar instead of classic OT. Ryan’s use of weighted constraints and a smoothing factor means that all trigrams composed of grammatical bigrams are predicted to occur at least a very small percentage of the time. However, in a classic OT tableau, the statement quoted above no longer applies. This is because classic OT uses a categorical ranking system, so that candidates that violate high-ranked constraints are predicted to be completely unattested.

In tableau 2, the three high-ranked constraints are equally ranked, and the two low-ranked constraints are equally ranked. This ranking generates the two grammatical orderings (a) and (b) while ruling out the ungrammatical (c). The high-ranked group is comprised entirely of constraints against ungrammatical bigrams, and the low-ranked group is comprised of constraints against grammatical bigrams. Tableau 2 concerns only the three affixes -oe, -suli, and -mala. In order to handle data using -moga and -bali, the constraints used would be the same as those used in tableau 2, except that -mala would be replaced by -bali or -moga. In the case of words with more than one of the adverbial affixes, it would be necessary to add the top-ranked constraints *moga-mala, and *bali-mala.

<table>
<thead>
<tr>
<th>see (dake)</th>
<th>FUT (oe)</th>
<th>NEG (suli)</th>
<th>PL (mala)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>*sul-oe</td>
<td>*o-mala</td>
<td>*mala-suli</td>
</tr>
<tr>
<td>(a) dak-o-sur-mala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) dak-mal-o-suli</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) dak-sur-mal-oe</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>(d) dak-sul-o-mala</td>
<td>*</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>(e) dak-mar-sul-oe</td>
<td>*</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>(f) dak-o-mar-suli</td>
<td>*</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 2
Variable Affix Ordering in Kuna

This analysis accounts for the grammaticality of (a) and (b) and the ungrammaticality of (c). However, there is no independent confirmation of the existence of the low-ranked constraints on grammatical bigrams: why should *mala-oe and *suli-mala have an effect on the data, but not *oe-sul? Furthermore, there is no way of determining which of the two grammatical outcomes will surface in any given context. Assuming that the variation is random, then the situation can be modeled using stochastic OT (Bresnan, Dingare, and Manning 2001). According to this theory, as in traditional OT, constraints are ranked along a continuum from high to low, with higher constraints outranking lower ones. However, rather than occupying a particular point on the continuum, each constraint occupies a probabilistic cloud of locations, such that the constraint is most often ranked near the center of its range, but sometimes appears further away from the central point. Every time a word is generated, each constraint may appear at a slightly different location on the continuum. If two constraints are relatively close on the scale, then there is a low probability that the “lower” constraint will outrank the “higher” constraint in some instances of the grammar.

Using this theory to explain the data from Kuna, it would be necessary to propose that the constraints *-suli-mala and *-mala-oe are very near each other on the continuum, with the center of distribution of *-mala-oe ranked higher than that of *-suli-mala. In most instances of the grammar, *-mala-oe ends up being ranked higher; however, in some instances, when *-mala-oe is relatively low and *-suli-mala is relatively high, they cross. In these cases, *-suli-mala turns out to be ranked higher, allowing (b) to appear as grammatical. If these low-ranked constraints against grammatical bigrams are indeed driving the grammaticality of the two-bigram forms, then there may be evidence for the existence of these constraints in other areas of the grammar. The corpus study discussed in section 5 is an attempt to find such evidence.

5 Corpus Study

Section 4 raises the question of whether there is independent evidence for the existence of bigrams as a driving force in Kuna grammar. Hay and Baayen (2002) show that frequency can be used as an indicator of the productivity of an affix. They propose that if a given English suffix appears most often in highly parsable words, it is more productive. As a measure of parsability, they compare the frequency of a word with the frequency of its unaffixed base. They establish a threshold of relative frequency called the parsing line, based on comparing data on relative frequency to a study of response times. If the frequency of an affixed form is relatively high compared to the frequency of the unaffixed form, the word will fall below the parsing line, meaning that it is less likely to be parsed by speakers. Highly parsable words are words whose unaffixed base is relatively likely to appear as an independent form. The affixes of these words are those that are most likely to appear with a variety of bases. The study effectively establishes the idea that a frequency study of affixes can help us predict which base-affix combinations are likely to be grammatical.

There is no response-time data for Kuna, but frequency data is extractable from the corpus. The frequency of a bigram -a-b is compared with the frequency of words in which the two affixes appear in a different order, either -a-x-b or -b-x-a (in most cases -b-a is nonexistent). The hypothesis is that when two members of a bigram appear in a word, they are highly likely to appear as a bigram and not in some other order. This result will give support to the assertion that bigrams are important in Kuna grammar and that in some sense, they exist as composite units within the verb word.
The study also examines the relative frequencies of different bigrams in order to compare them to the ranking of bigram constraints in section 4.4. The top-ranked bigram constraints are inviolable and the corresponding bigrams are completely unattested. However, in a word with three affixes, the grammar has a choice of three different affix orders that all obey the inviolable constraints. In this case, the hypothesis is that the speaker will choose a word containing the more frequent bigrams. For a highly frequent bigram \(-a-b\), if \(-a\) and \(-b\) both appear in a given word, then they are highly likely to appear as the ordered bigram \(-a-b\) rather than in any other order.

The corpus study uses five published Kuna texts collected by Joel Sherzer (Sherzer 1998, 2004). All of the texts are spoken texts, and all but one are narratives. Kuna has various specialized registers whose vocabulary and phonology differ from those of the conversational language. Although it is unlikely that the affix order also differs, the survey is limited to narratives in the everyday register in order to avoid increasing the number of variables. The data includes all allomorphs and all senses of each affix. In the case of \(-bali\), homophonous morphemes glossed ‘to, towards,’ and ‘with’ were excluded; the other affixes do not have homophones. Affixes appearing as separate words were not included. The search was not limited to forms whose roots were verbs; however, the great majority of the roots are verbs.

Table 2 shows the number of tokens of each affix in the texts.

<table>
<thead>
<tr>
<th>morpheme</th>
<th>gloss</th>
<th>Text 1</th>
<th>Text 2</th>
<th>Text 3</th>
<th>Text 4</th>
<th>Text 5</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>-sunna</td>
<td>truly</td>
<td>126</td>
<td>86</td>
<td>22</td>
<td>78</td>
<td>99</td>
<td>411</td>
</tr>
<tr>
<td>-mala</td>
<td>PL</td>
<td>73</td>
<td>5</td>
<td>141</td>
<td>95</td>
<td>10</td>
<td>324</td>
</tr>
<tr>
<td>-oe</td>
<td>FUT</td>
<td>93</td>
<td>22</td>
<td>78</td>
<td>79</td>
<td>28</td>
<td>301</td>
</tr>
<tr>
<td>-suli</td>
<td>NEG</td>
<td>29</td>
<td>16</td>
<td>73</td>
<td>57</td>
<td>7</td>
<td>182</td>
</tr>
<tr>
<td>-bali</td>
<td>again, too, back</td>
<td>8</td>
<td>17</td>
<td>17</td>
<td>53</td>
<td>28</td>
<td>123</td>
</tr>
<tr>
<td>-moga</td>
<td>also</td>
<td>17</td>
<td>3</td>
<td>15</td>
<td>45</td>
<td>10</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 2: Corpus frequencies of selected Kuna suffixes

Table 2 shows that \(-oe\), \(-mala\), and \(-sunna\) are the most frequent of the six affixes in question. Note that \(-mala\) appears as the plural marker on many nouns. As for \(-sunna\), it appears more frequently in the corpus than it does in ordinary conversation, because the corpus is comprised as narrative texts, and the repeated use of \(-sunna\) is a common rhetorical device in narratives.

Table 3 shows data on all of the words that contain at least two of the affixes \(-oe\), \(-suli\), and \(-mala/-moga/-bali\). It combines the data for the affixes \(-bali\), \(-mala\), and \(-moga\), because they all have similar behavior with respect to \(-oe\) and \(-suli\). In the table, \(-bali\), \(-mala\), or \(-moga\) is represented by \(x\). Data on \(-sunna\), which has different ordering properties, is not included. The column marked ‘total’ shows the total number of words containing the two affixes in question, regardless of their position in the word. The column marked ‘adjacent’ shows the number of times that the affixes appear adjacent to each other. The adjacent pairs always appear in the order shown; there are zero tokens of adjacent pairs with reversed order for these three bigrams. The

---

23 The texts are The One-Eyed Grandmother from Sherzer (2004) and The Hot Pepper Story, Muristo Counsels a New Chief, The Report of a Kuna Medicinal Specialist, and The Agouti Story from Sherzer (1998). Muristo Counsels a New Chief is the non-narrative text included in this study.
column marked ‘non-adjacent, same order’ lists the number of times that the affixes appear in the same order, but separated by one or more affixes. For example, the pair -mala-oe may appear as -mala-moga-oe. The column ‘non-adjacent, reversed,’ lists the number of times that a pair of affixes appears non-adjacently in reversed order; for example, -mala-oe may appear as -oe-suli-mala (in fact, this is the only possible string where -mala and -oe appear in reversed order).

<table>
<thead>
<tr>
<th>bigram</th>
<th>adjacent</th>
<th>non-adjacent, same order</th>
<th>non-adjacent, reversed</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>-oe, -suli</td>
<td>10 (83%)</td>
<td>2 (17%)</td>
<td>0 (0%)</td>
<td>12</td>
</tr>
<tr>
<td>-suli, -x</td>
<td>16 (94%)</td>
<td>0 (0%)</td>
<td>1 (6%)</td>
<td>17</td>
</tr>
<tr>
<td>-x, -oe</td>
<td>47 (87%)</td>
<td>5 (9%)</td>
<td>2 (4%)</td>
<td>54</td>
</tr>
</tbody>
</table>

Table 3: Corpus frequencies of bigrams involving -oe, -mala/-moga/-bali (-x), and -suli

The most frequent combination is -mala, -moga, or -bali followed by -oe. Less frequent are -oe-suli and bigrams composed of -suli followed by any one of the affixes -mala, -moga, or -bali. If any two of these three affixes appear in the same word, they are highly likely to occur as the appropriate bigram -o-suli, -sur-mala, or -mal-oe. They are unlikely to occur non-adjacently in either order, and they never occur adjacently in reversed order. In this sense, the corpus study supports the hypothesis that bigrams are important to affix ordering.

However, in section 4.4, the violable constraints *mala-oe and *suli-mala were ranked higher than *-oe-suli. This would suggest that combinations like -mala-oe would be less frequent in the corpus, and -oe-suli more frequent. This seems to show that the corpus study does not provide support for the frequency hypothesis. One reason for the relatively high frequency of -x-oe could simply be that the combination contains both -oe and -mala which are the two highest-frequency affixes of those listed in table 2.

Note that a similar count of combinations of -sunna, -suli, and -oe shows similar results as for the other adverbials. The main difference is that in the case of -sunna and -oe, the reversed order is possible. There is no inviolable constraint against the combination -oe-sunna, so it occurs in about one-quarter of words containing the two affixes.

6 The Role of Focus and the Inner-Outer Affix Boundary

The data from the corpus study supports the idea that bigrams are important in Kuna morphology, but the frequency of the bigrams in the corpus does not provide any support for the use of low-ranking constraints against grammatical bigrams. The new analysis presented in this section differs from the previous analysis only in the area of the lower-ranked constraints. It retains the top-ranked constraints against ungrammatical bigrams, which still provide a perfect account of all of the single-bigram forms (root-a-b) as well as any longer forms that do not exhibit non-cumulativity. However, in order to better account for non-cumulativity in the double-bigram forms (root-a-b-c), the lower-ranked bigram constraints are replaced by an alignment constraint.

24 These non-adjacent, same order forms never appear for the pair -suli-mala because of the cumulative effect of a group of top-ranked bigram constraints. For all affixes -x there exists a top-ranked bigram constraint either *-suli-x or *-x-mala.
(45) ALIGNINNER-OUTER:
This constraint penalizes candidates that violate the inner-outer boundary by containing outer affixes that are ordered before (to the left of) an inner affix. The constraint is gradient: if one outer affix is ordered before an inner affix, the candidate receives one violation. If two outer affixes are ordered before an inner affix, the candidate receives two violations, etc.

The classes of inner and outer affixes are those defined in section 2. In short, inner affixes attach only to verb stems, while outer affixes may attach to any root, and may also appear as independent words. According to this definition, -oe is an inner affix. It is in fact the “rightmost” inner affix, in that for every inner affix -x, there is a top-ranked constraint *-oe-x.

Note that constraint does not specify a stronger penalty for having two inner affixes ordered after an outer affix. The penalty for such candidates will never determine the outcome of a tableau, since all such candidates ruled out by inviolable bigram constraints before they reach the level of ALIGNINNER-OUTER. These bigram constraints have not been shown in tableau 3 because they involve morphemes that are not included in the input.

The constraint ALIGNINNER-OUTER replaces the lower-ranked constraints on grammatical bigrams in the constraint hierarchy:


Tableau 3 using this ranking shows that form (a) is the winner, while (b) is a runner-up. Since (b) is actually attested, a system in which the runner-up candidate could sometimes win would be preferable. However, stochastic OT is not a possibility because there is only one low-ranked constraint.

<table>
<thead>
<tr>
<th>see (dake)</th>
<th>FUT (oe)</th>
<th>NEG (suli)</th>
<th>PL (mala)</th>
<th>*sul-oe</th>
<th>*o-mala</th>
<th>*mala-suli</th>
<th>ALIGNINNER-OUTER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>☁ (a) dak-o-sur-mala</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) dak-mal-o-suli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>(c) dak-sur-mal-oe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>(d) dak-sul-o-mala</td>
<td>*</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>(e) dak-mar-sul-oe</td>
<td>*</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>(f) dak-o-mar-suli</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>**</td>
</tr>
</tbody>
</table>

Tableau 3

This again raises the question of why order (b) is sometimes grammatical, and whether it is possible to predict when it will occur. Clearly, both (a) and (b) have semantics that include the meanings FUT, NEG, and PL; however, there may be subtle differences between the two. For more
than one consultant, order (b) was considered to be more appropriate if it was interpreted as a question with the meaning “Aren’t you going to x?” This was often given as the translation of the out-of-context verb forms with this ordering. This suggests that order (b) has a semantic quality that makes it more appropriate in certain contexts, including when it is being used as a question. A naturally occurring example of order (b) used in another type of question in an unrecorded conversation is given in (47):

(47) \[ \text{gab=mal-oe or gab=mal-oe-suli} \]  
\[ \text{sleep=PL-FUT or sleep=PL-FUT-NEG} \]  
‘[Are you] going to sleep or not?’  
(unrecorded conversation, Jan. 7, 2009)

An example of the use of order (b) in the recorded conversations is repeated here from example (38):

(48) \[ \text{na=mal-o-suli} \]  
\[ \text{go=PL-FUT-NEG} \]  
‘[The children] won’t go [to school tomorrow].’  
(July 3, 2008, rec. 2, 7:08)

The context of this utterance is that one (so-called) speaker (myself) was assuming that the children would go to school the following day, but the speaker who produced (43) knew that they would not, because it was a holiday. The speaker who produced the verb was therefore directly refuting the content of the previous utterance.

What does this situation have in common with the use of order (b) as a question? In all three cases discussed, there is an emphasis on the negative. When order (b) is used as a question in (47), it uses an exaggerated intonation expressing surprise and disbelief regarding the possibility that the addressee is NOT going to do x. What is at issue is whether or not \( x \) will happen; in other words, whether or not \( \text{NEG} \) can appropriately be used to describe the situation. In the case of example (48), it is again at issue whether or not \( x \) will happen, since there is a temporary disagreement between the speakers over whether \( \text{NEG} \) can appropriately be used to describe the situation. In short, it appears that in both cases \( \text{NEG} \) is in focus. In order to show that this is the case, 18 more-or-less spontaneously produced examples of the two different morpheme orders in contexts are given in the appendix, most of which appear to support the hypothesis.

Based on this data, it appears that a focus position exists at the end of the verb word. This is the case at least when -suli is at the end of the verb word, and possibly in other cases as well. The focus position can be modeled using an alignment constraint on the location of the focused affix, as shown in (49):

(49) \text{ALIGNFoc-R: A focused affix appears at the right edge of the word.}

This constraint is violated when a focused affix appears somewhere other than the right edge of the word. Note that in this paper, the constraint is only used for cases in which the negative morpheme -suli is focused; however, it is possible that it could have a broader application. The constraint is ranked below the bigram constraints, and above \text{ALIGNINNER-OUTER}:

(50) \[ *\text{suli-oe, *oe-mala, *mala-suli} \gg \text{ALIGNFoc-R} \gg \text{ALIGNINNER-OUTER} \]

In tableau 4, the input shows that the semantic feature \text{NEG} has the feature +FOC. If this were not
the case, then the output would be identical to that of tableau 3. In tableau 4, order (b) is a
winner. However, it will only win when \textsc{neg} is focused. When \textsc{neg} is not focused, (a) will still
win. This means that the analysis can account for the grammaticality of both (a) and (b). The
lower frequency of (b) in the elicitations is explained by the fact that it has more specialized
semantics than (a). In absence of any context; for example, when translating an isolated sentence
from Spanish, speakers are more likely to produce the unmarked form (a); but (b) is not an
incorrect response.

\begin{table}
\begin{tabular}{|l|c|c|c|}
\hline
\textbf{see (dake)} & \textbf{*sul-oe} & \textbf{*o-mala} & \textbf{*mala-suli} \textbf{ALIGNFOC-R} \\
\textbf{FUT (oe)} & \textbf{ALIGNINNER-OUTER} \hline
\textbf{NEG (suli) (+FOC)} & & & \\
\textbf{PL (mala)} & & & \\
\hline
(a) dak-o-sur-mala & & & \\
\hline
\textbf{3} (b) dak-mal-o-suli & & & \\
\textbf{(c) dak-sur-mal-o} & & & \\
\textbf{e} (d) dak-sul-o-mala & & & \\
\textbf{(e) dak-mar-sul-oe} & & & \\
\textbf{(f) dak-o-mar-suli} & & & \\
\hline
\end{tabular}
\end{table}

Tableau 4

This ranking accounts for the grammaticality of (a) and (b) and effectively rules out order (c)
because it has two outer affixes ordered before an inner affix. However, there are other forms
that do in fact contain two inner affixes ordered before an outer affix, and we need to make sure
that the analysis does not rule them out as well. A constructed example representative of these
forms is given in (51), and tableau 5 shows it as the winning form (omitting the irrelevant *sul-
oe and *mala-suli).

(51) dak=mar-mog-oe

\texttt{see=}\texttt{=}\texttt{PL-also-FUT}

‘They will also see.’

\begin{table}
\begin{tabular}{|l|c|c|c|}
\hline
\textbf{see (dake)} & \textbf{*o-mala} & \textbf{*o-moga} & \textbf{*moga-mala} \textbf{ALIGNFOC-R} \\
\textbf{FUT (oe)} & \textbf{ALIGNINNER-OUTER} \\
\textbf{also (moga)} & & & \\
\textbf{PL (mala)} & & & \\
\hline
\textbf{3} dak-mar-mog-oe & & & \\
\textbf{e} dak-mo(g)-mal-o & & & \\
\textbf{e} dak-o-mo(g)-mala & & & \\
\hline
\end{tabular}
\end{table}
Variable Affix Ordering in Kuna

Tableau 5 contains some additional top-ranked bigram constraints because of the introduction of -moga, but otherwise the constraint ranking is the same as in tableau 4, with all of the bigram constraints being equally ranked and inviolable. The new constraint ALIGNFOC-R is not active in this tableau because there is no focused affix. Unlike tableau 5, because -suli does not appear in the input form, there is no affix that can provide a buffer between -oe and -mala, the only way that a candidate can avoid violating *-oe-mala is by having the affixes appear in reversed (and non-adjacent) position, even though the presence of -moga means that this will result in a double violation of ALIGNINNER-OUTER.

7 Conclusion

In the final constraint ranking, constraints against ungrammatical bigrams remain in the highest-ranked position, and are never violated. Constraints on adjacency are clearly a powerful force in Kuna affix ordering, and it is impossible to ignore the fact that the high-ranked bigram constraints are never violated. The bigram constraints alone account for all of the single-bigram forms, and are also consistent with two-bigram (three-affix) data. Furthermore, in terms of token frequency, the bigram constraints account for the great majority of the data: of forms that contain two or more of the six affixes in the study, the great majority contain only two of the affixes, or a single bigram. Non-transitivity is not so much accounted for but ignored by this analysis: it is allowed because most of the constraints are local rather than global.

Given that forms with three of the exceptional affixes are relatively rare in conversation, it is not surprising that there should be some variability in their ordering. When forming a word, speakers always obey the top-ranked bigram constraints. However, if they are still left with some choice in ordering after obeying these constraints, they do not always choose the same order. In order to explain the ordering of these three-affix combinations, two analyses were attempted; the first using low-ranked bigram constraints, and the second using zone-boundary and focus constraints.

The alignment constraint appears to be a better choice than the low-ranked bigram constraints. As shown in section 2, there is robust evidence for the existence of the inner/outer zone boundary in other aspects of the grammar, while there is no independent evidence for the existence of the low-ranked bigram constraints. The surprise in this case is not that the constraint ALIGNINNER-OUTER exists, but rather that it can be violated at all.

The lowest-ranked constraint, verb-final focus, allows us to account for the cases where order (b), which violates the alignment constraint, wins out over order (a). This constraint makes a very interesting claim about Kuna verbs; namely that information structure can influence morpheme order. This is a strong claim supported by little data; the corpus must be expanded in order to provide more evidence for it. It will be especially interesting if morphemes other than the negative -suli can be focused in this way.
Appendix: The Effect of Focus on Kuna Affix Ordering

8.1 Introduction

This appendix gives contextualized data on forms containing three of the affixes in question, with the aim of supporting the hypothesis that focus on the negative influences affix order. Because verb forms containing three of the morphemes in question (the negative, the future, and one of the relevant adverbial affixes) are so rare in the corpus, additional forms were obtained through free-form conversational elicitation. Elicitation was carried out by having conversations with speakers about their lives and opinions, so that the context of the verb form would be available for study. In order to avoid priming effects, I avoided using the relevant verb forms, instead using verb-free utterances or starting sentences and hoping that the speakers would finish them. In some cases, the data given comes from unelicited conversations that happened to contain the forms in question. For the data given in this section, please note that lines in italics were spoken by the researcher and should not be taken as data. Because the passages are relatively long, glosses are omitted in the contextualizing sentences.

The conversational data is expected to support the following hypothesis, which is relevant to forms containing three-morpheme combinations of the relevant morphemes (-oe, -suli, -mala/-moga/-bali):

The NEG-final ordering (-x-oe-suli or -ADV-FUT-NEG) is preferred over the alternative grammatical ordering (-oe-suli-x or -FUT-NEG-ADV) in contexts where the negative is in focus.

Determining whether or not a negative is in focus can be difficult, because it is possible that the only indication of the speaker’s decision to focus the negative could be the affix order itself. However, for the purposes of this study, the conversational context will be used as a way of determining whether or not the negative is in focus. The negative is likely to be in focus if the negative statement contradicts or rebuts a positive version of the proposition appearing in the preceding context.

In order to determine whether there is a contrast between negative and positive, examples will be examined for the presence of relevant negatives or positives in the preceding context. The contextualizing sentences need not necessarily appear immediately before the negative form in question, but they must appear within a few conversational turns. The positives sometimes take the form of a question. In some cases, there is both a positive and a negative in the preceding context. These passages also show a large number of uses of the marker -di, which I have glossed as CONTR (contrastive). This marker has been identified in Smith (2009) as a subject marker, probably because of the extensional overlap between contrastive topics and overtly expressed subjects. However, it can appear on objects and other arguments as well as subjects. It is used either to highlight a contrast between real-world entities or situations, or to show a shift

---

25 Elicitation sessions were carried out mostly in Kuna Nega, a Kuna town near Panama City. A few of the sessions were carried out in Concepción, another small town near Panama City, and one was carried out in the city center. All speakers were raised in Kuna Yala, where Kuna is the dominant language. Most speakers are bilingual in Spanish and Kuna and have children or grandchildren whose dominant language is Spanish. They use Kuna as their primary language at home and in the community, but are more likely to use Spanish words than speakers living Kuna Yala. Spanish words are not italicized because it is unclear which are borrowings and which are instances of codeswitching. Spanish spelling is used in order to make the words more recognizable to readers.
in conversational topic. While the contrastive marker is relevant to the discussion, it is used so frequently that correlating its distribution with any particular morpheme order is difficult. Furthermore, it is homophonous with the nominalizer -di, which can sometimes lead to confusion.

8.2 Neg-Final Orderings

The first data set consists of forms with Neg-final ordering. According to the hypothesis, the context of these verbs should highlight a contrast between the negative statement and a corresponding positive proposition.

In the first example, the speaker used one of the forms in question while telling me about a phone conversation. I then turned on the recorder and asked her if she could repeat what she said. The form has the order mala-oe-suli (PL-FUT-NEG):

(52) nanaye, sabadogi an namaloenye, beseye,
    “Mom, we were going to go to you on Saturday,” (I said to her),

    emi-di an na-mal-o=sur-dakenye.
    now-CONTR 1ST go-PL-FUT=NEG-QUOT
    “but now we will not go.”

    gwenad nailikgumaye, uellemadakenye.
    “A relative is lying having been hurt; [she’s] lying sick.” (Aug. 12, 2010, rec. 1, 0:15)

Note that the sentence containing the Neg-final form contains the affix -di which combines with emi (‘now’ or ‘today’) to form the meaning ‘but instead.’ In example (52), emidi signals a contrast between the positive statement ‘we were going to go’ and the negative statement ‘we are not going to go.’ This conforms to the hypothesis that the Neg-final order is found in contexts where negative and positive are contrasted.

In example (53), a contrast is drawn between the positive proposition that gay marriage will be accepted in Panama, and the negative proposition that it will not be accepted in Kuna Yala. The positive proposition is reiterated in the sentence immediately preceding the Neg-final verb form. In this conversation, both A and B are native speakers of Kuna; C is the researcher. The order of morphemes is mala-oe-suli (PL-FUT-NEG).

(53) A: allá en-- bia noticias oyolesad aa, geb nikgunananai, matrimonio, aceptaron.
    Over there in-- where it was shown in the news, [they] just started marrying, marriages, [they] accepted it.

    B: hombre y hombre?
    A man and a man?

    C: eye.
    Yes.

    [...]
C: igi be dake?  
How do you see it?

A: eye, igi be binsa, neg amarga gwae gwae o igi?  
Yes, what do you think, is everything changing on us, or what?

igi an sogmoga, eye gwaydapo an sogodo.  
What can I say, yes, it’s changing, I say.

San Blasgindi suli; San Blasgindi  
But not in San Blas; in San Blas

accept-sa=mal-o-suli  
accept-do=PL-FUT-NEG  
[they] won’t accept it.  

As in example (52), in example (53), the contrastive marker -di appears on San Blas (Kuna Yala), marking the contrast between Panama (City) and San Blas, and therefore between the positive and negative versions of the statement. Therefore, it seems that this example also conforms to the hypothesis. However, it is important to note that the passage also contains a negative statement immediately preceding the negative form in question.

In example (54), which also has NEG-final ordering, a contrast is established between the positive statement (‘I don’t mind making molas’), and the negative statement (‘young girls won’t wear molas anymore’). In this example, there is no contrastive marker -di; however, the negative polarity item bar ‘anymore’ is used. This serves to emphasize the contrast between the two propositions, because it contrasts between the past and present generations. The form has the order mala-oe-suli (PL-FUT-NEG).

(54) A: bar bunormala yoobisuli we mor. bar abesuli.  
The young girls don’t want to wear this mola anymore. [They] don’t want [it] anymore.

soge arbaed bipisuliye.  
[They] say it’s a lot of work [to make molas].

[...]

suli! andi arbaed du-- gwen dummad an itoged an, an sae!  
No! As for me, I don’t feel like it’s a lot of work, I do it!

B: tegite be ... ee ... igi nuga, ee *yoomarsuli-- yoo--

26 When -bar (-bali) appears as a suffix, it is not a negative polarity item; appearing in both positive and negative constructions.
27 The form yoomarsuli spoken by the researcher is problematic for two reasons: First, the researcher may have primed the speaker by using this verb form, which is the same as the speaker’s subsequent verb form except that the future -oe is omitted. Second, this form is ungrammatical; the order of suffixes should be reversed.
Then you ... uh ... what’s it called, [they] not wear it ...

A: bar yoosuli.
   [They] don’t wear it anymore.

B: *bar ee, mordina?*
   *Anymore, uh, and the mola?*

A: bar yogguoe gwable. mor yogguoe, win yogguoe, daksaa.
   [It] will all be lost. The mola will be lost, the beads will be lost, see.

   bar ar yoo=mal-o-suli  
   again well wear=PL-FUT-NEG  

This example appears to support the hypothesis. However, as in example (2), there is a statement with a negative meaning immediately preceding the negative form in question.

In example (55), repeated from (38) with additional context, the positive version of the proposition enters the context in the form of a yes/no question. The speaker contradicts the question’s suggestion that the children will go to school tomorrow, saying that they will not go to school because there is a holiday. There is no contrastive marker in this example; this would be impossible since the subjects, objects, and settings of both versions of the sentence are identical.

(55)   A: banedsedina escuelase namarmoga ...
   *And tomorrow, are [they] going to school too?*

   [...]

   B: bane clasesuli, bane ibua?
   Tomorrow there’s no school; what’s tomorrow?

   [...]

   A: *igi be sogsa, bane clase sate?*
   *What did you say, there’s no class tomorrow?*

   B: clase sate bane.
   There’s no class tomorrow.

   A: degite bur- igi ... mimmigan
   *So – how ... the children*

   B: na=mal-o-suli  
   go=PL-FUT-NEG  
   [They] won’t go.   (July 3, 2008, rec. 2, 7:08)

In this case, although there is no contrast marker, the context provides ample evidence for a
contrast pair between positive and negative. As in example (54), there is a sentence with a negative meaning appearing before the negative form in question. However, the context still appears to be contrastive.

In example (56), repeated from (47), the utterance has no conversational context since it was the first utterance of a short conversation:

(56) gab=mal-o-suli
sleep=PL-FUT or sleep=PL-FUT-NEG

‘[Are you] going to sleep or not?’

In this example, a direct contrast is drawn between the negative and positive versions of the proposition. Furthermore, the direct juxtaposition of the two orderings shows how the NEG-final ordering preserves the ordering of the positive proposition, simply adding the negative at the end. This suggests that priming effects may contribute to the tendency to use the NEG-final order in contexts containing the positive ordering. There is no other negative sentence in the context, so this example strongly supports the hypothesis.

In example (57) the negative statement that land in Kungiler Nega (a group of homes being built adjacent to the municipality of Kuna Nega) will not be cleared is being contrasted with the positive statement that land in Kuna Nega will be cleared. The form has the order mala-oe-suli (PL-FUT-NEG).

(57) A: suli, we Kuna Nega suli, we Kungilermeg
No, that’s not Kuna Nega, that’s Kungiler Nega.

B: urgoneg sobnaidgi?
Where they are building the wooden houses?

A: eye, urgomarneg sobarbukwamala we Kungilerneg
Yes, where they’re building the wooden houses, that’s Kungiler Nega.

wedi aaga napa neg uilesmala. (--aa!) 
That’s why they gave [us, Kungiler Nega people] that land.

Kuna Nega degsoggudi egi dogarnaisuli, Kuna Nega gwapa, 
It’s not Kuna Nega, so [they] aren’t going in there, all the Kuna Nega people,

... Kuna Negadi campo eminaidi. 
... the people that clear the land around Kuna Nega.

nuedi anmarga soge  
“Good,” they tell us,

be-mar-di campo emi-le=mal-o-sur-dake 
2ND-PL-CONTR land clear-PASS=PL-FUT-NEG-QUOT 
“As for you [people in Kungiler Nega], that land won’t be cleared.”

(Aug. 11 2010, rec. 5, 13:43)
Again, the NEG-final order occurs in a context where there is a marked contrast between positive and negative. Although the positive version of the negative appears inside a relative clause, it is still overtly present. The contrast marker -di appears on the pronoun you, which refers to the people of Kungiler Nega, who have no obligation to clear land, as opposed to the people of Kuna Nega, who do. This has the effect of contrasting the positive and negative versions of the statement.

In conclusion, all six of the examples involving NEG-final ordering appear in contexts containing a positive statement that is contradicted, supporting the hypothesis that this ordering appears in contrasts in which NEG is focused. However, it is important to note that four of the examples (53, 54, 55, and 57) also contain statements with a negative meaning preceding the negative form in question.

8.3 Adverbial-Final Orderings

The following examples have the adverbial-final affix orderings (ending with -mala, pl., -moga, ‘also,’ or -bali, ‘again, back’). If the inverse of the hypothesis is true, then the conversational context should lack a marked contrast between the negative sentence and a positive counterpart. There should be no overt statement of the positive version of the negative. If this is the case, the example in question will be seen as supporting the hypothesis. Of course, given this situation, it is expected that the context should be more likely to contain statements with a negative meaning, such that the negative form in question agrees with them.

In example (58), taken from the same conversation (and same speaker) as example (57), there is a contrast between the negative sentence (‘In Kungiler Nega, you won’t clean up the land’) and a positive counterpart in the preceding context (‘In Kuna Nega, they always clean up the land’). The contrastive marker -di is also used to distinguish the two locations, which correspond to the negative and positive propositions. The example contains three verb forms with the ordering -oe-suli-mala (FUT-NEG-PL).

(58) A: Kuna Negagindi campo emileardaed.
In Kuna Nega, they always clean up the land.

   ee dake, gaganmar weararananai,
See, [they] are gathering up the leaves,

   emilergebe aa emilesurdibe multarguleged igi mani, geb ...
[you] have to clean up the land; if [you] don’t, [you] get fined, how much money, and ...

B: webadina?
And over there?

A: webardin naanmarga soglesa, suliye,
As for over there, they told us, “No,”

   be-ga im-mar akar gu-o=sur-mar-ye
2ND-DAT what-PL different become-FUT=NEG-PL-QUOT
“things won’t go bad on you,
One explanation for the unexpected ordering may be that because the statements in (58) are reiterating the negative statement made earlier in the conversation, shown in (57). Even though a positive statement is also present, the speaker is agreeing with a previous negative statement. Of course, the last two of the three negative verb forms directly follow other negative statements with similar meanings. These three forms must be considered exceptions to the hypothesis, but their exceptional nature is explained by the other negatives in the context.

Similarly, in example (59), the positive version of the negative statement is suggested in a question (‘will they forget?’), but the speaker does not opt for the NEG-final ordering (‘they will not forget’). Instead, he uses the *oe-suli-mala* (FUT-NEG-PL) ordering. The contrastive marker *-di* is also present in the context, but it is highlighting the speaker’s identity in a meta-statement that has little to do with the negative proposition. This is another counterexample to the inverse of the hypothesis.

(59) A: *itogena. anmardi geg deggue. nabir, igi anmardi, iedgu anmar gaya. suli.*
Listen. As for us, we can’t do that, really, how could we, how could we forget our language? No.

anmar mimmigan odurdaksuli. degi aa? be itogena.
We don’t teach the children. That’s it, right? Listen.

B: *degite, iegoe?*
*Then, will [they] forget?*

A: *iegosuli* an dakdo. aadi ee escuelagi, emigi escuelagine maestramar odurdaksokarmoga.
[They] won’t forget, I see. As for that, in school, now the teachers are about to teach it in school too.

B: *mmm.*

A: *aye. an-di an dak-oe bar ieg-o=ur-mala an dak=sun-do, dak-sa*\(^{28}\)
*yes 1ST-CONTR 1ST see=FUT again forget=FUT=NEG-PL 1ST see=truly-well see-PFV*
Yes. I see that [they] won’t forget it, I see, [you] see.

(Aug. 11, 2010, rec. 3, 4:14)

---

\(^{28}\) This example also includes *bar*, but in this context, the word has the sense ‘again,’ as in ‘they won’t forget it again once they learn it.’ This sense of *bar* does not have a contrastive meaning, so it is not problematic for the inverse of the hypothesis.
The reason for this unexpected ordering could be that another negative verb (iagosuli, ‘they won’t forget’) directly precedes the negative form in question. As in example (58), the speaker may see the negative form as a continuation of the previous negative statement, and therefore not mark it as contrastive. The previous negative could also have the effect of priming its ordering; the second verb being identical to the first except for the addition of the final -mala (‘PL’).

Example (60) is consistent with the inverse of the hypothesis in that it does not contain any overt statement of the positive version of the negative sentence, ‘they won’t speak Kuna anymore.’ It also does not contain the contrastive marker -di, although the negative polarity item bar, ‘anymore,’ arguably indicates contrast.

(60) bar anmar dulegaya yoggudoe binna binna bar dulegayaba bar sunmaksurmala,
Our Kuna language will be lost, bit by bit, they don’t speak in Kuna anymore

dake emisgwa bato gudidiimala aku itomalad.
Look, now already there are [kids] running around that don’t understand.

tegido ep– yoo–
So, back then– in the fut–

yooedse burguoedsunna bur buledo,
in the future, it will really die, even more,

bar sunmak-o=sur-mala dule-gaya
again talk-FUT=NEG-PL person-mouth

This does not technically contradict the inverse of the hypothesis, so it is considered as supporting the hypothesis.

Example (61) is an unelicited example that was captured during a break in elicitation sessions. In this example, a working mother scolds her children, saying ‘I won’t buy anything for you.’ This example fits the hypothesis in that there is no overt positive statement to be contrasted with the negative statement. Instead, the negative is part of an if/then construction, where the if-clause also has a negative meaning. The adverbial prefix is -bali, which can be interpreted here as ‘again’ (with respect to buying things) or ‘either’ (with respect to not going to work).

(61) Yanita, Yamar, aide! uaya bemar sater be it – itosa,
Yanita, Yamar, get down! If you don’t listen, [you] hear,

bega an immar amb an arbanaosuli,
I won’t go to work for you anymore,

an be-ga ib-mar bak-o=sur-ba
1ST 2ND-DAT what-PL buy-FUT=NEG-again
I won’t buy anything for you either.   (Aug. 20, 2010, rec. 1, 16:09)
This example is consistent with the inverse of the hypothesis, since there is no positive statement and no contrast markers.

Example (62) is another case of the adverbial-final ordering (-oe-suli-mala, FUT-NEG-PL), in which there is no overtly stated positive version of the negative statement. Again, the negative statement is part of a sequence with an if/then meaning, ‘you have to do a lot of paperwork or they won’t let you in.’ There is no contrastive marker -di. The example conforms to the inverse of the hypothesis.

(62)  
A: doctorse arbi?  
Did he go to the doctor?  

B: yoo an nae. wegindi bulë arbaed geb bin nanalege.  
I haven’t gone yet [with him]. Here [in the city] it’s a lot of work to go.  

hospitalse nadbardibe an oibogoe, daksa.  
If [I] go to the hospital I will stay up all night, see.  

papel immak-dag-oe abgayala be o-dog-o=sur-mala  
paper make-come-FUT at.first 2ND CAUS-enter-FUT=NEG-PL  
[You] will come to do a lot of paperwork first [or] [they] won’t let you in.  

(Aug. 19, 2010, rec. 6, 10:06)

This example is taken as supporting the hypothesis.

In example (63), there are two verbs with adverbial-final orderings, (o-sur-moga, o-sur-mala). In the context of the first verb, there is no direct statement of a positive version of the negative sentence (‘My mom won’t stay here either.’). However, the sentence does follow the statement (‘what about here?’); which could be seen as invoking the positive version of the negative. The first verb is also in a quoted sentence containing the contrastive marker -di, which refers to the first-person speaker of the quote (‘my mother’). However, the morpheme is not marking a positive-negative contrast, since the marker is contrasting two different people who are both doing the same thing (not staying here). This verb must be taken as a counterexample.

In the second verb in this example, there are positive sentences (‘it’s up to you’) (‘if you want to stay’) in the context, but they are in agreement with the negative sentence (‘I won’t obligate you’). There is also a contrastive -di in the sentence, but it serves only to clarify the identity of the speaker. This can be considered consistent with the hypothesis.

(63)  
A: annan anba namogoe anga,  
My mom will go with me too [she said] to me  
akarsuli beba namogoeeye.  
“No problem, I’ll go with you too.”  

geb despues San Blasse nader.  
“Then afterwards [I] might go to San Blas.”
Variable Affix Ordering in Kuna

mientras bemar an visitar sagwichi nabir samaloe.
“In the meantime I’ll stand visiting you, we really will do [that].”

B: bar wegi ... be ... wegindina ...
Here anymore ... you ... what about here ...

A: an-nan-di wegi yoggu-o=sur-moga-d pudar wegi yoggu=moga-d
1ST-mother-CONTR here stay-FUT=NEG-also-RHET sick.of here stay=also-RHET
My mom won’t stay here either. She doesn’t feel like staying here either.

[...]

anmarga soge, bedi yoggubier, wegi be yoggumalo.
[She] told us, as for you, if you want to stay, you’ll stay here.

be-gi sii=mar-ye pero an-di be obligar-sa-o=sur-mala
2ND-in sit=PL-QUOT but 1st-CONTR 2ND obligate-do-FUT=NEG-PL
It’s up to you. But I won’t obligate you.

(Aug. 18, 2010, rec. 5, 31:29)

This example therefore has one verb out of two that is consistent with the hypothesis. In example (64), another adverbial-final ordering (o-sur-moga), there is also no direct statement of a positive version of the negative statement (‘You won’t see them anymore’). Instead, the negative statement is a continuation of similar statements in the preceding context. The word bar appears in the negative sentence, but here it has the meaning ‘(not) again,’ rather than a contrastive meaning.

(64) ... garmakdiidamoga, daksa. degisoggu be igi nugde, igar nikasuli itolesundo. ... always running around, see. So you, what’s it called, without direction, it is heard truly.

machergan deobmoga machergan arbaed baid amisdibe naed, dikasur; Men are like that too, if men find another job, they go, far away;

bar be dak-o=sur-moga
again 2ND see-FUT=NEG-also
you won’t see them anymore either;

bane be gudiid nabiedbarsundo.
one day you’re here, then [you] want to leave again. (Aug. 17, 2010, rec. 3, 0:10)

This example therefore conforms to the hypothesis. In example (65), there are two adverbial-final sentences, both with the same affix ordering (-oe-suli-bali, or FUT-NEG-AGAIN). These sentences are a special case, because affix -bali has a

29 Here, the plural -mala refers to the object, rather than the subject. This is one of the reasons why it is considered an adverbial rather than an agreement affix.
few different senses depending on the context. In positive verb forms, it can mean either ‘back’ (as in a reversative) or ‘again.’ When combined with the negative, -bali can mean either ‘not back,’ ‘not again,’ or ‘not either,’ where either sometimes acts as an intensifier, as it does in colloquial English. The ‘not back’ meaning appears to correlate with the order -ADV-FUT-NEG (-bali-oe-suli), while the ‘not either’ meaning appears to correlate with the order -FUT-NEG-ADV (-oe-suli-bali). This makes sense in that the ‘either’ or intensifier meaning places the morpheme in question in word-final position, which is arguably a focus position. In any case, the order of morphemes in the two forms in (65) is probably determined by the semantics.

Similarly to example (64), the first of these sentences (‘You won’t go back to Panama’) is a continuation of a previous discourse which also has a negative meaning. Earlier in the discourse (not shown) there is a positive sentence that is contradicted (‘I want to study’), therefore this sentence is technically a counterexample.

The second sentence containing the same string of affixes actually seems like it would be a good candidate for the NEG-final ordering, because the speaker is relating how he directly contradicted his sister’s suggestion, a positive statement. In this way, it also appears to contradict the hypothesis.

(65) andi an ega soge, suliye, estudiadose bergus begad.
    I said to [my children], “No, about studying, it’s finished, yours.”

    estudio bar saosunna mani sateye estudiarsaed.30
    “Studying [you] will [not] do again, truly. There’s no money for studying.”

    Panama-se na-o=sur-bali
    Panama-to go-FUT=NEG-again
    “[We] won’t go to Panama [City] either.”

    […]

    geb an bun soggali, suliye, an be bendakoye.
    Then my sister said, “No, I’ll help you.”

    […]

    bat begad trabajo gwakwa sii Panamagi.
    “A job sits ready for you in Panama.”

    an soge suli-ye, an na-o=sur-bali-ye
    1ST say no-QUOT 1ST go-FUT=NEG-again-QUOT
    I said, “No, I won’t go either.”

(Aug. 15, 2010, rec. 4, 4:35)

Although these two examples seem to contradict the hypothesis, their ordering may be explained by the special circumstances of the affix -bali. If -bali appeared in its other position, it might be interpreted as ‘back.’ However, in this case, it has the meaning ‘not either.’

Table 4 summarizes the data. If the hypothesis were followed perfectly, the first column

---

30 Here, the negative polarity item bar appears without a negative morpheme, but its meaning is still negative.
would show 100% for positives and 0% for negatives; while the second column would show 0% for positives and 100% for negatives. Note that among the adverbial-final examples each verb form is counted separately, so that the number of examples is 12 rather than 8.

<table>
<thead>
<tr>
<th>Context</th>
<th>NEG-final ordering</th>
<th>Adverbial-final ordering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contradicting positive</td>
<td>100% (6/6)</td>
<td>50% (6/12)</td>
</tr>
<tr>
<td>Agreeing negative</td>
<td>50% (3/6) (4/6)</td>
<td>92% (11/12)</td>
</tr>
<tr>
<td>Contrast marker</td>
<td>(3/6)</td>
<td>(6/12)</td>
</tr>
</tbody>
</table>

Table 4: The effect of focus on affix ordering

The data on NEG-final verb forms supports the hypothesis that the NEG-final order arises when there is a focus on the contrast between negative and positive. The data on adverbial-final verb forms shows only weak support for the inverse of the hypothesis. However, among the adverbial-final sentences, all exceptional cases except for those in example (63) can be explained either by the fact that the negative form is preceded by a similar negative sentence (examples 58 and 59) or by the argument that the adverbial-final order is used for semantic reasons (example 65). The two forms in example (63) remain a problem for the hypothesis.

### 8.4 Word-Final Focus and the Habitual

Finally, in order to explore further the idea the word-final position is some form of focus position, data is presented using a different affix, the habitual. These examples do not contain the future or the other adverbial affixes; they contain only the negative and the habitual. The ordering of the negative -suli and the habitual -dae is more straightforward than the three-affix combinations because the two morphemes can appear adjacently in either order. Unfortunately, these examples are rare in the text corpus because there is a competing construction with a similar meaning. From the small amount of data available, it appears that the NEG-final order is used to emphasize the negative as opposed to a positive version of the sentence, just as in the three-morpheme orderings in the previous section.

Beginning with the habitual-final orderings, the context shows no particular emphasis on the contrast between negative and positive. In example (66), there is a positive statement (‘I think women’s beads are beautiful’) earlier in the discourse. However, the verb form in question follows a negative statement (‘They’re messy, loose.’). There are no contrast morphemes in the sentence. Therefore, this example may possibly support the hypothesis.

(66) yer an dakdae omegan win aa?
    I think women’s beads are beautiful, right?

    win nasis nuedsunnaaidibe.
    If the beads are strung really well.

31 The competing construction is periphrastic, with the negative morpheme attached to the verb ‘know’:

    an wis=sur e-bo sunmake
    1ST know=NEG 3RD-COM speak
    ‘I never speak to him.’
wege as-- win nued we sunba sunba akala sinni sinnigwa.
Here-- beads well, they’re messy, loose.

an abe= **sur-dae**, isdar an dak=dae
1st want= **NEG-HAB** badly 1ST see=HAB
I don’t tend to like that, I think they look bad. (Aug. 17, 2010, rec. 2, 27:35)

In example (67), there is no positive version of the negative statement (‘half of them don’t usually eat’); instead the negative statement functions as an explanation for the preceding sentences which state that a cafeteria is going to be built. This conforms very well to the hypothesis.

(67) geb irgwen sigsokarbalid ee, comedor infantil burwiga.
They’re about to sit another thing here, uh, a children’s cafeteria, for the kids.

aa egilenaai burwiganga mas gunmalad.
That’s what’s being asked for, so that the kids eat.

abar mas **gun-sur-da-mala**
half food eat= **NEG-HAB-PL**
Half [of them] don’t usually eat. (Aug. 8, 2010, rec. 1, 1:14)

In example (68), the **NEG-final order** is shown. Unfortunately, this is the only example of the **NEG-final order** with the habitual. In this situation, the speaker actually uses two different constructions with very similar meanings, (‘I didn’t use to drink’). The first time, he makes the statement using the negative and the imperfective, in a normal tone of voice. The second time, he makes the statement using the negative and the habitual with **NEG-final ordering**. There is no overt contrasting positive statement (‘I drink now’) or contrastive markers; however, it is safe to say that the negative aspect of the sentence was being emphasized, since it was spoken very strongly with dramatic hand motions.

(68) gebesaila an cerveza gob=*suli-na* 32
at.first 1ST beer drink= **NEG-IMPV**
At first I didn’t drink beer,

gob= **da-suli**
drink= **HAB-NEG**
I didn’t use to drink at all. (unrecorded conversation, Aug. 20, 2010)

Of course, since stress always falls on the first syllable of *-suli* when it is the last morpheme in a word, emphasizing the word results in emphasizing *-suli*. When *-suli* is followed by an adverbial, stress may fall either on *-suli* or on the adverbial, and in any case, *-suli* appears in its short form.

---

32 It is ungrammatical to use a **NEG-final ordering** with the imperfective; that is *-na-suli.*
Variable Affix Ordering in Kuna

The small amount data on the ordering of NEG and HAB suggests that the NEG-final order might have the interpretation of “never” while the HAB-final order might have the interpretation of “usually not.” This seems consistent with the trend of the NEG-final endings giving a more emphatically negative meaning. It also seems consistent with the predictions of the semantic scope hypothesis as proposed by researchers such as Rice (2000). More data is necessary in order to determine whether this is true.

There are various reasons why it makes sense that the NEG-final order should be used to mark focus. First, when the negative morpheme occurs at the end of a verb word, it is always stressed. Second, the word-final position makes the negative morpheme more perceptually salient. Third, the NEG-final order may be primed by a preceding positive, which has the same ordering except that it lacks the negative morpheme at the end. It is unclear which of these explanations, if any, may be affecting affix order. However, the negative-focus hypothesis is satisfying because it suggests that there is a reason why speakers have a choice of orders.

References


Ryan, Kevin. 2008. *Morphotactics: Affix Order Variation and Learnability in Tagalog and Other Languages*. Ms, UCLA.


Lindsey Newbold
University of California, Berkeley
Department of Linguistics
1203 Dwinelle Hall
Berkeley, CA 94720-2650
lindseynewbold@berkeley.edu
Passive Constructions in Kwakwala

Daisy Rosenblum
University of California, Santa Barbara
University of California, Berkeley

1 Introduction

Franz Boas, in his 1947 grammar of Kwakwala, identifies a set of passive suffixes used to make syntactic subjects from various non-subject arguments. This paper furthers our understanding of Kwakwala passivizing morphosyntax with an analysis of the syntactic and semantic factors determining the distribution of passive morphemes. After some background on the language and its grammar, I present an overview of the multiple passivizing morphemes available to Kwakwala speakers and describe the function of each passive suffix. I compare the syntactic alignment of pronominal arguments with that of lexical arguments, and explore the various discourse motivations for passive constructions. Discussion and conclusions follow.1

2 Background

2.1 Speakers and Location

![Map of the Kwakwala, Chimakuan, and Salishan Families](adapted from Suttles 1990:ix)

Figure 1: The Wakashan, Chimakuan, and Salishan Families (adapted from Suttles 1990:ix)

1 This research was made possible by the support of the Phillips Fund for Native American Research, the Jacobs Research Fund, the Jacob K. Javits Fellowship Program, and Robert Oswalt Fund for Research on Endangered Languages. I gratefully acknowledge the wisdom, humor and patience of the elders who have shared their knowledge of Kwakwala with me: Lillian Johnny, Beverly Lagis, Gertrude Robertson, Daisy Sewid Smith, and Spruce Wamiss. I am also grateful to Bernard Comrie, Marianne Mithun, Patricia A. Shaw, and Sandra A. Thompson for their generous and insightful feedback. All errors are my own.
Kwak'ala (Wakashan, KWK), formerly identified as Kwakiutl, is the language of the Kwak'waka'wakw nation located on Northern Vancouver Island and the neighboring mainland of British Columbia. It is severely endangered, with approximately 190 fluent speakers remaining.

The data discussed here are drawn primarily from Boas and Hunt’s documentation of the language in grammar, dictionary, and texts from 1893 to 1948, as well as some elicited data published by Levine in 1980, and a few examples from recent fieldwork (2008-present). This analysis should be understood as referring primarily to the language as spoken in the years between 1895-1980. Before we proceed to a description of passive constructions in Kwak'ala, the next section provides an overview of grammar relevant to voice and valence, emphasizing patterns of alignment and reference tracking.

2.2 Grammatical Overview

This section provides an introduction to the grammatical structure of Kwak'ala clauses and argument structure. First, a word about lexical classes: Kwak'ala belongs to the Pacific Northwest Sprachbund for which the appropriateness of lexical classes such as ‘noun’ and ‘verb’ has been questioned (Jacobsen 1979; Kinkade 1983). In this paper, I assume that Kwak'ala nouns and verbs exist as syntactic categories, defined according to derivational and inflectional marking in the context of the clause. The question of their status in the lexicon is more complex and will be set aside. A few derivational suffixes are specific to predicates, while others are restricted to arguments, but as Boas says: “[a]ny ‘verb’ preceded by an article is a noun … and any noun with predicative endings is a verb …” (Boas 1947:280). To avoid confusion, I will usually refer to predicates and arguments rather than verbs and nouns.

Kwak'ala is polysynthetic. The language employs three core argument cases and one oblique case. Alignment of both lexical and pronominal arguments is thoroughly nominative-accusative. For this reason, I use the terms ‘subject’ and ‘object’ in a syntactically-constrained sense, to describe the grouping of single arguments of intransitive predicates (‘S’ in the sense used by

---

2 ‘Kwakiutl’, an anglicized orthographic representation of the name Kwaqwaq, applies only to the band at Fort Rupert (Caycillos) with which Franz Boas and George Hunt did most of their documentation. There are 16 bands in the Kwak'waka'wakw nation and 5 dialects, each with their own name. Some bands prefer the alternate language name ba'mala. I use Kwak'ala here to refer to all dialects.


4 Future work examines the use of passive constructions in new corpora gathered from 2008-present, comparing and contrasting this with earlier documentation. Examples from 2008 were gathered during the InField 2008 at UCSB, in the course “Kwak’ala Field Training” with Patricia A. Shaw, Beverly Lagis, and Daisy Sewid Smith, made possible by a Social Sciences and Humanities Research Council of Canada (SSHRC) Aboriginal Strategic Research Grant to Patricia A. Shaw.

5 Referring to the lexicon, Boas comments that “[s]trictly speaking there are only three classes of words: predicative terms, syntactic particles which define the function of predicative terms, and exclamatory forms. Not withstanding the occurrence of nominalizing suffixes there is no clearcut distinction between noun and verb. Any “verb” preceded by an article is a noun … and any ‘noun’ with predicative endings is a verb. Stems are neither nominal nor verbal. A division may be made between stems of static and active meaning” (Boas 1947:280). In contrast to Boas’ use of the terms active/stative to describe a semantic contrast between stems, the term active is used in this paper for constructions which lack passive morphosyntax.

6 The term ‘polysynthetic’ refers here to the encoding of core arguments on the verb, allowing a single phonetic word to serve as a complete clause, as well as the language’s rich morphological resources, which combine in morphologically complex words.
Passive Constructions in Kʷaḵwala

Comrie 1978 and Dixon 1979) with the ‘A’ (actor or agent) argument of a transitive or ditransitive predicate, as opposed to the ‘P’ (most patient-like argument) of a transitive predicate. The three core argument types are identified here as ‘SUBJECT’ (S), ‘PRIMARY OBJECT’ (O₁), and ‘SECONDARY OBJECT’ (O₂). These terms correspond with Boas’ terms ‘subject’, ‘object’, and ‘instrumental’ (Boas 1947) and with Levine’s terms ‘subject’, ‘object’, and ‘oblique’ (Levine 1980).

My use of the terms ‘primary’ and ‘secondary’ for Kʷaḵwala objects departs from previous traditions in order to avoid the use of the term ‘oblique’ for what I analyze as a third core argument, and to emphasize the core status of secondary objects. Paradigms of pronominal and adnominal enclitics exist for each of the three core arguments — subjects, primary objects, and secondary objects — while non-core arguments occur in prepositional phrases at the end of a clause. Boas and Levine refer to these prepositionally-marked arguments as ‘indirect’ (Boas 1947:206), but I reserve the term ‘OBLIQUE’ (OBL) for non-core arguments. Boas distinguishes between ‘prenominal’ and ‘postnominal’ adnominal case marking of lexical arguments; this paper adopts both terms.

Pragmatically-neutral phrases are predicate-initial, with ‘VSO’ word order. Predicates are distinguishable by their position at the front of the clause, the encliticized flagging of core pronominal arguments, and/or the adnominal marking of lexical subjects, and in some cases the use of derivational suffixes specific to predicate forms. Lexically-specified arguments are case-marked with preceding prenominal enclitics and deictic demonstratives (which, like case-marking, are phonologically grouped with the preceding constituent). A clause with three lexically-specified core arguments is presented below.

(1)

\[
\begin{array}{ll}
\text{hə́n} & \text{ʔid} = \text{id} = \text{da} \\
\text{bə́gʷánom} & \text{χa} = \text{χ} = \text{sa} \\
\text{hə́n} & \text{lə} = \text{m} = \text{i}
\end{array}
\]

‘The man shot the black bear with a gun.’ (Shaw: 2008_07_21_003DS)

Arrows direct our attention to the marking of lexical arguments on the preceding constituent with enclitics; in this case, the subject bə́gʷánom ‘man’ is marked with the prenominal enclitic =i, the lexical primary object χayi ‘black bear’ with prenominal enclitic =χ(a), and the lexical secondary object hə́nɬəm ‘gun’ with the prenominal enclitic =s(a) (and the third-person postnominal sentence-final =i).³

The general shape of prenominal primary-object-marking is a set of variations on =χ(a), and for prenominal secondary object markers on =s(a). Meanwhile, pronominal object-marking tends to include the uvular stop -q for primary objects and again the -s for secondary objects. Based on these shapes, I will sometimes refer to ‘χ-marking’ and ‘s-marking’ to discuss differences between primary and secondary objects. As we will see in section 3.1, these two object cases are also consistently distinguished through the use of different passive morphemes:

---

³ Following the tradition of labeling the primary agent or actor of a transitive verb with ‘A’ and the object argument of the transitive as ‘P’ (Comrie 1978).

⁸ The forms =χ(a) and =s(a) vary to reflect features of deixis and visibility; complete sets of related forms are provided in the appendix.
-suʔ for primary object and -ayu, -ano, and -əm for secondary objects.

A simple clause with specified lexical arguments has VS(O₁)(O₂)(X) order, with V representing a singly- or multiply-expressed predicate, S representing the subject, O₁ representing the primary object, O₂ representing the secondary object, and X representing an oblique argument in a prepositional phrase. The sequence of pronominal enclitics attached to the predicate corresponds directly with the VSO₁O₂ sequence of lexical arguments in a clause for which all three arguments are explicit. Thus one can form a complete transitive or ditransitive clause with a single prosodic word as in (2) and (3) below.

(2) ˚nikənλaq
 ˚nik=ən(ƛ)=aq
say=1s.SBJ=3.OBJ1
I said to him… (B1947:281, CX12.9)

(3) ˚χʷəsʔидəqs
 ˚χʷəs-ʔid(ə)=Ø=q=s
strike-MOM=3.SBJ=3.OBJ1=3.OBJ2
He struck him with it. (B1947:281)

In example (2), the predicate ˚nik- ‘to say’ encodes both the first-person singular subject =ən(ƛ) and the third-person primary object (O₁) =aq. In example (3), the third-person subject (S) is marked with a zero-morpheme =Ø, while both third-person primary (O₁) and secondary (O₂) objects are encoded on the verb with =q and =s, respectively. Returning to example (1), the prenominal SUBJECT marker =i (and definite marker =da) attach to the predicate stem hən-t-ʔid-
‘shoot-PST-MOM’ preceding the lexical subject bagʰənəm(a) ‘man’. The prenominal PRIMARY OBJECT enclitic =χa precedes the lexical primary object ˚χəʔi ‘black bear’, and the prenominal SECONDARY OBJECT enclitic =sa precedes the lexical secondary object hənəm ‘gun’. The postnominal distal demonstrative =i follows hənəm. By comparing examples (2) and (3) with example (1), we can see that third person pronominal enclitics are in complementary distribution with lexical arguments. The domain of attachment for pronominal enclitics is the predicate, but in cases with multiply-expressed predicates, pronominal clitics can be distributed; the subject pronominal enclitic may attach to the first (auxiliary) predicate, and the object pronominals or prenominals may attach to the second predicate.⁹ See also examples (6) and (13). With lexical arguments, the prenominal demonstrative forms occur attached to the predicate or preceding element and specify deictic information about the following lexical arguments.

It is useful, while considering a dual-object system, to mention the connection between argument structure and typologies of ditransitive alignment. The valence of a Kʷəkʷala predicate stem is lexically determined, and can be increased or decreased with derivational affixes. Some Kʷəkʷala stems are inherently transitive and can take objects, while others are inherently intransitive and do not take objects (Boas 1947:280). This is true of English verbs as well: say can take two non-subject arguments, the thing-said and the recipient (the person to whom something is said); for English, the thing-said is marked as a DIRECT OBJECT and the recipient of

⁹ Because the third-person subject pronominal marker is -Ø, and because the initial discourse markers are in the process of grammaticalizing (only sometimes taking person-marking inflection), it is not always clear how to interpret where the third-person subject enclitic attaches.
Passive Constructions in Kwak'wala

a statement as an INDIRECT OBJECT in a prepositional phrase (‘I said hello to him’). The Kwak'wala stem ńik-, translated as ‘say’ by Boas (1948:243), has a different argument structure. The recipient is the PRIMARY OBJECT of this stem, while the theme (the thing said) is marked as a SECONDARY OBJECT (see (21)).

Languages such as Kwak'wala for which recipients of a ditransitive predicate are consistently marked in the same way as the single object of a transitive verb have been described as having secundative alignment, in contrast with indirective languages for which the theme of a ditransitive predicate groups with the single object of a transitive verb, as is the case for the English ditransitive verb say (Malchukov, Haspelmath and Comrie 2010:3).

Many of Kwak'wala’s ditransitive predicates of transfer such as λíqala- ‘to name’, ćw- ‘to give’, and hələqa- ‘to pay’ follow a secundative pattern, as do stems describing communicative events such as ńik- ‘say’, ẃəłə- ‘ask’ and λiřə- ‘invite, call’.

Surprisingly, predicates of motion such as qas- ‘to walk’ and sin- ‘to paddle’ — which are intransitive in English — are transitive in Kwak'wala, with the destination marked as PRIMARY OBJECT. Although we might think of ‘walk’ as an intransitive verb which (in English) would take oblique arguments indicating destination (I walk to the store) or accompaniment (I walk with my brother), this is not the case for the Kwak'wala stem qas-. In an active construction, the primary object of the predicate qas- ‘walk’ has a predictable interpretation as the goal or destination. For example, see (4) (also (29) and (30)).

(4) Active: qas- ‘walk’ with PRIMARY OBJECT destination

<table>
<thead>
<tr>
<th>DISC</th>
<th>AUX</th>
<th>PRED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wə, ləʔlaʔi</td>
<td>qástalkə</td>
<td>ńaqʷala</td>
</tr>
<tr>
<td>Wə, ləʔlaʔi</td>
<td>qás-(ʔi)d-o=(i)χa</td>
<td>ńaqʷala</td>
</tr>
</tbody>
</table>

Well then it is said walk-MOM-away-OBJ1=T.DEM light-T.DEM
Well, then it is said, he walked toward the light. (B1906, III1.4)

The labelling of objects as ‘primary’ and ‘secondary’ thus also reflects the typological profile of ditransitive constructions in Kwak'wala, and acknowledges some resonance with the systems discussed for other languages by Dryer (1986) and Genetti (1997). Recognizing the secundative pattern of alignment in Kwak'wala exposes the syntactic properties of some of language’s passive suffixes, allowing us to identify a persistent contrast between the passive suffix -suʔ, which promotes primary objects, and the suffixes -ayu, -ano, and -əm, which promote secondary objects. These are described in 3.1.

In contrast to the three core arguments (SBJ, OBJ1, OBJ2), OBIQUE arguments are indicated in a prepositional phrase, constructed from a small set of grammaticalized predicates including la- ‘go’ and gχə- ‘come’, combined with deictically-appropriate demonstratives indicating

---

10 For this reason, it may be more appropriate to translate ńik- with the English verb ‘tell’, for which the recipient is also marked as a direct object. Thanks to Marianne Mithun for pointing this out.

11 English, like many languages, has different alignment patterns for different verbs.

12 Here, as in all examples, the morphophonemic parsing provided in the second line of each example follows the rules governing fusion of phonological segments as identified by Boas in his 1947 grammar (Boas 1947:210-215).

13 ‘Primary object’ and ‘secondary object’ are used here to refer only to morphosyntactic alignment in Kwak’wala grammar, not to the cross-linguistic generalizations proposed by Dryer in comparing direct/indirect object systems with primary/secondary object systems (Dryer 1986).
proximity, visibility, and (sometimes) possession. In example (5), the preposition la- includes prenominal =χ marking the argument gukʷ ‘house’\(^{14}\) and the third-person possessive enclitic =is (specifically, the form used for referents which are both invisible and distal).

(5) Prepositional phrase marking OBL

<table>
<thead>
<tr>
<th>Pred</th>
<th>SBJ</th>
<th>OBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>kwəʔ-tələʔi</td>
<td>Xaticən</td>
<td>laχis</td>
</tr>
<tr>
<td>kwəʔ-iləla=i</td>
<td>Xaticən</td>
<td>la=χ=is</td>
</tr>
</tbody>
</table>

sit-in.house-CONT=SBJ Xaticən (NAME) PREP=OBJ1=3.POSS house

Xaticən was sitting in his house. (B1947:282, CII 2.1)

As is true for many languages, certain types of lexical arguments in Kʷakʷala, such as places, are predictably oblique in active clauses.

Pronominal and prenominal flagging on the predicate and adnominal case-marking on arguments allow referent-tracking at a high level of detail. We, as well as Kʷakʷala speakers themselves, can thus confidently interpret the argument structure of most passivized predicates. Table 1 provides a table of pronominal and prenominal paradigms in order to enable readers to track referents in active and passive examples. Both sets of enclitics express an almost complete set of distinctions between subject (S), primary object (O₁) and secondary object (O₂), with the exception of the first-person forms (discussed below).

<table>
<thead>
<tr>
<th>Pronominal</th>
<th>SBJ</th>
<th>OBJ1</th>
<th>OBJ2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>=ən(ƛ)</td>
<td>---</td>
<td>=ən(ƛ)</td>
</tr>
<tr>
<td>1INCL</td>
<td>=ənʔs</td>
<td>---</td>
<td>=ənʔs</td>
</tr>
<tr>
<td>1EXCL</td>
<td>=ənuχʷ</td>
<td>---</td>
<td>=ənuχʷ</td>
</tr>
<tr>
<td>2ND</td>
<td>=əs</td>
<td>=uƛ</td>
<td>=us</td>
</tr>
<tr>
<td>3RD</td>
<td>=Ø</td>
<td>=q</td>
<td>=s</td>
</tr>
</tbody>
</table>

Table 1: Verbal enclitic pronouns and pronouns (adapted from Boas 1947:252)

I analyze the third-person subject pronominal as a morpheme with the shape -Ø; when third-person subject pronominals are tagged on the verb, there is no ambiguity about the intended referent, because all other types of marking occur. Number is only marked in first-person, which also makes a distinction between inclusive and exclusive forms. Aside from marking number, the first-person forms in Kʷakʷala are unusual in other ways. S and O₂ marking are identical for first-person. Meanwhile the cells marking first-person O₁ are ‘empty’, reflecting the fact that first-person primary objects are not indexed on the verb, but are instead expressed using a clause-final prepositional phrase derived from the verb gaχ- ‘come’. This speaker-oriented construction based on gaχ- ‘come’ echoes other-directed prepositional phrases marked with la- ‘go’, as seen above in example (5). Speakers know that a first-person primary object is expressed when the prepositional phrase gaχən occurs. The example below illustrates the encoding of the first-person

\(^{14}\) Prepositional phrases can be analyzed as embedded predicates, but I see this as a diachronic fact rather than a synchronic one.
primary object with the prepositional phrase *gaχən* ‘to me’.

(6) First person primary object in a prepositional phrase

<table>
<thead>
<tr>
<th>la-µisas</th>
<th>λιqala-λας</th>
<th>aλaŋm</th>
<th>gaχən</th>
</tr>
</thead>
<tbody>
<tr>
<td>la-µis=əs</td>
<td>λιqala-λ=ς</td>
<td>aλaŋm=ən</td>
<td></td>
</tr>
</tbody>
</table>

AUX-CONN-2.SBJ name-FUT=OBJ2 wolf PREP=1.SBJ

And so you will name me (with) wolf.15 (Anderson 2005:17)

The secundative alignment of the stem *λιqala*- ‘name’ is clear from the secondary object case marking of the theme *aλaŋm* ‘wolf’, the name given to the speaker. If the person (or object) being named were second- or third-person, the primary-object status of the speaker would be encoded on the verb with -uλ (2.OBJ1) or -q (3.OBJ1), but for a first-person argument, the primary-object status becomes clear through the use of the phrase *gaχən*.

In addition to the pronominal and prenominal paradigms presented above, additional sets of third-person demonstrative forms express a six-way deictic contrast: demonstrative enclitics distinguish proximal, medial and distal locations, with visible and invisible status encoded for each. These distinctions are just as fully expressed in sets of possessive suffixes (which also encode a distinction between a subject possessor and a non-subject possessor), as well as in forms used for embedded purpose clauses. For those who wish to follow reference tracking in examples, the following additional tables are included in the appendix: (i) third-person ‘verbal’ (affixed to the predicate) demonstrative enclitics for subjects and prenominal forms; (ii) third-person pronominal demonstrative enclitics for subjects, primary and secondary objects; (iii) subject/primary object combinations; (iv) subject/secondary object combinations; (v) possessive forms; (vi) purposive clause forms; and (vii) special possessive marking in pronominal predicate clauses. In glossing third-person demonstrative forms, I will comment where the form allows one to disambiguate between multiple third-person arguments in the clause.

The next section provides a description of the morphosyntactic and combinatorial properties of passive suffixes in *Kʷaḵʷala*.

3 Passive Morphosyntax in *Kʷaḵʷala*

Before moving on to the descriptive portion of the paper, I briefly address my use of the term ‘passive’. Much has been written about what should be considered ‘passive’, and traditions of interpretation vary (Comrie 2008, Fox and Hopper 1994, Givon 1994, Klaiman 1991, Payne 1997, Siewierska 1984, Shibatani 2003, *inter alia*). In what follows, *Kʷaḵʷala* passives are described in terms of both morphosyntactic and discourse-contextual properties.

Comrie 2008 provides some useful cross-linguistic criteria for identifying passive phenomena, and as we will see, *Kʷaḵʷala* passives conform to these tendencies. The first criterion is markedness: as is true of morphological passives in other languages, *Kʷaḵʷala* passive constructions are formally ‘marked’ by the addition of a derivational affix to a base stem which, in its unmarked state, is considered active. Furthermore, the direction of derivation is

---

15 As one can see by comparing the secondary object marking of the theme *aλaŋm* ‘wolf’ with the first-person primary object marking, *λιqala*- ‘to name’, like the verb *iik*- ‘to say’, marks the recipient (R) of a name (‘me’) as the primary object, and the name being bestowed upon the recipient (T) (as the secondary object).

16 Certain ceremonial objects, such as coppers and canoes, also receive names.
from active to passive\textsuperscript{17}, and passive forms are less frequent than active forms in connected discourse. These patterns of distribution will be clear in the description below. Secondly, Comrie suggests that passives share a quality of being ‘P-oriented’ rather than ‘A-oriented’. K\textsuperscript{wa}k\textsuperscript{wa}la passives can be described as P-oriented: they promote non-A arguments to subject position, whether for reasons of topic continuity or to make an argument accessible to relativization.

Cross-linguistically, passive suffixes reduce the transitivity of an active predicate stem and restructure the case-marking of arguments in certain cross-linguistically predictable ways. If an active predicate has subject ‘A’, the addition of a passive suffix to an active stem allows the promotion of non-‘A’ arguments to subject position and the demotion of ‘A’ from subject position. In many languages, active transitive predicates are made intransitive by passivization. But because K\textsuperscript{wa}k\textsuperscript{wa}la has three core arguments, and because demoted subjects are marked with SECONDARY OBJECT case (=sa), passivized predicates are not necessarily intransitive, as they might be in languages with only two core arguments. While K\textsuperscript{wa}k\textsuperscript{wa}la passive morphemes reduce transitivity, they do not seem to reduce valence. In many passive clauses, the case-marking is restructured, but all three arguments would still be considered core according to the formal criteria established above. On the other hand, demoted subjects, while marked as secondary objects, are not obligatory, as we will see in many of the examples provided below. By this criterion, one might identify wish to separate demoted subjects from other arguments marked as secondary objects.

K\textsuperscript{wa}k\textsuperscript{wa}la is typologically unusual for having several passivizing suffixes with different functions.\textsuperscript{18} Passive morphology has been well described for another Wakashan language, Nuu-chah-nulth, but is limited to a single form with broad functional scope (Nakayama 1997), in contrast to the six forms used in K\textsuperscript{wa}k\textsuperscript{wa}la.\textsuperscript{19} The passive suffixes of K\textsuperscript{wa}k\textsuperscript{wa}la are presented in Table 2.

| PRIMARY OBJECT | -suʔ |
| SECONDARY OBJECT | -ayu, -əm, -ano |
| EXPERIENTIAL | -ɬ |
| LOCATIVE | -ʔas |

Table 2: K\textsuperscript{wa}k\textsuperscript{wa}la Passive Suffixes

These morphemes occur in contrastive distribution in a slot suffixed to the predicate. The first two rows list syntactically-selected passive forms, while the third and fourth rows list semantically-selected passive forms.

The PRIMARY OBJECT passive -suʔ is used to promote primary objects of an active transitive or ditransitive predicate to subject status, while the SECONDARY OBJECT passives -ayu, -əm, and -ano promote secondary objects with an active ditransitive predicate to subject status (the variation among these three forms is discussed below). As mentioned, many K\textsuperscript{wa}k\textsuperscript{wa}la

\textsuperscript{17} This is clear from the boundary effects of certain ‘hardening’ and ‘weakening’ suffixes on preceding coda consonants (Boas 1947:226).

\textsuperscript{18} Some Philippine languages, such as Tagalog, have voice systems also described as having multiple passives. A comparison between the K\textsuperscript{wa}k\textsuperscript{wa}la system and similar Austronesian systems is beyond the scope of this paper but is in preparation.

\textsuperscript{19} Nakayama finds ‘passive’ to be an inadequate term to capture the full range of functions of the ‘-at’ suffix he describes in Nuu-Chah-Nulth.
Passive Constructions in Kwak'ala
ditransitive predicates with meanings such as ‘say’, ‘give’, and ‘pay’ mark the recipient as primary object and the theme — the item said, given, or paid, for example — as secondary object. Identifying the pattern of secundative alignment in Kwak’ala thus reveals the consistently syntactic distribution of Kwak’ala passive suffixes promoting primary and secondary objects in the data explored below.

Boas’ discussion of -suʔ and -ayu partially captures the syntactic character of these morphemes. In keeping with Boas’ analysis of primary objects as “objective” and secondary objects as “instrumental”, he labelled -suʔ as the “passive governing (the) objective form” and -ayu, -ano, and -əm as “passive of verbs with instrumental” (Boas 1947:242). On the other hand, Boas later appeals to a semantic distinction, saying “the passives in -əm and -ayu designate the thing used for doing something, while -suʔ designates the thing to which something is done”, while -ɬ is “(the) passive of verbs expressing sensations and mental actions; also sensations produced by some outer action” (Boas 1947:270).

In his 1980 paper, Levine also described the function of these suffixes in partly syntactic terms, saying that “[t]he suffixes -suʔ and -ayu belong to a class of morphemes I refer to as focus elements, to emphasize that these forms shift attention to various participants in the situation, which is specified by the meaning of the stem […]. (F)ocus morphemes permit the apparent promotion to subject status of NPs containing -χ and -s as determiners” (Levine 1980:242). On the other hand, Levine renames -suʔ a ‘goal focus passive’, -ayu an ‘instrument focus passive’, -ɬ a focus morpheme referring to ‘lack of control’, and -ʔas a ‘location focus’ morpheme, suggesting a semantically-grounded interpretation of these forms.20 Levine does not address -əm and -ano, the other secondary-object promoting passives.

I argue here that both syntactic and semantic criteria are necessary for a complete description of the Kwak’ala passive paradigm. As is clear from the terms employed in Table 2, the data show the PRIMARY OBJECT and SECONDARY OBJECT passives to be syntactically-selected, based on the argument structure of an active predicate stem. Meanwhile, one must look beyond syntax to explain the distribution of the remaining passives. Section 3.1 discusses the use of -suʔ to promote PRIMARY OBJECTS to subject status and the use of -ayu, -ano and -əm to promote SECONDARY OBJECTS to subject status. Section 3.2 discusses the semantically-governed use of the EXPERIENTIAL -ɬ and the LOCATIVE -ʔas.

3.1 Syntactic Roles

Examples (7) and (8) below illustrate a contrast between active and passive forms of the same predicate, showing how the Kwak’ala passive suffix -suʔ allows the promotion of a PRIMARY OBJECT into SUBJECT position.

---

20 I believe Levine intends the term ‘focus’ to refer to the shared attention of the speaker and listener captured by the subject position in a Kwak’ala clause, rather than concepts of topic and focus as commonly used in discussing information structure in discourse.
(7) Active: wəƛ- ‘ask’

ƛálaʔi  wəƛí  Gíχdənaxa  bagʷən̓ámi: “Məʔinəχəʔas?”
ƛá-laʔi  wəƛ=í  Gíχdən=χa  bagʷən̓ámi: “Məʔinəχəʔas?”
SEQ-QUOTE ask=SBJ Gíχdən (name)=OBJ1 man     “Of what tribe are you?”

Then Gíχdən asked the man, “What tribe are you from?” (B1895: M665.10)

(8) Passive: wəƛ(a)- ‘ask’ with primary object passive -suʔ?

ƛálaʔi  wəƛásuwa: “Másus yálagilišax?”
ƛá-laʔi  wəƛá-suʔ-a  Más=us yála-giš=ax
SEQ-QUOTE ask-PASS-T.DEM Q=2.SBJ do-TR-on.beach=DEM

Then he was asked: “What are you making on the beach?” (B1895: M666.23)

In example (7), the protagonist of the story, Gíχdən, asks a question. His status as subject is clear from the prenominal subject-marking clitic =i preceding his name. The man of whom he asks the question is χ-marked as a primary object with the prenominal enclitic =χa. Later in the story, a question is asked of Gíχdən; as the recipient of a question, Gíχdən would be the primary object of the active predicate wəƛ(a)- ‘ask’. Instead, the PRIMARY OBJECT passive suffix -suʔ in example (8) allows him, as the protagonist of the story, to remain in subject position. Meanwhile, the person asking the question does not appear.

Example (9) is from the same story containing examples (7) and (8) above, and this example illustrates the use of the SECONDARY OBJECT passive morpheme -ayu represented in bold type. At the moment excerpted below, the protagonist Gíχdən has finally found the final magical treasure he has been seeking, the decapitated heads of his rival chiefs; the decapitated heads (also in bold) are the subjects of these sentences.

(9) Passives: čəw- ‘give’ and tikʷ- ‘hang.on.body’ with secondary object passive -ayu

Gáχlaʔi  čøyida  qágukʷ  laχ  Gíχdən.
gáχ-laʔi  čəw-ayu=i=da  qágukʷ  laχ  Gíχdən
come-QUOTE give-PASS=SBJ=DEF heads PREP Gíχdən (name)

Now it is said the heads were given to Gíχdən.

Laʔám  tikʷiʔidayu  laχ  Gíχdən.
La-ʔám  tikʷ-ʔid-ayu=Ø  laχ  Gíχdən.
SEQ-OI hang.on.body-MOM-PASS=3.SBJ PREP Gíχdən (name)

Then they (the heads) were hung onto Gíχdən’s body. (B1895, M667.6-667.7)

The argument structures of these two predicates (čəw- ‘give’ and tikʷ- ‘hang.on.body’) require the secondary object passive -ayu rather than the primary object -suʔ in order to promote the heads to subject position. The predicate čəw- ‘give’ is a ditransitive predicate with secundative alignment, for which the recipient is marked as primary object and the theme as secondary object; the SECONDARY OBJECT passive -ayu is thus used to promote the qágukʷ ‘heads’ given to Gíχdən, rather than Gíχdən himself (the recipient). In the first clause, the lexically-expressed subject is preceded by the subject enclitic =i; in the second clause, the third-person subject is
Passive Constructions in Kʷaḵʷala

represented with a zero pronominal enclitic. Some elicited examples provided by Levine offer useful minimal pairs to illustrate the contrasting effects of the passive suffixes -suʔ and -ayu on a predicate and the surrounding clause. The first sentence is active.

(10) Active: SBJ is ‘child’

PRED SBJ O1 O2
nāpīdida gənānəmχa gukʷsa tisəm.
nāpīd=ı=da gənānəm=χa gukʷ=sa tisəm
throw=SBJ=DEF child=OBJ1 house=OBJ2 rock
The child hit the house with a rock by throwing.
(The child pelted the house with a rock. - DR; Levine 1980:241)

The verb nāpīd- ‘throw’ is perhaps closer to the English word ‘pelt’; the recipient gukʷ ‘house’ is the primary object, while the thing thrown (tisəm, ‘rock’) is the secondary object. When the primary object passive suffix -suʔ is added to the predicate stem nāpīd- ‘throw’, we get example (11), with gukʷ ‘house’ promoted to subject position, and the demoted subject marked with =sa as a secondary object.

(11) PRIMARY OBJECT passive: promoted SBJ is ‘house’

PRED SBJ O2 O2
nāpīdsowıda gukʷsa gənānəmsa tisəm.
nāpīd-suʔ=ı=da gukʷ=sa gənānəm=sa tisəm
throw=PASS=SBJ=DEF house=OBJ2 child=OBJ2 rock
The house was hit by a rock thrown by the child.
(The house was pelted with a rock by the child. - DR; Levine 1980:241)

The house, which would be a primary-object marked with =χa in an active predicate, here becomes the subject of the sentence with the addition of the primary object promotion morpheme -suʔ. The rock-throwing gənānəm ‘child’, the subject of the active sentence, is demoted to secondary object status. Meanwhile, the rock tisəm retains secondary object status, leading to a clause with two secondary objects with different syntactic origins. In many languages, the actor of a passivized transitive verb is marked as an oblique or non-core argument, but in Kʷaḵʷala, the actor of a verb passivized with -suʔ is demoted to secondary object marking, rather than a prepositional phase. Should one consider the erstwhile subject of a

21 In both clauses, Giχən appears in a prepositional phrase. While this seems natural to English speakers, the argument structure of the Kʷaḵʷala predicate čəw- ‘give’ actually marks the recipient as primary object. In this case, the presentation of the primary object recipient Giχən in a prepositional phrase (rather than case-marked as a primary object) is due to the extraposition of a primary object in some double-object clauses, described by Boas: “Since Kwakiutl transforms the direct object -q into the indirect object laq whenever the verb takes an instrumental s, these forms must be considered as a substitute for the direct object …” (B1947:283).

22 Thanks to Bernard Comrie for this gloss.

23 The ordering of these two ‘s-marked’ objects is also interesting; according to Levine, the order given, with animate (but adjunct) ‘child’ preceding inanimate ‘rock’, is the preferred order. Further research is underway to explore the strength of this preference among speakers.
passive construction to be a core argument? Despite the identity of the surface form =sa marking secondary object arguments and demoted subjects, other criteria might lead us to identify demoted subjects as adjuncts; as we saw above in examples (7), (8), and (9), the demoted subject can be omitted. More work is needed to determine the syntactic status of the demoted subject.

Example (12) shows that the use of the SECONDARY OBJECT promoting suffix -ayu as a passive suffix allows the rock to be promoted to subject instead.

(12) SECONDARY OBJECT passive: promoted SBJ is ‘rock’

\[
\begin{align*}
\text{plodey}=\text{sid} & \quad \hat{\text{tisom}}=\text{xa} & \quad \text{guk}^{\text{w}}=\text{sa} & \quad \text{go}\text{nanom}. \\
\text{plod-ayu}=\text{id}=\text{da} & \quad \hat{\text{tisom}}=\text{xa} & \quad \text{guk}^{\text{w}}=\text{sa} & \quad \text{go}\text{nanom}.
\end{align*}
\]

throw-PASS=SBJ=DEF rock=OBJ1 house=OBJ2 child

The rock was (what was) thrown at the house by the child. (Levine 1980:241)

In contrast to the double secondary-object marking of example (11), here the house remains as an χ-marked primary object of the passivized predicate; the demoted subject, the child, is again marked as a secondary object with =sa. Knowing that Boas termed secondary objects ‘instrumental case’, we can understand why he called these suffixes ‘instrumental passives’: in this particular case, the rock fits an analysis based on its semantic role as the instrument of throwing. However, as we will see below, -ayu and the other so-called ‘instrumental passives’ are not governed by the semantic role of the promoted argument, but by its syntactic role.

Do the three SECONDARY OBJECT passive morphemes contrast semantically or otherwise? Boas doesn’t find recoverable differences between them; synchronically, their distribution seems lexically-determined and unpredictable. There may have been a historical difference; Boas describes -əm as being restricted from combining with stems with certain endings, and -ano as used with stems that form transitivies with the MOMENTANEOUS allomorph -nd (Boas 1947:270). But of -ano, Boas says “this suffix is used with a few words only, and is not freely available” (Boas 1911:36). As we will see in the discussion of frequency in discourse, -ayu is significantly more frequent than either of the other forms, and there is some indication that in the modern language, it is becoming the dominant alternative to -su?, retaining its productivity while -əm and -ano become increasingly restricted. Further research is exploring distributional contrasts within the group of secondary object passives.

Samples from narrative discourse published by Boas and Hunt confirm the syntactic distribution of these passivizing suffixes, with -su? governing the primary object case, and -ayu governing the secondary object case, as we can see below by comparing active and passive forms for predicates with consistent argument structure. Several of these clauses are explored in detail below, in examples (13) through (31). These examples demonstrate the consistency with which -su? can be predicted to promote a primary object and -ayu, -əm, and -ano can be predicted to promote a secondary object, whatever the semantic role of the argument being promoted.

In active examples (13) and (14), wací ‘dog’ is the subject of the clause, the one who bites; the patient of the predicate ʔax- ‘bite’ (an unspecified third-person) is marked as a primary

\(^{24}\) Here Levine notes that the order of ‘house’ and ‘child’ is interchangeable. It would be surprising to find a secondary object before a primary one in the clause, but this could be a result of the animacy of the child. Current research is pursuing these questions.
Passive Constructions in Kʷakʷala

object, indicated with the pronominal suffix =q.25

(13) Active sentence: qəx- ‘bite’

AUX.PRED   PRED   SBJ
laʔam    qəxʔidida    wacaq
laʔam    qəxʔiđ=da    wac=(a)q
SEQ-DE    bite-MOM=DEF  dog=3.OBJ1

Now the dog bit him. (B1947:286, CIII12.19)

(14) Active sentence: qəx- ‘bite’

PRON.PRED   SBJ  PRED
hiʔida    wacʔi    qəxʔidadq
hiʔam=i=da    wacʔi    qəxʔiđ=(ə)q
3.PRED-DE   dog-DEM  bite-MOM=3.OBJ1

That is the dog that bit him. (B1947:286)

(15) Passive sentence: qəx- ‘bite’ with PRIMARY OBJECT passive -suʔ

PRON.PRED   PRED  SBJ  O2
hiʔam    qəxʔicəʔiʔida    gəنانəмəsəʔa    ʔwacʔi
hiʔam=i=da    gəنانəмə=əʔa    ʔwacʔi
3.PRED-DE   bite-MOM-PASS=DEF  child=OBJ2  dog

That is the child (who was) bitten by the dog. (B1947:286)

In example (15), the PRIMARY OBJECT passive suffix -suʔ attached to the predicate qəx- ‘bite’ promotes gəنانəм ‘child’, the one bitten, to subject status. One can see the effect of the passive suffix -suʔ in both the word order, which has moved child into subject position following the predicate, and in the secondary object marking of the dog ʔwacʔi, who bit the child. Examples (14) and (15) are translated with initial ‘that’ in English, indicating additional discourse factors at work in relation to the use of passive here; these are actually predicative third-person pronouns. I return to the description of pronominal predicates in section 4.2.

As we saw in the earlier example (9), the stem čəw- ‘give’ also marks recipient as primary and theme as secondary. Example (16) shows the predicate čəw- in an active form with secondary object marking of the theme yaxʷsəʔiʔa ‘a bad thing’. The recipient in this clause occurs in a prepositional phrase as a result of extraposition of primary objects in certain syntactic contexts.27

---

25 Examples (14) and (15) are complicated by the use of the person-marking predicate hi-, also described as a ‘verbal form’ of an independent pronoun, which leads to an appositive clause for which the second predicate is a relative-clause type complement. For this reason, the subject precedes the embedded predicate ‘bite’ (Boas 1947:258). These clauses are described in section 4.2.

26 Examples (13) and (14), in which a dog is subject-marked without triggering any special morphology in relation to a primary-object-marked human, suggest that passivization is not obligatory according to an animacy hierarchy, at least in the case of dogs biting humans. Inverse alignment systems are not uncommon in the region but the alignment here does not suggest such a pattern. For further argumentation against an inverse pattern in Kʷakʷala, see section 3.3 below. Further research will pursue this hypothesis.

27 Close observers will notice that in the active clause, the pronominal mention of the recipient is actually marked as an oblique in the clause-final prepositional phrase laq, rather than with the primary object pronominal morpheme =q on the predicate, as we would expect. Boas notes that while subject, primary, and secondary arguments...
(16) Active clause: ṭə-‘give’

```
AUX.PRED  PRED  O2     OBL (O1)
  ła-mi  čása  ｙαχʷsə́ma  laq
  la-ｍ=Ø=i  ṭə́w=ša  ｙαχʷsə́ma  la-q
```

Then he gave him something bad. (B1947:342)

However, when the predicate ṭə-‘give’ is passivized with PRIMARY OBJECT -suʔ, the recipient becomes the subject, and the demoted subject ｙαχʷsə́ma retains secondary object-marking.

(17) Passive clause: ṭə-‘give’ passivized with -suʔ

```
AUX.PRED  PRED  O2
  laʔmyə́nti  čásuʔsa  ｙαχʷsə́ma
  la-ｍ-χənt=Ø=i  ṭə́w-suʔ=ša  ｙαχʷsə́ma
```

Then he was evidently given something bad. (B1947:342, CII32.13)

On the other hand, when ṭə-‘give’ is passivized with SECONDARY OBJECT -suʔ, the theme ｙαχʷsə́ma is promoted to subject.

(18) Passives: ṭə-‘give’ passivized with -ayu

```
Gáχəʔi  čáyida  qáguʔw  laχ  Gíχdən.
  gáχ-laʔi  čə́w-ayu=i=da  qáguʔw  laχ  Gíχdən
```

Now it is said the heads were given to Gíχdən.

As mentioned earlier and seen in the examples below, the predicate ṭʊik- (translated as ‘to say’ by Boas but glossed as ‘to tell’ here) also has secundative alignment and marks hearers (the ‘recipient’ of the words spoken) with the primary object case. The active example (2) is reproduced below to illustrate this secundative alignment.

(19) Active clause: ṭʊik- ‘to say, tell’

```
  ṭʊikə̃laq
  ṭʊik=ə̃λəq
```

say=1s.SBJ>3.OBJ1

I told him. (B1947:281, CX12.19)

When ṭʊik- is passivized with the PRIMARY OBJECT passive -suʔ, the hearer qəmtalať is promoted...
Passive Constructions in K'wak'wala

to subject, immediately following the predicate and preceded by the prenominal subject-marker =i.

(20) Passive clause: ńik- ‘to say, tell’ passivized with -su?

```
<table>
<thead>
<tr>
<th>PRED</th>
<th>SBJ</th>
<th>O2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ńixwsuʔi</td>
<td>q̓ámtalaʔas</td>
<td>q̓aniqilakʷ.</td>
</tr>
<tr>
<td>ńik-suʔ=fa=i</td>
<td>q̓ámtalaʔ=as</td>
<td>q̓aniqilakʷ.</td>
</tr>
</tbody>
</table>
```
say-PASS-QUOT=SBJ name=OBJ2 name
q̓ámtalaʔ was told by q̓aniqilakʷ. (B1947:256, III100.19)

As we would expect, the demoted subject q̓aniqilakʷ, who is telling q̓ámtalaʔ something, is demoted to secondary object status and appears in clause-final position preceded by the secondary object enclitic =as. The example below shows the same verb ńik- ‘to say, tell’, passivized with the SECONDARY OBJECT passive -əm (within a second-person purposive clause marked with the subordinator qaʔ) to promote the theme rather than the hearer.

(21) Passive clause: ńik- ‘to say, tell’ passivized with -əm

```
<table>
<thead>
<tr>
<th>SUBORD</th>
<th>PRED</th>
<th>O2</th>
</tr>
</thead>
<tbody>
<tr>
<td>...qaʔs ńígəmusaʔs</td>
<td>fəwənəmʔus</td>
<td></td>
</tr>
<tr>
<td>...qaʔs ńik-əm=ʔ=us=əγs</td>
<td>fəwənəm=əγus</td>
<td></td>
</tr>
<tr>
<td>2.PURP say-PASS=3.SBJ=2.OBJ2=2.POSS husband=2.PURP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
...in order that it be said to you by your husband. (B1947:270, CX249.40)

As we see in this example, the subject-promotion of the theme (the thing being said) requires the SECONDARY OBJECT passive suffix -əm rather than the PRIMARY OBJECT passive -suʔ, but instead.

A similar contrast is evident with the stem həlaq- ‘pay’. When a primary object-marked recipient becomes the subject, as in examples (20) and (21), the suffix -suʔ is used. In (20) the secondary object s-marking of the theme (the four blankets) is maintained, as can be seen from the =sa marking on the predicate həlaqasusə.

(22) Passive clause: həlaq(a)- ‘pay’ passivized with -suʔ

```
<table>
<thead>
<tr>
<th>AUX</th>
<th>PRED</th>
<th>ADJ</th>
<th>O2</th>
</tr>
</thead>
<tbody>
<tr>
<td>le  həlaqasusə</td>
<td>mú̕xsə</td>
<td>Ḵə̕lasgəmaʔ</td>
<td></td>
</tr>
<tr>
<td>le  həlaqa-suʔ=ʔ=sa</td>
<td>mu-γsə</td>
<td>Ḵə̕lasgəm=ʔaʔ</td>
<td></td>
</tr>
<tr>
<td>CONN</td>
<td>pay-PASS=3.SBJ=OBJ2</td>
<td>four-POSS.NSUB blanket=DEM</td>
<td></td>
</tr>
</tbody>
</table>
```
She is paid with four blankets. (B1947:270, CX249.40)

28 The case-marking of the second-person hearer as a secondary object with =us is unexpected. We would expect the hearer to be marked as a primary object. In addition, the demoted subject, təwənəm ‘husband’, is not case-marked at all. I do not have an explanation for this, although several factors could be at work: the subordinate syntax of the purposive clause, the second-person status of the hearer, and/or the possessed status of the husband.
(23) Passive clause: ḥəlaq(a)-‘pay’ passivized with -su?

AUX PRED
le ḥəlaqasu?
le ḥəlaq-su?=Ø
CONN pay-PASS=3.SBJ
Then he is paid. (B1947:241)

However, when the s-marked theme, the payment, becomes the subject, as in example (24), the secondary object passive suffix -əm is used to promote the secondary object.

(24) Passive sentence: ḥəlaq(a)-‘pay’ passivized with -əm

PRED SBJ PRED O1
hiʔəmxəʔáwise wàxə ḥəlaqəmaxa mámaʔutšila
hiʔ-əmxəʔáwis-i wàxə ḥəlaq-əm=χa mámaʔutšila
3P.PRED-OI-also-CONN-SBJ amount pay-PASS=OBJ1 midwife
That is also the amount that was paid to the midwife. (B1947:270, R 670.92)

The amount wàxə is subject-marked with the prenominal enclitic =i, and the enclitic =χa preceding mámaʔutšila ‘midwife’ marks it as retaining primary object status. As we saw with examples 14 and 15, the subject, the amount paid, or wàxə, precedes the stem ḥəlaq-‘pay’, rather than following it, because it is the subject of the verbal independent pronoun hi-. These ‘verbal pronouns’ are discussed in section 4.2.

The derived verb ḥəmjil(a)-‘to feed’, constructed from the stem ḥəm-‘eat’ and the transitivizing suffix -gil(a), again illustrates the contrast between passivization of the primary object with -su? and passivization of the secondary object with a secondary object promoting form, in this case -ayu. In (25), Stone Body, who is being fed, is the subject of the predicate ḥəmjil(a)-‘feed’ passivized with the primary object passive -su?.

(25) Passive sentence: ḥəmjil(a)-‘feed’ passivized with -su?

DISC ADV PRED SBJ
li maʔɪpəná ḥəmjilasəwí tisəmgiti
li maʔ-ɪpəná ḥəmjil(a)-su?=i tisəmgiti=i
then two-times feed-PASS=SBJ Stone.Body=DEM
Then Stone Body was fed twice. - DR29 (B1947:270, CIII220.30)

In example (26), the subject of the predicate passivized with the secondary object passive -ayu must be understood as the food, rather than the recipient of the food.

29 Boas translates the first sentence as “Then Stone Body was twice given to eat,” and the second as “… to be given to eat (with it) to his tribe.” The translations provided above are my own. Boas’ translations of this and the following examples are based on a gloss of ḥəmjil as ‘give to eat’, which allows for a less-intuitive translation. The English verb ‘feed’ can have secundative alignment, marking recipient as primary (direct) object (‘He feeds someone with something’), or indirective alignment (‘He feeds something to someone’).
Passive Constructions in Kwakwala

(26) Passive: həmgil(a)- ‘feed’ passivized with -ayu

PURP  PRED  O2
...qa?  həmgilayusəχis  gūkʷəloti
...qa?  həmgil-ayu-s=Ø=iχ=is  gūkʷəlot=i
...3.PURP feed-PASS=-?=3.SBJ=OBJ1=3.POSS tribe=DEM
...in order that it be fed to his tribe. - DR (B1947:270, CIII 7.6)

The recipient, gūkʷəlot ‘(his) tribe’ retains case-marking as the primary object with the prenominal enclitic =iχ.30

The verb gʷəqəq(a)- ‘pour over’ also has secundative alignment, and two examples illustrate the contrast between subject-promotion of the primary object with -suʔ and promotion of the secondary object with -əm. In the first example, passivized with -suʔ, the subject is understood as the ‘recipient’ or ‘destination’ of the liquid being poured while the material being poured is s-marked as a secondary object.

(27) Passive sentence: gʷəqəqa- ‘to pour over’ passivized with -suʔ

PRED  ADJ  O2
gʷəqəqa-suʔsa  wədaʔstá  wapa
gʷəqəqa-suʔ=Ø=sa  wədaʔstá  wap=a
pour.over-PASS=3.SBJ=OBJ2  cold  water=T.DEM
It is poured over with cold water. (B1947:270, R516.16)

In contrast, in example (28), passivization with -əm results in the subject being understood as the liquid being poured.31

(28) Passive sentence: gʷəqəqa- ‘pour over’ passivized with -əm

PURP  PRED  PREP  OBL
...qaʔs  gʷəqəqəmiʔ  laχa  līqəstənʔiʔ
...qaʔs  gʷəqəqə-əm=Ø=iʔ  la=χa  līqəstən=iʔ
...3.PURP pour.over-PASS=3.SBJ=DEM.OBJ1  PREP  seaweed=DEM
...it is poured on the seaweed. (B1947:270, CIII 7.6)

Finally, the pair of examples below illustrate the interpretive contrast between -suʔ and -ayu marking on the motion predicate qas- ‘walk’. For active motion predicates, the destination is encoded as primary object In (29), qas- is passivized with PRIMARY OBJECT -suʔ. The protagonist is being pursued. (That is, he is the destination of those ‘walking towards’ him.) He is expressed as a third-person subject, marked with -Ø, rather than the pronominal -q used to mark a primary object referent.

---

30 I am not sure what interpretation we should give to the morpheme -s following the -ayu in this predicate. It may be a pronominal marker indicating secondary object, in which case the translation should be ‘in order that it be fed to his tribe by him/them’; but in most sentences like this we would find the primary object ‘tribe’ extraposed to a prepositional phrase.

31 The seaweed, līqəstən, the destination (or ‘recipient’) of the liquid, is marked in a prepositional phrase here. Again, this may be a result of the extraposition of primary object to a clause-final prepositional phrase.
(29) Passive: *qas-* ‘walk’ passivized with *-su?*

```
DISC   DISC   PRED
Loʔám lawis qásʔidsaʔa
Loʔám lawis qásʔ-idi-suʔ-Ø-a
Then it is said walk-MOM-PASS-3.SBJ-T.DEM
Then it said they went after him.
(Then, it is said, he was pursued by them. - DR; Boas 1895, M727.17)
```

On the other hand, in (30a) and (30b), *qas-* is passivized with *-ayu*, and in both examples the interpretation is that the protagonist is being walked by another character.

(30) Passive: *qas-* ‘walk’ passivized with *-ayu*

```
a. DISC   PRED    O2
  LálaʔI qásʔidayusa wiwaʔokʷ
  LaʔlaʔI qásʔ-idi-ayu=sa wiwaʔokʷ
  SEQ-QUOT walk-MOM-PASS=OBJ2 wolf
  Then he was walked by the wolf. (B1895: M 666.21)

b. DISC   PRED
  LálaʔI qásʔidayu
  LaʔlaʔI qásʔ-idi-ayu=Ø
  SEQ-QUOT walk-MOM-PASS=3.SBJ
  Then it is said they walked with him

  DISC   OBL
  lālaʔI laŋa ƛiŋkala
  lālaʔI la=ŋa ƛiŋkali-a
  SEQ-QUOT PREP=OBJ1 beating.of.boards-T.DEM
  then to the beating of boards.
  (Then it is said he was walked (then) to the beating of boards. - DR; B1895: M 683.4)
```

I have not yet found an active example of *qas-* with a secondary object, but we might predict that a person being made to move in a certain way, or a pet being walked by its owner, would be encoded as the secondary object.

Examples (13) to (30) show that the distribution of the PRIMARY OBJECT passive morpheme *-su?* and the SECONDARY OBJECT passive morphemes *-om*, *-ayu*, and *-ano* reflect the case-marked syntactic role of an argument rather than its semantic role. Further evidence can be found in the first clause of the example below, in which a semantic instrument (*λʔmgayu* ‘wedge’) is promoted to subject status, but with the PRIMARY OBJECT passive morpheme *-su?*. If the distribution of *-su?* and *-ayu/-om/-ano* morphemes were based on semantic categories, the first predicate *ʔaľ*- ‘take’ would be marked with *-ayu*, *-om*, or *-ano*, one of the passive morphemes identified as ‘instrumental’ by Boas, in order to promote the semantic instrument *λʔmgayu* to subject status. But the distribution is syntactically rather than semantically determined, as we see here.
Passive Constructions in K\textsuperscript{akwala}

(31) Passive: ʔəχ- ‘take’ passivized with -suʔ, ʔəχʷ ‘beat (cedarbark)’ passivized with -ayu

<table>
<thead>
<tr>
<th>AUX</th>
<th>PRED</th>
<th>PREP</th>
</tr>
</thead>
<tbody>
<tr>
<td>la</td>
<td>ʔəχʔidsəwida</td>
<td>laq</td>
</tr>
<tr>
<td>la</td>
<td>ʔəχʔ-ʔid-suʔ=i=da</td>
<td>la=q</td>
</tr>
</tbody>
</table>

Now is taken the wedge,

<table>
<thead>
<tr>
<th>PURP</th>
<th>PRED</th>
<th>PREP</th>
</tr>
</thead>
<tbody>
<tr>
<td>qaʔs</td>
<td>təlχʔwidáyuwiʔ</td>
<td>laq</td>
</tr>
<tr>
<td>qaʔs</td>
<td>təlχʷʔ-ʔid-ayu=iʔ</td>
<td>la=q</td>
</tr>
</tbody>
</table>

and it is used for beating it (the cedarbark). (B1947:312, R296.82)

The active predicate ʔəχ- ‘take’ requires primary object marking for the thing being taken; passivization with -suʔ promotes this object — in this case a wedge, ƛəmgayu. Semantic interpretations of -suʔ would identify it as an ‘objective passive’ (Boas 1947:270) or a ‘goal focus’ morpheme (Levine 1980b:58), in contrast with -ayu, -əm, or -ano, identified as the passives used to promote the “the thing used for doing something” or the “instrumental” (Boas 1947:270). The semantic role of the wedge taken here would certainly fit into the category ‘instrument’ and one might expect the passivizing suffix to be one of the ‘instrumental’ suffixes. However, because of the argument structure of the predicate ʔəχ- ‘take’, for which the wedge is a primary object of the stem, it must be passivized with -suʔ. In the subsequent predicate ʔəχʷ- ‘beat cedar bark’, the same wedge is passivized with -ayu, because of the argument structure of the active lexeme ʔəχʷ- ‘beating it (the cedarbark)’. (B1947:312, R296.82)

The suffix -ayu also occurs in the lexicalized word for wedge, ƛəmgayu, derived from the stem ƛəmk- ‘to split wood with wedge’. Here, ƛəmgayu clearly functions as an argument; it follows the prenominal subject-marker =i attached to the predicate. The suffix -ayu functions in this form as an instrumental nominalizer. The question of whether these are the same morpheme or two polysemous morphemes is discussed below in section 3.4.

The section above argues that syntactic properties govern the distribution of passive morphemes -suʔ and -ayu/-əm/-ano. The next section addresses the distribution of the remaining passive suffixes in Figure 2, -ɬ (EXPERIENCER) and -ʔas (LOCATIVE), whose function reflects semantic, rather than syntactic, properties of the argument.

3.2 Semantic Roles

Unlike the PRIMARY OBJECT and SECONDARY OBJECT passives -suʔ and -ayu/-əm/-ano, the passive morphemes described below are sensitive to the different semantic qualities of the argument they promote. The EXPERIENCER passive -ɬ is sensitive to the semantic quality of an event: it only applies to situations in which an event’s semantic transitivity is reduced due to lack of control (Hopper and Thompson 1980). On the other hand, the LOCATIVE passive -ʔas is sensitive to the semantic role of the promoted argument: it is only used to promote places to

---

32 Again in this example, we see the extraposition of a third-person pronominal primary object (‘it’) to the prepositional phrase laq.
subject position. We will see examples of both below. The analysis of these two passives draws much from Levine’s presentation in two 1980 articles concerning the Kʷaçʷala passive.

3.2.1 EXPERIENCER Passive -ɬ and Event Structure

Returning to Boas’ 1947 grammar, we find that he defines the suffix -ɬ as being the passive of ‘sensations’, ‘mental actions’, and ‘sensations produced by outside actions’ (Boas 1947:270). He gave examples such as the ones below:

\[(32) \text{Stems which passivize with -ɬ (Boas 1947:377)}\]

- a. \(\text{wəɬəɬ}-\) heard (CII 30.17)
  \(\text{wəɬ}-\) hear-PASS
- b. \(\text{wəɬáɬəɬəɬ}-\) discovered by hearing (III 257.3, X 5.13)
  \(\text{wəɬ}-\) hear-SENSE.SUDD-PASS
- c. \(\text{dúɬəɬəɬəɬ}-\) seen suddenly (be discover by seeing) (CII 98.13, CII 124.10)
  \(\text{duq}^\wedge-\alphaɬəɬ-\) see-SENSE.SUDD-PASS
- d. \(\text{mísəɬəɬəɬ}-\) discovered by smelling
  \(\text{mís}-\alphaɬəɬəɬ-\) smell-SENSE.SUDD-PASS
- e. \(\text{pəɬəɬəɬəɬ}-\) discovered by touch
  \(\text{pəɬəɬəɬ}-\) feel-SENSE.SUDD-PASS
- f. \(\text{xičəɬt}-\) witnessed
  \(\text{xičəɬ}-\) examine-PASS
- g. \(\text{laG}^\wedge-\) affected by fire (burnt by fire wood) (CX208.32)
  \(\text{laq}^\wedge-\) fire.wood-PASS
- h. \(\text{qəbəɬ}-\) affected by upsetting something on oneself
  \(\text{qəp}-\) hollow.vessel.upsidedown-PASS
- i. \(\text{pəɬəɬəɬəɬ}-\) affected by a blister (B1947:270)
  \(\text{pəɬəɬəɬ}-\) be.blistered-PASS

Some of these predicates are derived from stems that would not take objects and therefore may not have an active counterpart to contrast with their passive form: \(\text{laG}^\wedge-\) ‘firewood’ (distinct from the stems \(\text{mix}(a)-\) ‘to start a fire’ or \(\text{xiq}(a)-\) ‘to be on fire’), \(\text{qəp}-\) ‘a hollow vessel upside down’, and \(\text{pens}-\) ‘be blistered’. Levine presents additional elicited examples of passivized predicates without active counterparts. The stem \(\text{xəɬəɬ}-\) ‘be moldy’ is passivized with -ɬ as seen below.
Passive Constructions in Kʷakʷala

(33) \(xəd'ə-'be moldy' + EXPERIENCER\) passive -t

<table>
<thead>
<tr>
<th>PRED</th>
<th>SBJ</th>
<th>O2</th>
</tr>
</thead>
<tbody>
<tr>
<td>xəd'ʔida</td>
<td>kʷənikʷ (=sa xəd'oxa)</td>
<td></td>
</tr>
<tr>
<td>xəd'ʔ-i=da</td>
<td>kʷənikʷ (=sa xəd'ox-a)</td>
<td></td>
</tr>
<tr>
<td>mold-PASS=SBJ=DEF bread(=OBJ2) mold-T.DEM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The bread is moldy (with mold). (Levine 1980a:39)

The agent of the change of state, \(xəd'ox\) ‘mold’, is optionally included in the clause as an s-marked (secondary object) argument, parallel to the way demoted subjects are marked in other passive constructions. Levine provides many parallel examples with meanings such as “the man is sore-ridden (with sores),” formed with the stem \(ʔamt\)- ‘be sore-ridden’, or “the man was burned,” formed on the stem \(xiq\)- ‘be on fire’. Surprisingly, Levine’s consultants seem to find it acceptable to include the inanimate agent of the event (mold, sores, flame) as a secondary object marked argument — even though this information is already contained in the predicate.

As Levine notes, what these examples seem to share is not the quality of mental sensation, but the lack of control exerted over an event by the experiencer, whether it is the experience of being discovered or the experience of having something suddenly spill on one’s lap. This is expressed in some cases by the English verb ‘affected by’, which also contains the sense of a lack of control on the part of the experiencer.

Levine presents several elicited examples to illustrate the non-control aspect of the -t suffix. In the first sentence, no deliberate actor is present.

(34) Passive clause: \(Gəl-s\)- ‘paint’

<table>
<thead>
<tr>
<th>PRED</th>
<th>SBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gəld'ʔuχʷ=da</td>
<td>?uχʔiχ</td>
</tr>
<tr>
<td>Gəls-ʔ=da</td>
<td>?uχʔiχ</td>
</tr>
<tr>
<td>paint-PASS=DEM=DEF wall-DEM</td>
<td></td>
</tr>
</tbody>
</table>

The wall is overpainted. (Levine 1980b:5)

In example (35), Levine notes that the inclusion of a deliberate actor makes the sentence ungrammatical.

(35) Passive clause: \(Gəl-s\)- ‘paint’, secondary object not permitted

<table>
<thead>
<tr>
<th>PREP</th>
<th>SBJ</th>
<th>O2</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Gəld'ʔuχʷ=da</td>
<td>?uχʔiχ=isa bəgʷənʔəm</td>
<td></td>
</tr>
</tbody>
</table>

The wall was overpainted by a person. (Levine 1980b:5)

If an argument marked as a secondary object is included in the sentence, it must be the paint, as illustrated in (36).
Passive clause: Gəls- ‘paint’

The wall is overpainted with paint. (Levine 1980b:6)

The ungrammaticality of (35) and the role of the secondary-object argument in (36) suggest, as Levine argues, that the inclusion of an agent with control is semantically incompatible with the passive suffix -ɬ:

“semantic considerations … determine the appropriateness of co-occurrence between various elements in the surface form of sentences” (Levine 1980a:13).

The two clauses below illustrate the use of the EXPERIENCER passive in clauses with the stem duqʷ- ‘see’, first in an active sentence, and then in a clause passivized with -ɬ.

Active: duqʷ- ‘see’

The man saw the child. (Levine 1980b:39)

Passive: duqʷ- ‘see’

That is my seen one it is a good play. (i.e., That one seen by me is a good play. - DR; B1947:286)

In example (38) passivization of the stem duqʷ- may be triggered by the use of the demonstrative third-person pronoun predicate hi-; this automatic passivization following independent pronoun predicates is discussed below in section 4.2. (Boas’ awkward translation is an expression of his analysis that the complement phrase beginning with duqʷ- is a possessed nominal.) In my current research, duqʷ- is the only predicate I have found to combine with either -suʔ or -ɬ passive suffixes, indicating a difference in degree of control by the patient. However, initial findings suggest that the combination of the morphemes duqʷ- and -ɬ to mean ‘discover’ is not transparent to speakers, suggesting that it has lexicalized.

The functional constraints determining the distribution of this morpheme might be more specific than ‘lack of control’ or an ‘experiencer’ role for the subject. In documentation recorded in 2008, a speaker used -suʔ to passivize the example below, despite the subject’s lack of control.

---

33 Passive suffixes can co-occur with transitivizing suffixes such -gila or the causative -amas, such as in the stem hám-gil-ayu ‘be fed’ (eat-TR-PASS) (B1947 R225.46). Levine’s example laGʷ-ɬ-amas-suʔ=i=da (burn-PASS-CAUS-PASS-SUB-DEF) ędaqaya bgʷanam ‘The woman caused the man to be burned’ contains two passive suffixes (Levine 1980a:7). This seems to be a matter of semantic scope, but more research is clearly needed.
Passive Constructions in Kʷaḵwala

(39) Passive: ʰkilak- ‘attack’ with -su?

PRED O2
ʰkilakasuʔọʔasa ʰɬaʔiʔɬa
ʰkilak-suʔ=ọʔ=(a)sa ʰɬaʔiʔ-ɬa
attack-PASS=SBJ.DEM=OBJ2 bear-DEM
He is being killed (attacked) by the bear.
(Elicited: The bear is killing him; Shaw: 2008_7_17DS.340)

Interestingly, the subject of this clause is semantically an ‘experiencer’ of this event, but there is an animate being in control of the attack: the bear. One potential analysis is that the -ɬ morpheme is only used in situations in which an event is perpetrated by a non-animate entity: blisters, a container full of liquid, one of the five senses, fire or wind. Another possibility is that -ɬ is a Kʷaḵwala ‘middle voice’; part of the voice paradigm, but not a passive (Kemmer 1993). Of course, it may also be the case that language shift has led to further grammaticalization of the passive suffixes. The same speaker who provided example (39) above used the SECONDARY OBJECT passive -ayu for another context in which one might have expected -ɬ to have been used previously: the stem ʸəʔi- ‘wind’ is stative, and the event described is not subject to animate control.

(40) Passive: ʸəʔi ‘wind’ with -ayu

PRED
ʸəʔisdandaŋyuʔoɬaɬuɬuʔw
ʸəʔi-sdana-ayu=ʊ=ʊ=ɨɬu
wind-DIE.OF-MOM-PASS=DIST.PAST=3.SBJ=DEM
He was blown overboard by the wind (and died; Shaw:2008_7_17DS.340)

Further research will allow us to map the new distribution of passive morphosyntax in the contemporary language.

3.2.2 LOCATIVE Passive -ʔas and Semantic Role of Argument

Levine adds another suffix, the LOCATIVE -ʔas, to the list of passive morphemes provided by Boas. This suffix is included in Boas’ long glossary of lexical suffixes, but he only provides examples of this suffix used in the creation of place names. Levine gives examples, however, where it is also used to promote a place to subject from what would otherwise be an oblique constituent of an active sentence.

(41) Active: SUBJECT is ᵇəqʷanəm ‘man’

PRED SBJ OBL
la laʔuɬuʔw da ᵇəqʷanəmɬu ɬaʔwə ʔəwiʔaŋwəis
la-la-uɬuʔw=da ᵇəqʷanəm=ɬu ɬa=ɬa ʔəwiʔaŋwəis
RED-go-DEM=DEF man=DEM.OBJ1 PREP=DEM village
The man goes to this village. (Levine 1980a:243)

In example (41), ‘village’ is expressed within a prepositional phrase, beginning with ɬaʔwə. In the
example below, it becomes the subject of the sentence.

(42) Passive: SUBJECT is ?əwíŋagʷis ‘village’

PRED          SBJ       O₂
lalaʔasuxʷda   ?əwíŋagʷis sa  bəgʷánəm
lala-ʔas=uxʷ=da ?əwíŋagʷis=sa  bəgʷánəm
RED-go-PASS-DEM=DEF village=OBJ2 man

The village is where the man goes. (The village is gone to by the man; Levine 1980a:243)

Levine notes that -ʔas can only be used to promote locative arguments to subject status; non-
locative arguments (i.e., laxa bəgʷánəm ‘to the man’) can not be promoted with -ʔas.

Thus, semantic properties motivate the use of both LOCATIVE passive -ʔas and EXPERIENCER
passive -ɬ suffixes. In the case of the LOCATIVE passive -ʔas, the semantics of the constituent
determines the appropriateness of -ʔas. Locative phrases are always marked in a prepositional
phrase, but not all prepositional phrases are eligible for subjection via the -ʔas suffix, only
those referring to a place. On the other hand, in the case of the EXPERIENCER passive -ɬ, it is the
semantics of event structure that determine whether -ɬ is appropriate. Constituents promoted to
subjection may be primary objects of a predicate (as with duqʷ- ‘see’), or they may be the
single patient of an intransitive verb which doesn’t have a simple active counterpart (as with qəbəɬ-
‘affected by upsetting something on oneself’). Meanwhile, neither suffix is syntactically
constrained as the PRIMARY OBJECT and SECONDARY OBJECT passives are. Further research
comparing historical and contemporary discourse will help us determine whether the passive
paradigm has further grammaticalized, and perhaps contracted, as a result of language change
over the past century.

The preceding two sections identified the properties governing the distribution of Kʷaɬʷala
passive morphology. The next section compares the alignment of lexical and pronominal
arguments.

3.3 Grammatical Relations in Passive Constructions

This section discusses the effect of passivization on the grammatical relations of a clause,
comparing lexically expressed arguments with pronouns. In passive constructions with lexically
expressed arguments, the resulting argument structure is predictable. The distribution of
pronominal arguments is less predictable. This section describes the patterns of alignment for
both lexical and pronominal arguments.

As we have seen in many examples, the promotion of lexical argument to subject in a passive
clause results in two changes that indicate subject status: (1) prenominal subject inflection on the
predicate, and (2) the immediate post-predicate position of an argument in the syntax. In most
languages, demoted subjects of passivized predicates are marked as oblique, but as we have
repeatedly seen, a Kʷaɬʷala demoted subject takes SECONDARY OBJECT marking, as in (43).
Passive Constructions in K\textsuperscript{wa}k\textsuperscript{wa}la

(43) Passive: \textit{síxʷ} ‘paddle’ with \textit{-ayu}

\begin{itemize}
  \item DISC          PRED\textsubscript{1}, PRED\textsubscript{2}
  \item \textit{Wä! Ləʔám gax síwodayúsis}
  \item \textit{Wä! Ləʔám gax \textit{sixʷ-ayu}=Ø=s=ịs}
  \item Now! Then come paddle-lead-PASS=3.SBJ=OBJ\textsubscript{2}=3.POSS
  \item Wa! Then they came, they took him home
\end{itemize}

\begin{itemize}
  \item OBJ\textsubscript{2}
    \begin{itemize}
      \item \textit{nə̓gʷə̓mp} \textit{láwis} \textit{gókulot}
      \item \textit{nə̓gʷə̓mp} \textit{láw}=is \textit{gókulot}
      \item father.in.law \textit{and}=3.POSS tribe
      \item his father-in-law and his tribe.
    \end{itemize}
\end{itemize}

(Then he came paddled home by his father-in-law and his tribe. - DR; B1895 M679.17)

The father-in-law and tribe who paddle the subject home, \textit{nə̓gʷə̓mp láwis gókulot}, are marked as secondary objects with the prenominal enclitic \textit{=s} following the passive suffix. Meanwhile, the pronominal subject, the protagonist of this story, is marked with the third-person pronominal zero morpheme on the predicate, as we would expect.

The possessive markers in the example above also help us track referents and identify syntactic roles. K\textsuperscript{wa}k\textsuperscript{wa}la third-person possessors distinguish between subject and non-subject possessors, and \textit{=is} marks a subject possessor, as opposed to \textit{-a~=Ø} for the corresponding non-subject possessor. Thus, we know that \textit{=is} refers to the syntactic subject — the protagonist being paddled home — and not \textit{nə̓gʷə̓mp}, his father-in-law. (See Appendix III, Tables 8 and 9 for the full paradigm of third-person possessors.)

The enclitics above confirm that third-person arguments, whether expressed lexically or pronominally, behave predictably in passivized clauses. What about other pronominal arguments? Many languages treat speech act participants (first- and second-person) separately; the grammar might resist demoting a speaker or listener out of subject position. Example (44) shows us that second-person pronominal arguments do not resist demotion from subjecthood.

(44) Pronominal arguments in passive constructions: second person

\begin{itemize}
  \item PRED          COMP
    \begin{itemize}
      \item \textit{dúqʷə̓taχs} \textit{ʔəχʔιχsdəsuʔəaʔus}
      \item \textit{dúqʷ-ə̓ta-χs} \textit{ʔəχʔιχsd-suʔ-ƛ=us}
      \item see-CONT-2.POSS desire-PASS-FUT-3.SBJ>2.OBJ\textsubscript{2}
    \end{itemize}
  \end{itemize}

See what will be desired by you. (See your desired-by-you thing; B1947:255, CIII409.29)

The active version of this phrase might be something like ‘see what you will desire’, with a second-person pronominal subject \textit{=a} of the verb \textit{ʔəχʔιχsd- ‘desire’}.\footnote{In this sentence, a matrix imperative ‘see’ precedes a nominalized (and possessed) complement predicate, ‘desired-by-you thing’; the second person \textit{prenominal} possessive form \textit{-χs} is used (for possessive paradigms, see tables 8 and 9 in Appendix III) but unlike third-person, no distinction is made between subject and non-subject, so this does not give us additional information about syntactic status of the possessor.} In the passivized clause above, the second-person actor is demoted to secondary object status, reflected in the \textit{SECONDARY OBJECT} enclitic \textit{=us}, used for transitive verbs with third-person subjects acting on second-person secondary objects (paradigms of these pronominal referents are provided in 34).
Tables 6 and 7 of Appendix III).

As we might expect, when a speaker is promoted to the syntactic subject of a passivized sentence, the argument structure follows suit, as in (45).

(45) First-person promoted subjects in passive constructions: ləʔiλ- ‘enter’

PRED₁       PRED₂ (ADV)

ləʔiλ-imənɬəγən  Gʷʔiɬcik
ləʔiλ-suʔ=ənɬəγə=gən  Gʷʔiɬ-c=ik
enter-PASS=1s.SBJ>3.OBJ1=1.POSS  thus.in.house-inside-DEM

I was the object of entering (i.e., someone entered and came to me) when I was in my house here. (I was entered upon in my house. - DR; B1947:270)

The person who enters the house does not appear in the clause; the speaker is subject, and the primary object marking in the 1.SBJ>3.OBJ1 form =ənɬəγə refers to the nominalized (and possessed) phrase Gʷʔiɬcik ‘inside (of) my house’. Example (46), with SECONDARY OBJECT passive =əm promoting the first-person from secondary object status to subject status, 35 illustrates this as well.

(46) First-person promoted subjects in passive constructions: ləʔiλ- ‘enter’

PRED       OBL

ləʔiλ-əmən  la=γə  gukʷ
ləʔiλ-əm=ən  la=γə  gukʷ
enter-PASS=1s.SBJ  PREP=OBJ1  house

I am used for entering (i.e., I am taken) into the house. (I am brought into the house. - DR; B1947:270)

However, first-person arguments do indeed seem to resist demotion from syntactic subject status. See the series of examples below, produced in 2012. In 47a, the money retains its secondary object marking; in 47b, Pearl takes secondary object marking as the demoted subject.

(47) Avoiding demotion of first-person subject

a. Active həlaq(a)- ‘pay’: third-person subject

Həłəqasuuwí  Pərlasa  dala.

həłaq-suʔ=ι  Pərl=sa  dala
pay-PASS=SBJ  Pearl=OBJ2  money

Pearl was paid the money. (Rosenblum 2012jul23_BL_09)

35 The motion predicate ləʔiλ- ‘to enter’, like qas- ‘to walk’ and siɬʷ- ‘to paddle’, seems to have a ditransitive argument structure, with primary object marking for goals/destinations and, in this case secondary object marking for a person who is made to move in this way (‘brought into’).
b. Passive ḥəlaq(a)- ‘pay’: first-person promoted subject

\[
\begin{align*}
&\text{Hətəqasuʔən\, Hərl.} \\
&\text{hətəqa-suʔ=ən\, Pərl} \\
&\text{pay-PASS.1.SBJ>3.OBJ2 Pearl} \\
&\text{I was paid by Pearl. (Rosenblum 2012jul23_BL_09)}
\end{align*}
\]

Elicited: Pearl was paid by me.
Translation: I paid Pearl. (Rosenblum 2012jul23_BL_15)

In (47c), however, despite the requested passive sentence, the speaker offered the active clause. Speakers easily provided full paradigms of passive constructions with second- and third-person demoted subjects, but consistently offered active formulations when prompted to translate English passive forms with demoted first-person subjects (expressed as oblique in ‘by’-phrases in English). Alternatively, they used the Kʷačʷala first-person pronominal predicate with an active predicate, as in (48).

(48) Alternate construction

\[
\begin{align*}
&\text{núGʷəʔm \ qʷáxʔidamasga gíŋənanəm} \\
&\text{núGʷə=m \ qʷáx-(x)?id-amas-sga gíŋənanəm} \\
&\text{1.PRED \ 1.SBJ \ grow-MOM-CAUS-DEM.OBJ2 \ children} \\
&\text{Elicited: The children were raised by me.} \\
&\text{Translation: I am the one who raised the children.}
\end{align*}
\]

In the narratives published by Boas and Hunt, relevant examples are sparse, but (49) presents the possibility that first-person pronominals resist other non-subject syntactic slots as well. According to the translation provided by Boas and Hunt, the speaker should be marked as the primary object (the recipient) of the payment; instead, it is marked as the subject.
Based on the translation, the first predicate *halaq*—‘to pay’ refers to the payment made to the speaker by someone else; the second predicate is a phrase meaning, in effect ‘the person of whom we were ashamed’. Both are passivized with *-ayu* suffixes, which we know make subjects from secondary objects. The embedded predicate is framed by a purposive clause construction from the first person point of view (see Appendix III, Table 10), but within the embedded predicate, the alignment is as one would expect; the third-person subject is marked with *=Ø* (rather than primary-object *-q* or secondary *-s*). On the other hand, the grammar of the matrix predicate suggests a completely different interpretation than the translation offers: a demonstrative form signalling a third-person pronominal primary object *=oqʷ* precedes the first-person acting on third-person OBJ2 form *=ənχas*. (These forms are available in Appendix III Tables 6 and 7.) Having become familiar with the secundative alignment of the verb *halaq*—‘pay’, we know that if the recipient of payment (the speaker) were the intended subject, the verb would have been passivized with *-suʔ*, but this clause refers to the amount paid, and is hence passivized with *-ayu*. And yet, the use of *=ənχas* shows that the pronominal marking maintains subject status for the speaker. Meanwhile, the referent of primary object demonstrative *=oqʷ* is unclear.36

Levine’s work adds to the puzzle; he provides three examples of passives in which speaker retains subject status despite passivization, without commenting on the argument structure (the examples are provided to illustrate a different argument).

(50) First-person pronominals in passive constructions

```
PRED
naχʔidsunukʷχəntən
naq-x?id-suʔ-nukʷ-χənt=ən
drink-MOM-PASS-have-EVID-1SBJ
It must be that I had something to drink. (Levine 1980a:51)
```
(51) First-person pronominals in passive constructions

PRED
mxʔid-suʔ-iʔsd=ən
mx-xʔid-suʔ-iʔsd=ən
strike-MOM-PASS-DESID=1.SBJ
I want to get hit. (Levine 1980a:51)

(52) First-person pronominals in passive constructions

PRED
duxʷʔidsunukʷən
duqʷ-xʔid-suʔ-nukʷ=ən
see-MOM-PASS-have=1.SBJ
I went to look at something. (‘I’ve got something looked-at. - DR; Levine 1980a:52)

The translations of these clauses rely on lexical passives in English (‘I’ve got…’), or other ways of reducing agency (and hence, transitivity) on the part of the speaker; however, a passive interpretation would mean that the speaker is not the subject. All three of these constructions, unlike the example taken from Boas, contain aspectual suffixes following the passive derivation (-nukʷ ‘have’ and -iʔsd ‘DESID’); both act to reduce the telicity (and thus the transitivity) of the predicate (Hopper and Thompson 1980). This may also play a role in argument structure. Perhaps these are lexicalized stems for which second- and third-person subjects would also behave as if the predicate is active. Or it may be that the presence of valence-reducing suffixes such as -nukʷ and -iʔsd following a passive derivation trigger different argument structures. Further research will explore the relationship of aspectual suffixes to passive morphology and alignment of first-person arguments. The preliminary analysis suggests a different pattern of alignment for passives with first-person arguments than for other pronominal arguments, but additional data from connected and spontaneous speech are necessary to support the claim.

Prompted by the prevalence of inverse systems in neighboring languages, some might suggest that the divergence of first-person pronominal alignment is evidence of an inverse system (Forrest 1994, inter alia). Such inverse systems reflect the grammaticalization of some type of a semantic hierarchy ranking the expected topicality of participants in a clause: first- or second-person arguments may be more topical than third-person arguments, or human arguments over animal arguments, leading to special marking in clauses which reverse this hierarchy (with a horse kicking a man, for example).

Does the divergence of the first-person pronominal agreement in Kwak̓ala suggest an inverse system? I would argue that it does not, for several reasons. In a prototypical inverse system, such as the Plains Cree system described by Dahlstrom, neither the transitivity of the predicate nor the grammatical relations of arguments are changed: “both the inverse and direct form are transitive and active: that is, both map agent onto subject and patient onto object” (Dahlstrom 1986:74). But as we have seen in the examples above, passive derivations consistently reduce transitivity and reorganize argument structure in predictable ways in Kwak̓ala clauses. For example, in (53), we see again that the second-person argument (the agent of the predicate wəƛá- ‘ask’) is marked as a SECONDARY OBJECT with the suffix =us.
(53) Pronominal arguments in passive construction with -su?

PRED  SBJ
hímaʔis  wəxásəwus
hi- mâ-is  wəxâ-suʔ=us
3.DEM.PRON-OI-3.POSS.SBJ  ask-PASS=3.SBJ>2.OBJ2

That is what was asked by you. (B1947:286, III64.4)

Nakayama’s analysis of the “passive” suffix -’at from the neighboring Wakashan language Nuu-chah-nulth also explored the question of whether an inverse analysis might be appropriate. While he found that -’at is sensitive to an animacy hierarchy (based on speaker empathy, Nakayama 1986:429), he concludes that the pattern of pronominal indexing in -’at does not support an inverse analysis (Nakayama 1986:422). 37

K'wa:k'ala grammar does not seem to be sensitive to an animacy hierarchy, as we might remember from examples (12), (13), and (14), in which a dog bites a child. The active clause (12), with no special marking, is reproduced here.

(54) Active sentence: ʔəx- ‘bite’

AUX.PRED  PRED  SBJ
laʔəm  ʔəxʔídida  wəc=q
la- mâ=Ø  ʔəx-ʔíd=ı=da  wəc=(aq)
SEQ-OI=3.SBJ  bite-MOM=SBJ=DEF  dog=3.OBJ1

Now the dog bit him. (B1947:286, CIII12.19)

The third-person pronominal object =aq encliticized to the subject wəc- ‘dog’ refers to the person being bitten. Languages with an animacy hierarchy triggering inverse-marking might require ‘dog’ to be marked as ‘obviative’ and the third-person object to be marked as ‘proximate’, but there is no such marking here.

On the other hand, as we saw in examples (47) through (52), K'wa:k'ala alignment of passive constructions does seem to be sensitive to a hierarchy with respect to the speaker in relation to other participants (1 > 2,3). I would argue that these clauses should tentatively be interpreted as passives with a resistance to demotion of the first-person speaker from subject position. The behavior of pronominal marking in passive constructions in contemporary K'wa:k'ala discourse is a key target for further research.

The next and last part of this section addresses the combinatorial properties of these suffixes with respect to predicates and arguments, and the question of their status within the lexicon.

3.4 Derivation, Lexicalization, and Polysemy

As mentioned briefly in the introduction to the language, some derivational suffixes attach only to predicates, and others only to arguments; these suffixes contribute to the distinctions we can draw between syntactic predicates and arguments in the K'wa:k'ala clause. However, there is also ambiguity; some derivational suffixes can be used with either predicates or arguments. This is the case with four of the passive suffixes in K'wa:k'ala. The SECONDARY OBJECT suffixes -ayu and -əm — two of Boas’ so-called ‘instrumental’ passives — are used extensively to nominalize

37 Nor does Nakayama choose to characterize the Nuu-chah-nulth suffix -’at as “passive” (Nakayama 1986:429).
Passive Constructions in $K^wak^wala$

transitive stems, usually to create a word for the tool or instrument which performs an action. The LOCATIVE suffix -ʔas is used widely in placenames. On the other hand, the PRIMARY OBJECT suffix -suʔ and EXPERIENCER -t seem never to occur in a non-predicative context.

This raises the question of the status of such morphemes in the lexicon: are the nominalizers the same as the passivizing morphemes used to derive predicates? Boas considered these morphemes to be one and the same, serving different functions in a predicative context versus a nominalizing one. I would argue that they are not. Based on the lexicalization of instrumental forms in nouns, the grammaticalization of -suʔ and -ayu/-ʔm/-ano in relation to the syntax of argument structure, and the fact that not all of the passive forms have a nominalizing function, I consider the nominalizing forms of -ayu, -ʔm, -ano, and -ʔas to be separate lexemes, distinct from the homophonous passivizing suffixes to which they are historically related.

We have seen some of the nominalizations formed by attaching -ayu and -ʔm to transitive stems. Example (31) contains the word λáŋgayu ‘wedge’, a combination of the stem λámk- ‘to split wood with wedge’, and the suffix -ayu. The word Gályayu in example (34) combines the stem Gáls- ‘to paint’ with ‘-ayu’ to form ‘paint’ (the substance). The $K^wak^wala$ lexicon is full of such forms. The word ñítobayu ‘shoes’ derives from the reduplicated form of ñip- ‘to step on’ (because there are two shoes) and -ayu (Boas 1948:176). The word ñónáţi ‘rope’ combines the stem ñon- ‘to pull a rope’ and -ʔm (Boas 1948:151). The words sìwayu ‘paddle’, Káwayu ‘knife’, kيلة ‘fishing net’, and qámdám ‘song’ are further examples of nouns created with instrumental suffixes.

The suffixes are clearly very productive, but it is not likely that these forms are created anew each time. There is further evidence that certain combinations of forms are lexicalized. Boas contrasts qáñyu ‘thread’ with qánąyu ‘needle’, both from combinations of qáñ- ‘sew’ with -ayu (B1947:312). Another contrasting pair employs two different ‘instrumental’ suffixes with the same stem náq- ‘to drink’, leading to contrasting semantic values: -ayu in náGayu ‘drinking tube’ and -ʔm in náGám ‘bucket’. In still other cases, the productivity of these suffixes leads to formal variability with semantic constancy: both máñayu and máñyám ‘ruler, instrument for measuring’, derived from máñ- ‘measure’, are judged acceptable by Boas’ consultants, although he says that máñayu was more commonly used (B1947:302).

The different functional distributions of the argument and predicate suffixes further support a polysemous interpretation. As we have seen, the semantic role of a syntactic SECONDARY OBJECT is sometimes ‘instrument’, but not always. For ‘transfer’ ditransitives such as ‘give’, ‘pay’, or ‘send’, the SECONDARY OBJECT will be the transferred object (while the recipient is marked as PRIMARY OBJECT), and for ‘motion’ predicates such as ‘walk’ and ‘paddle’, the SECONDARY OBJECT will be the one made to move in that way (while the destination is marked as PRIMARY OBJECT). As we saw in example (31), -ʔm and -ano do not always promote an ‘instrument’ to subject status: the semantic role of λáŋgayu ‘wedge’ is the instrument which is taken (and used) in this clause, but the selection of the object-promoting passivizing suffix depends on the syntactic case of an argument. The predicate ?ṣəʔid- ‘take’ must be passivized with PRIMARY OBJECT passive -suʔ to allow promotion of λáŋgayu to subject status.

It is not yet obvious whether some stems used in a predicative context have lexicalized with

---

38 This is likely the origin of Boas’ use of the term ‘instrumental’ for the secondary object case, because he linked this suffix in its form as a nominalizer to its relationship with the secondary object paradigm of pronouns and demonstratives.

39 The two forms -ayu and -ýu are allomorphs of the same form: -ýu tends to follow vowels, /ɪ/, /ɬ/, /ɬ/, and /w/, but Boas notes “it (the passive morpheme -ayu) is somewhat irregular in its behavior” (B1947:312).
passive morphology. Levine gives the example *hanɨxʔidsonuŋʷən*, ‘I’ve got something to eat’, as an example of the suffix -suʔ in a lexicalized context.

\[(55) \text{PRED} \]
\[
\begin{align*}
\text{hanɨxʔidsonuŋʷən}^{40} \\
\text{ham-x?id-su?-nukʷ=ən} \\
\text{eat-MOM-PASS-have=1.SBJ}
\end{align*}
\]
\[I’ve \text{ got something to eat. (Levine 1980a:242)}\]

Like examples (50) to (52), this stem includes a valence-reducing suffix -nukʷ following the passive suffix -suʔ. Again, the first-person subject pronominal =ən is attached to the predicate, and the translation includes a first-person subject. It is not clear what function the passive suffix plays here. As mentioned earlier, perhaps this is a lexicalized derived stem which can now behave as an active predicate and take subject-like subjects. More research on the alignment of pronominal arguments in active and passive constructions will help answer this question definitively.

### 4 Discourse Motivations for Passivization

Section 3 focused on the morphosyntactic picture of passive constructions in Kʷaḵʷala: what the individual morphemes are, what their functions are, what syntactic and semantic considerations determine their distribution, and what alignments of arguments they trigger. But passive constructions are grammaticalizations of the patterns of use in a speech community: the tendency to promote a participant to subject arises in certain discourse contexts. Over time, these tendencies can become an obligatory and unconscious feature of the grammar. This section examines the discourse context in which passive constructions are used in Kʷaḵʷala. Both types of discourse motivations are explored in this section: (1) passivization as an optional strategy within a speaker’s repertoire, used in ways that reflect the speaker’s choice to focus the listener’s attention on something topical; and (2) passivization as an obligatory strategy triggered by certain syntactic patterns within discourse. The preliminary analysis of discourse data presented here draws on Boas’ texts and grammar alongside elicited material from Levine and data from recent fieldwork.

Figures 2 and 3 summarize the findings from seven interlinearized narratives recorded by Boas and Hunt (Boas 1895; Boas 1906).

---

40 This example is misspelled in Levine; it should be *ham- ‘eat’*. 
Passive constructions are far less frequent than active predicates; out of a total of 650 clauses (one predicate per clause), only 36 predicates, or 5.5%, were passivized. Kwakwala passive constructions can thus be considered pragmatically and functionally ‘marked’ in contrast to neutral active forms.

Perhaps because the PRIMARY OBJECT promotion suffix -suʔ applies to both transitive and ditransitive stems, it occurs more frequently in the texts examined than the three SECONDARY OBJECT promotion suffixes (-ayu, -ano, -əm) combined: 22 uses of PRIMARY OBJECT -suʔ compared with 12 uses of SECONDARY OBJECT -ayu (8), -ano (1) and -əm (3). The EXPERIENCER passive -ɬ was found in one predicate. No locative passives were found in the narrative texts. The

---

41 Clauses were counted according to predicates. Because some of the secondary object passive forms also occur in lexicalized nominals, passive forms were only counted when they were suffixed to a predicate stem and influenced the argument structure of the clause.
use of passives in the examined texts was primarily promotional, to allow a protagonist to maintain subject status, for example, or to maintain an important topic in subject position. These frequency counts should be considered in the light of the genre-limitation to origin tales. Forthcoming work analyzes the use of passive morphosyntax in a newly collected corpus of spontaneous interaction.

In many of these contexts, passivization is a narrative tool sensitive to speaker decisions about what is topical. Commonly-cited cross-linguistic motivations for using passive morphosyntax include topic continuity, argument structure of relativized clauses, and ‘foregrounding’ or ‘featuring’ of a significant participant (cf. Shibatani 1988 inter alia). These expectations are fulfilled in this sample of Kʷakʷala narrative discourse. However, passivization is obligatory in two syntactic contexts, described in the next two sub-sections: (1) relative clauses and (2) with independent pronoun predicates. These syntactic patterns are relatively rare in discourse, and much of the existing documentation relies on elicitation, which poses some challenges to interpretation, discussed below.

4.1 Passivization and Relative Clauses

In Kʷakʷala, relativized arguments must be presented as subjects of embedded predicates in relative clauses. The hierarchical constraint on accessibility to relativization, first described by Keenan and Comrie (1977), is common to many languages. In Kʷakʷala, when one wishes to relativize a non-subject argument of an active clause, one passivizes the embedded predicate to promote the argument to subject status and allow relativization. Levine’s elicited examples suggest that almost any argument from the matrix clause can be relativized if it can be made a subject (Levine 1980a:245-248). The syntactic status of the relativized argument determines whether a passive is needed and if so, which passive is used in the embedded predicate. In the first three examples, no passivization is needed because the relative is the subject of the active form of the embedded predicate.

(56) Active RC: Relative is OBJ1 of Main, SBJ of RC

\[\text{duq}^w\text{len}\lambda\chi\text{a}\quad \text{bog}^w\text{an}\text{om}\quad \text{nopida}\text{sa}\quad \text{tis}\text{om}.\]
\[\text{duq}^w\text{-l}=\text{en}(\chi)=\chi\text{a} \quad \text{bog}^w\text{an}\text{om}\quad \text{nopida}=\text{sa} \quad \text{tis}\text{om}.\]

see-PST=1.SBJ=OBJ1\quad \text{man}\quad \text{throw}=\text{OBJ2}\quad \text{rock}

I saw the man who threw the stone. (Levine 1980a:245)

(57) Active RC: SBJ of Main, SBJ of RC

\[\text{q}\text{olk}^{\w}\text{ka}\text{da}\quad \text{mex}^?\text{idi}\chi\]
\[\text{q}\text{olk}=\text{u}\chi^{w}=\text{da}\quad \text{mex}^?\text{id}=\text{i}\chi\]
\text{tired=OBJ1=DEF}\quad \text{strike=T.TDEM}

The one who hit is tired. (Levine 1980a:245)
Passive Constructions in Kʷakʷala

(58) Active RC: OBJ1 of Main, SBJ of transitive

duqʷələnλαχʷa       mexʔidiχ
duqʷ-l=ən(ƛ)=χʷa      mexʔid=iχ
see-PST=1.SBJ=OBJ1    strike=T.DEM
I saw the one who hit. (Levine 1980a:245)

In the next examples, the relativized arguments are not the subject of an active form of the predicate within the relative clause; passivization of the embedded predicate thus promotes the relativized argument to subject status and allows relativization. In (59), the primary object of the stem duqʷ- ‘see’, the one seen by the speaker, would be the primary object of an active form of mexʔid-. Thus, the PRIMARY OBJECT passive -suʔ is used to promote the relativized argument, ‘the one who got hit’, to become the subject of the relative clause and allow it to be relativized.

(59) Passivized RC: OBJ1 of Main, SBJ of passivized RC (< OBJ1)

duqʷələnλαχʷa       mexʔiawahidse
duqʷ-l=ən(ƛ)=χʷa      mexʔid-suʔ=iχ
see-EXP=1.SBJ=OBJ1    strike.with.fist-MOM-PASS=T.DEM
I saw the one who got hit (punched - DR). (Levine 1980a:245)

In the next example, both predicates are passivized. The subject of the main predicate, the speaker (indicated by pronominal =ənƛ) is the recipient of the gift, hence the use of PRIMARY OBJECT passive -suʔ to promote the recipient to subject of the matrix clause. As a result of passivizing the matrix predicate, the demoted actor (who gave the gift) is s-marked as a SECONDARY OBJECT with =sa. This actor, the gift-giver, is the relativized argument, ‘the one who was hit’.

(60) Passivized RC (and passivized main PRED):

OBJ2 of Main, SBJ of passivized RC (< OBJ1)

ˈcəswəwənλəsa       mexʔidsuʔ
ˈcəw-suʔ=ən(ƛ)=sa      mexʔid-suʔ
give-PASS=1.SBJ=OBJ2    strike.with.fist-MOM-PASS
I was given (it) by the one who was hit (punched - DR). (Levine 1980a:245)

He (or she) is the recipient of the punch, and would be marked as PRIMARY OBJECT in an active clause, but in order to allow relativization, PRIMARY OBJECT passive -suʔ promotes this person to subject. Interestingly, the gift itself, which would also be marked as a secondary object does not appear in the Kʷakʷala clause; there seems to be no space for it.

The example below shows the addition of further arguments to the relative clause; passivization within an embedded relative requires the same process of demotion of the actor (and secondary object marking) that one would expect in a main clause.
(61) SBJ of Main, SBJ of RC (<OBJ1)

\[
\begin{align*}
\text{hiʔəm} & \quad \text{bəgʷənəm} \quad \text{məxʔidswəsa} \quad \dot{c}ədaq \\
\text{hi-ɨm} & \quad \text{bəgʷənəm} \quad \text{məxʔid-suʔ=sa} \quad \dot{c}ədaq \\
3\text{PRON.PRED-OI} & \quad \text{man} \quad \text{strike-PASS=OBJ2} \quad \text{woman}
\end{align*}
\]

That is the man the woman hit (hit by the woman.) (Levine 1980a:245)

Note that the lexically-expressed relativized argument bəgʷənəm ‘man’ precedes the embedded predicate for which it is the subject, and no subject-marking appears on məxʔidswəsa. (With the lexical subject following the predicate, we would expect the prenominal subject marker =i). Note also that this is translated as a presentative clause, with ‘That is…’. The matrix predicate hiʔəm is an independent pronoun predicate (which for the third-person forms are also equivalent to demonstratives meaning ‘this’, ‘that (near)’, and ‘that (far)’ (Boas 1947:258); we have seen other examples of the pronominal predicates above. These so-called ‘verbal pronouns’ (Boas 1947:256) also trigger passivization and are further discussed in the next section.

Levine presents the next example as proof of the ungrammaticality of non-subjects in relative clauses. This seems very likely, but there are other ambiguities which may have obscured the grammaticality judgement of a consultant; the pronominal subject marker =i refers to the woman cədaq, but confusingly, the pronominal marker =χa is used in the constructed example to refer to the man (bəgʷənəm) who was hit, which precedes rather than follows the relativized predicate.

(62) * SBJ of Main, OBJ1 of RC

* \[
\begin{align*}
\text{hiʔəm} & \quad \text{bəgʷənəm} \quad \text{məxʔididaχa} \quad \dot{c}ədaq \\
\text{hi-ɨm} & \quad \text{bəgʷənəm} \quad \text{məxʔi=da}=\chi a \quad \dot{c}ədaq \\
3\text{PRON.PRED-OI} & \quad \text{man} \quad \text{strike=SBJ=DEF=OBJ1} \quad \text{woman}
\end{align*}
\]

That is the man the woman hit. (Levine 1980a:250)

It is not clear how the relativized argument should be expressed in the embedded predicate, and whether a different type of marking (such as the third-person PRIMARY OBJECT pronoun =q) or any marking at all would have been appropriate.

This example highlights the challenge of constructing examples for the elicitation of grammaticality judgements. Further research will pursue the question of relativization in both the corpus of spontaneous speech and in a context of carefully considered elicitation, and focus on the pronominal and prenominal marking of subjects in relative clauses.

4.2 Independent Pronoun Predicates

Another situation in which passivization seems to be obligatory involves the unusual Kʷaχʷala paradigm of demonstrative pronoun ‘predicates’ used in appositive constructions and presented in Table 3 below. These forms occur in clause-initial position and take predicate derivation and inflection as necessary although they do not take subject-marking enclitics. They must be followed by a complement which completes the clause and was analyzed by Boas as a type of relative clause; the complement takes special terminal marking triggered by the pronominal predicate.
Passive Constructions in Kʷakʷala

<table>
<thead>
<tr>
<th>1s</th>
<th>nūgoṿ</th>
<th>=Ø</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p.INC</td>
<td>nūgoṿənc</td>
<td>=Ø</td>
</tr>
<tr>
<td>1p.EXC</td>
<td>nūgoṿənuʔχ̣</td>
<td>=Ø</td>
</tr>
<tr>
<td>2</td>
<td>su</td>
<td>=əs</td>
</tr>
<tr>
<td>3.PROX</td>
<td>ga</td>
<td>=ək</td>
</tr>
<tr>
<td>3.MED</td>
<td>yu</td>
<td>=ux</td>
</tr>
<tr>
<td>3.DIST</td>
<td>hi</td>
<td>=i</td>
</tr>
</tbody>
</table>

Table 3: Predicate demonstrative pronouns and terminal markers (adapted from B1947:258)

When the complement is a simple NP, two constituents are juxtaposed, the predicate and the complement, as in the example below.

(63) Second-person demonstrative pronoun predicate

PRED    COMP
súʔəm    Síwidəs
sú=əm    Síwid=əs
2.PRON.PRED-OI Siwid=2.PR.DEM
You are (the one who is) Siwid. (B1947:258)

Sometimes, the complement is another predicate, and the complement predicate is stativized, as below with -nukw.

(64) Second-person demonstrative pronoun predicate

PRED    COMP
súʔəm    ʔísʔunukw
sú-əm    ʔísʔu-nukw
2.PRON.PRED-OI own.crest-having.NOM
You are the one who owns the privilege. (…who is the crest-owner - DR; B1947:258, CX 66.18)

However, in other cases, when the complement predicate is transitive, Boas observes that “when the subject is emphasized by a demonstrative pronoun, the predicate is expressed by a passive” (B1947:286). Boas’ ‘subject’ seems to refer to the subject of the complement predicate, and it seems that passivization ensures an interpretable relationship between the referent of the ‘pronominal predicate’ and the subject of the embedded predicate. In the first example below, passivization with -suʔ allows coreference between the demonstrative hiʔəm and the liked-thing to which it refers (which would normally be expressed as a primary object).
(65) Predicate demonstrative pronouns and passivization

<table>
<thead>
<tr>
<th>PRED.PRON</th>
<th>PRED.COMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>hīʔəm</td>
<td>?əχʔʔxsdəsəʔs</td>
</tr>
<tr>
<td>hí-ʔə-Ø</td>
<td>?əχʔʔxsdə-suʔ=əs</td>
</tr>
<tr>
<td>3.DEM.PRON.PRED-OI-3.POSS</td>
<td>like-PASS=2.PR.DEM</td>
</tr>
</tbody>
</table>

That is what he likes. (lit. that is his liked one; B1947:286)

In the next example, the construction is very similar. Here the first-person possessor -ən precedes the complement. Again, passivization permits coreference between the pronominal predicate híʔə- ‘that’ and the subject of the complement, [=en λúma ?əχʔʔxsdəsəʔwi] ‘my very-much liked thing.’

(66) Predicate demonstrative pronoun and passivized complement

<table>
<thead>
<tr>
<th>PRED</th>
<th>COMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>híʔən</td>
<td>λúma  ?əχʔʔxsdəsəʔwi</td>
</tr>
<tr>
<td>hí-ʔə-ən</td>
<td>λúma  ?əχʔʔxsdə-suʔ=i</td>
</tr>
<tr>
<td>3.DEM.PRON.PRED-OI-1.POSS</td>
<td>a lot  like-PASS=3PR.DEM</td>
</tr>
</tbody>
</table>

That is what I like very much. (That is my liked-thing. - DR; B1947:286)

(67) Predicate demonstrative pronoun and passivized complement

<table>
<thead>
<tr>
<th>PRED</th>
<th>COMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>híʔəʔis</td>
<td>wəƛásəwəs</td>
</tr>
<tr>
<td>hí-ʔə=is</td>
<td>wəƛá-suʔ=</td>
</tr>
<tr>
<td>3.DEM.PRON.PRED-OI=2.POSS</td>
<td>ask-PASS=3PR.3POSS</td>
</tr>
</tbody>
</table>

That is what was asked by you. (That is your asked-for-thing. - DR; B1947:286)

The prenominal possessive markers are familiar (see Appendix III), but the postnominal possessive markers =s seen in example (61) and (63) above and =əs in example (65) are not the same forms found in the standard possessive paradigm (see Appendix III, Tables 8 and 9), and are described as a special set of possessive markers used only in predicative pronominal constructions such as these (Boas 1947:259; see Appendix III, Table 11).

The three examples provided above are relatively straightforward to interpret; they also seem to be elicited examples, rather than spontaneously occurring ones. The next example, from one of Boas and Hunt’s texts, is less intuitively understood.

---

42 Despite the homophony between 1.SBJ =ən and 1.POSS =ən, a first-person subject would not make sense in this context.
(68) Second-person predicate demonstrative and passivized complement

<table>
<thead>
<tr>
<th>PRED</th>
<th>COMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>súʔəm</td>
<td>ma-ʔid-suʔ-gʷɨt=ɬ</td>
</tr>
<tr>
<td>súʔəm</td>
<td>ma-ʔid-suʔ-gʷɨt=ɬ</td>
</tr>
</tbody>
</table>

2.PRON.PRED-OI     punch-MOM-PASS-reason=3PR.3POSS
You are the reason of his being struck. (… the reason he was punched. - DR; B1947:258, CX 66.18)

In this example, the second-person demonstrative pronoun predicate súʔəm precedes a passivized complement meaning ‘reason for being punched’. The derivational suffix -gʷɨt ‘reason for’ follows the primary object passive suffix -suʔ; we don’t know the argument structure of an active form of ma-ʔid- ‘punch’ combined with -gʷɨt ‘reason for doing something,’ but based on our analysis of the distribution of -suʔ, the primary object of the active verb form would have to be ‘the reason’ for which someone was punched. Passivization seems to allow coreference between the listener and the promoted subject of the complement predicate, the reason for punching.

The preceding examples employ the PRIMARY OBJECT passive -suʔ, but other passives also occur in these appositive constructions. Example (38) is repeated below.

(69) Passive: duqʷ- ‘see’

<table>
<thead>
<tr>
<th>PRED</th>
<th>PRED_COMP</th>
<th>SBJ_COMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>hiʔmən</td>
<td>duqʷ-ʔida</td>
<td>ʔixpʰmatə</td>
</tr>
<tr>
<td>hi-m-ən</td>
<td>duqʷ-ʔ=ʔ=da</td>
<td>ʔixpʰmatə-a</td>
</tr>
</tbody>
</table>

3.PRON.PRED-OI-1.POSS see-PASS=SBJ=DEF good-play-T.DEM
That is my seen one it is a good play. (i.e., That one seen by me is a good play. - DR; B1947:286)

Here, the embedded predicate duqʷ- ‘see’ is passivized with the EXPERIENCER passive -t enabling the play ʔixpʰmatə to be the subject, coreferential with the third-person demonstrative pronoun ‘that’.

The use of passivization in special syntactic contexts is intriguing, but definitive claims are limited by the rarity of these constructions and the reliance on elicitation contexts to explore their occurrence. The development of searchable corpora for both legacy texts and newly collected spontaneous speech will hopefully further illuminate the use of passivization to serve the needs of speakers.

5 Discussion and Conclusions

The rich passive morphosyntax of Kʷaḵʷala adds to our typology of voice systems in the world’s languages. The Kʷaḵʷala passive paradigm reduces transitivity without reducing valence. Meanwhile, the distribution of passive morphology is shaped by both syntactic and semantic forces. The PRIMARY OBJECT passive suffix -suʔ and the SECONDARY OBJECT suffixes -ʔəu, -ʔəm and -ʔano are governed by syntactic criteria: the former used to promote χ-marked primary objects into subject position, the latter used to promote s-marked secondary objects into subject position. On the other hand, the functions of passive suffixes -t and -ʔas are determined by...
semantic considerations of different types. The use of the EXPERIENCER suffix -ɬ is determined by the event structure of the clause, expressing a lack of control on the part of the patient or an experience resulting from a non-animate agent. The LOCATIVE suffix -ʔas is only used to make subjects from a geographic place.

This description of the passive morphosyntax of Kʷaḵʷala reveals only the tip of the iceberg. There are several issues remaining to be explored in emerging corpora. The behavior of pronominal arguments, and in particular, the behavior of first person pronominal arguments in passivized predicates, invites further study to determine whether first person arguments indeed retain their subject status. The passivization of stative predicates and the question of what voice is expressed by EXPERIENCER -ɬ also invite further study. One might also ask whether predicates containing passive suffixes have lexicalized and can be used as the stems of active constructions.

The corpus of contemporary interactive speech in Kʷaḵʷala in progress will allow us to explore the use of passive suffixes in more recent speech. By comparing modern examples with those taken from Boas and Hunt’s texts recorded a century ago, we will be able to examine ways in which the usage or distribution of such morphemes may have changed, and flesh out our picture of passive morphosyntax in Kʷaḵʷala.

References


Passive Constructions in Kwakwa’ala


Daisy Rosenblum
University of California, Berkeley
Department of Linguistics
1203 Dwinelle Hall
Berkeley, CA 94720-2650
drosenblum@umail.ucsb.edu

Appendices

Appendix I: List of Abbreviations Used in Glossing

1 first person
2 second person
3 third person
A agent-like argument of canonical transitive verb
ADJ adjective
ADV adverb(ial)
AUX auxiliary
CAUS causative
CONN connective
CONT continuous aspect
DEF definite
Passive Constructions in KwaKwala

DEM demonstrative
DET determiner
DIST distal
DISTR distributive
EXCL exclusive
FUT future
IMP imperative
INCL inclusive
INS instrumental
INTR intransitive
LOC locative
N neuter
NEG negation, negative
NMLZ nominalizer/nominalization
OI old (known) information
OBJ1 primary-object
OBJ2 secondary object
P patient-like argument of canonical transitive verb
PR.POSS special possessive postnominal marking complements of predicative pronouns
PASS passive
PL plural
POSS possessive
PRED predicative
PROG progressive
PROX proximal/proximate
PST past
PURP purposive
Q question particle/marker
QUOT quotative
RECP reciprocal
REFL reflexive
REL relative
RM.PST remote past
RES resultative
S single argument of canonical intransitive verb
SBJ subject
SG singular
T.DEM sentence-closing demonstrative
TOP topic
TR transitive
Appendix II: Orthographies and Phonetic Correspondences

Note: All examples have been transliterated to the ‘NAPA’ (North American Phonetic Alphabet) orthography.

<table>
<thead>
<tr>
<th>NAPA</th>
<th>Boas</th>
<th>U’mista</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>p</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>p!</td>
<td>p!</td>
<td>p!</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>t!</td>
<td>t!</td>
<td>t!</td>
</tr>
<tr>
<td>g</td>
<td>g</td>
<td>g</td>
</tr>
<tr>
<td>g^w</td>
<td>gw</td>
<td>gw</td>
</tr>
<tr>
<td>k</td>
<td>k</td>
<td>k</td>
</tr>
<tr>
<td>k^w</td>
<td>kw</td>
<td>kw</td>
</tr>
<tr>
<td>k!</td>
<td>k!</td>
<td>k!</td>
</tr>
<tr>
<td>k^w</td>
<td>k!w</td>
<td>k!w</td>
</tr>
<tr>
<td>G</td>
<td>g</td>
<td>g</td>
</tr>
<tr>
<td>G^w</td>
<td>gw</td>
<td>gw</td>
</tr>
<tr>
<td>q</td>
<td>q</td>
<td>k</td>
</tr>
<tr>
<td>q^w</td>
<td>qw</td>
<td>kw</td>
</tr>
<tr>
<td>q!</td>
<td>q!</td>
<td>k!</td>
</tr>
<tr>
<td>q^w</td>
<td>q!w</td>
<td>k!w</td>
</tr>
<tr>
<td>λ</td>
<td>L</td>
<td>d†</td>
</tr>
<tr>
<td>λ†</td>
<td>L†</td>
<td>d†</td>
</tr>
<tr>
<td>dz</td>
<td>dz</td>
<td>dz</td>
</tr>
<tr>
<td>ts</td>
<td>ts</td>
<td>ts</td>
</tr>
<tr>
<td>ts!</td>
<td>ts!</td>
<td>ts!</td>
</tr>
<tr>
<td>s</td>
<td>s</td>
<td>s</td>
</tr>
<tr>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>x^w</td>
<td>xw</td>
<td>xw</td>
</tr>
<tr>
<td>x'</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>x^w</td>
<td>xw</td>
<td>xw</td>
</tr>
<tr>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td>m</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>m^m</td>
<td>‘m</td>
<td>‘m</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>n^m</td>
<td>‘n</td>
<td>‘n</td>
</tr>
<tr>
<td>l</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>l!</td>
<td>l!</td>
<td>‘l</td>
</tr>
<tr>
<td>y</td>
<td>y</td>
<td>y</td>
</tr>
<tr>
<td>y^m</td>
<td>‘y</td>
<td>‘y</td>
</tr>
<tr>
<td>w</td>
<td>w</td>
<td>w</td>
</tr>
</tbody>
</table>
Appendix III: Tables of Pronouns and Demonstratives

Table 4: 3rd person demonstrative verbal enclitics and postnominals (adapted from Boas 1947:252)

<table>
<thead>
<tr>
<th>3.DEM</th>
<th>Attached to Predicate</th>
<th>Postnominal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pronominal</td>
<td>Prenominal</td>
</tr>
<tr>
<td>PROX.VIS $^{43}$</td>
<td>=k</td>
<td>=ga(da)</td>
</tr>
<tr>
<td>PROX.INV</td>
<td>=gaʔ</td>
<td>=ga(da)</td>
</tr>
<tr>
<td>MED.VIS</td>
<td>=uχ</td>
<td>=uχ(da)</td>
</tr>
<tr>
<td>MED.INV</td>
<td>=uʔ</td>
<td>=uχ(da)</td>
</tr>
<tr>
<td>DIST.VIS</td>
<td>=iq</td>
<td>=i(da)</td>
</tr>
<tr>
<td>DIST.INV</td>
<td>-iʔ</td>
<td>=i(da)</td>
</tr>
</tbody>
</table>

Table 5: 3rd person pronominal and adnominal demonstrative enclitics (adapted from Boas 1947:252)

<table>
<thead>
<tr>
<th>3.DEM</th>
<th>Pronominal</th>
<th>Prenominal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SBJ</td>
<td>OBJ1</td>
</tr>
<tr>
<td>PROX.VIS</td>
<td>=k</td>
<td>=qʔk</td>
</tr>
<tr>
<td>PROX.INV</td>
<td>=gaʔ</td>
<td>=χgaʔ</td>
</tr>
<tr>
<td>MED.VIS</td>
<td>=uχ</td>
<td>=qʷ</td>
</tr>
<tr>
<td>MED.INV</td>
<td>=uʔ</td>
<td>=qʷ, =quʔ</td>
</tr>
<tr>
<td>DIST.VIS</td>
<td>=iq</td>
<td>=q</td>
</tr>
<tr>
<td>DIST.INV</td>
<td>-iʔ</td>
<td>=qı</td>
</tr>
</tbody>
</table>

$^{43}$ Boas named these demonstratives according to their proximity to speech participants, as ‘Demonstrative of 1st person, visible, Demonstrative of 2nd person, visible, etc.’ It is not clear whether this reflects additional referential qualities other than proximity, such as discourse relevance. However, the labels Proximal, Medial and Distal are more transparent and one can infer some metaphoric or deictic extension.
Table 6: Transitive predicates with primary object

| Subject | Primary Object | | | | |
|---|---|---|---|---|
| 1SG | 1INCL | 1EXCL | 2ND | 3RD |
| --- | --- | --- | =ənƛoƛ | =ənƛaƛ |
| --- | --- | --- | --- | =ənc aʔ |
| 2ND | =ənƛoƛ | =ənƛoƛ | =ənuʔƛʷaʔ |
| --- | --- | --- | --- | =siq |
| 3RD | =ənƛoƛ | =ənƛoƛ | =ənuʔƛʷaʔ | =q |

Table 7: Transitive predicates with secondary object (adapted from Boas 1947:253)

| Subject | Secondary Object | | | | |
|---|---|---|---|---|
| 1SG | 1INCL | 1EXCL | 2ND | 3RD |
| --- | --- | --- | =ənƛos | =ənƛas |
| --- | --- | --- | --- | =əncas |
| 2ND | =secən | =secənʔƛʷ | =secənʔƛʷ | =sis |
| 3RD | =ən | =ənc | =ənuʔƛʷ | =us | =s |

Table 8: Possessive enclitics for 1st and 2nd person (adapted from Boas 1947:253)

<table>
<thead>
<tr>
<th>3.DEM</th>
<th>Prenominal</th>
<th>Postnominal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1SG</td>
<td>1INCL</td>
</tr>
<tr>
<td>PROX.VIS</td>
<td>=gin</td>
<td>=ginc</td>
</tr>
<tr>
<td>PROX.INV</td>
<td>=g=</td>
<td>=ga=</td>
</tr>
<tr>
<td>MED.VIS</td>
<td>=ən</td>
<td>=ənc</td>
</tr>
<tr>
<td>MED.INV</td>
<td>=q=</td>
<td>=q=</td>
</tr>
<tr>
<td>DIST.VIS</td>
<td>=us</td>
<td></td>
</tr>
<tr>
<td>DIST.INV</td>
<td>=is</td>
<td></td>
</tr>
<tr>
<td>=a=</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9: Possessive enclitics for 3rd person (adapted from Boas 1947:254)

<table>
<thead>
<tr>
<th>3.DEM</th>
<th>Possessor subject of sentence</th>
<th>Possessor not subject of sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prenominal</td>
<td>Postnominal</td>
</tr>
<tr>
<td>PROX.VIS</td>
<td>=gas</td>
<td>=k</td>
</tr>
<tr>
<td>PROX.INV</td>
<td>=gaʔ</td>
<td>=gaʔs</td>
</tr>
<tr>
<td>MED.VIS</td>
<td>=q (=iƛ)</td>
<td>=q (=aƛ)</td>
</tr>
<tr>
<td>MED.INV</td>
<td>=q (=aƛ)</td>
<td>=q (=aƛ)</td>
</tr>
<tr>
<td>DIST.VIS</td>
<td>=i</td>
<td>=as</td>
</tr>
<tr>
<td>DIST.INV</td>
<td>=a</td>
<td>=as</td>
</tr>
</tbody>
</table>
Passive Constructions in *Kʷakʷala*

<table>
<thead>
<tr>
<th>1SG</th>
<th>qən…aʔən</th>
</tr>
</thead>
<tbody>
<tr>
<td>1INCL</td>
<td>qənc…aʔnc</td>
</tr>
<tr>
<td>1EXCL</td>
<td>qənuʔχʷ…(aʔnʔχʷ)</td>
</tr>
<tr>
<td>2ND</td>
<td>qaʔs…aʔus</td>
</tr>
<tr>
<td>3RD (POSSESSOR ≠ SUBJECT)</td>
<td>qaʔ…is</td>
</tr>
<tr>
<td>3RD (POSSESSOR = SUBJECT)</td>
<td>qaʔs…a</td>
</tr>
</tbody>
</table>

Table 10: Purposive clauses (adapted from Boas 1947:274)

<table>
<thead>
<tr>
<th>1.POSS</th>
<th>2.POSS</th>
<th>3.POSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>---</td>
<td>nugʷʔəms …=us</td>
</tr>
<tr>
<td>2</td>
<td>súʔmən …=s</td>
<td>---</td>
</tr>
<tr>
<td>3</td>
<td>híʔmən …=Ø</td>
<td>híʔəms …=Ø</td>
</tr>
</tbody>
</table>

Table 11: Terminal markers on possessed nominals occurring with pronominal predicates (adapted from Boas 1947:259)
Dialect Contact, Convergence, and Maintenance in Oregon Athabaskan

JUSTIN SPENCE
University of California, Davis

1 Introduction

The Oregon Athabaskan languages were once spoken in an area stretching from the upper reaches of the Umpqua River in present-day Douglas County in Oregon to Del Norte County in northern California. Tensions fueled by an influx of Euro-American settlers from the 1840s onwards culminated in the Rogue River Indian Wars of 1855-1856 (Miller and Seaburg 1990, Schwartz 1997). Ostensibly for their own protection, most indigenous people of the region were dispossessed and consolidated on the Siletz and Grand Ronde reservations in the northwest part of the state, far from their ancestral homelands. Traditional food sources were scarce, and chronic undernutrition and disease took a heavy toll (Kent 1973, Bureau of Indian Affairs 1979). Those who managed to survive found themselves thrust elbow-to-elbow in a small number of reservation communities organized according to coarse-grained ethnolinguistic similarities.

As elsewhere in North America, the long-term linguistic consequences of this forced resettlement are by now all too familiar. Population loss, reservation life, and adaptation to a wage economy created an imperative for indigenous people to communicate in new ways, both with each other and with agents of the foreign culture that had uprooted them (Zenk 1988). Regional contact languages – initially Chinook Wawa (also known as Chinook Jargon), eventually English – gained new prominence in everyday life and gradually supplanted traditional forms of speech. By the mid-twentieth century, all of the Oregon Athabaskan languages were highly endangered; with the marginal exception of Tolowa, as far as is known none has survived to the present day (Pierce and Ryherd 1964, Golla 2011).

While the sociolinguistic setting leading to the obsolescence of Oregon Athabaskan is thus readily apparent, less well-understood are the system-internal consequences of reservation life on the Athabaskan languages themselves. Language shift did not happen overnight: most Oregon Athabaskan languages were still being acquired by children for the first half-century or more after resettlement, and at least three generations of speakers were multilingual to varying degrees in Athabaskan languages, Chinook Wawa, and English. There was thus ample opportunity for contact-induced linguistic changes to take place, including those due to the mutual influence of different varieties of Oregon Athabaskan on one another. Indeed, it would be surprising if such changes did not occur, since the overall demographic profile of resettlement on reservations – abrupt relocation of closely related speech varieties to a relatively restricted geographic area – is broadly similar to better-studied cases of dialect contact, such as large-scale migration to cities and colonies that led to the emergence of mixed varieties of English, Norwegian, and Hindi (Kerswill 1994, 2002; Kerswill and Williams 2000; Trudgill 1986, 2004; Kerswill and Trudgill 2005).

A common outcome of dialect contact is koineization, which can be understood as the gradual convergent leveling of dialect differences over successive generations of speakers raised in close proximity to each other. Apparently this was the result of the resettlement of Athabaskan-speaking people at Siletz as well. Reporting on his work there in the mid-1960s, Golla (1976:218) noted that most of the differences between the Lower Rogue River dialects of Oregon
Athabaskan, by then collectively known as “Tututni,” had all but disappeared: “The phonology and grammar of Mrs. Bensell’s [Euchre Creek] dialect appear to hold for all Tututni. The only major differences among the dialects that survived at Siletz at the time of my work were lexical.”¹ However, most details of what these differences might have been remain unexplored. As Golla (2011:70) has pointed out, much of what is known about Oregon Athabaskan is based on interviews conducted in the mid-twentieth century with speakers who had spent most of their lives in the reservation setting, presumably after any contact-induced changes had already run their course. It is only by examining documentation from earlier periods, closer to the time when resettlement occurred, that erstwhile differences between varieties can be detected and the effects of linguistic contact revealed.

These observations lead to the following questions of interest: which similarities between closely-related varieties of Oregon Athabaskan found in twentieth century documentation are extensions of a pre-reservation status quo, and which are the result of processes set in motion by intensive contact after resettlement? Just as importantly, which differences between varieties were maintained despite the leveling that took place in the same period? Working with documentation of the Oregon Athabaskan languages from the mid-nineteenth century onwards, the present study is a preliminary effort to answer these questions. Following an overview of the issues in the dialect contact literature that the data relate to (§2) and a summary of the languages (§3), two cases where leveling apparently did result from contact in the reservation period are considered: reduction of variation involving reflexes of Proto-Athabaskan *q and *χ̣ in Tututni (§4.1), and loss of several distinctive properties of the Chetco dialect (§4.2). Despite convergent pressures, however, some dialect idiosyncrasies survived, the Galice Creek dialect in particular maintaining several distinctive features into the twentieth century (§5.1). Even closely related dialects formerly spoken on the Rogue River avoided complete erasure of differences, Mikwantu-ni apparently maintaining a tendency to affricate onset coronal fricatives (§5.2).

Beyond its contribution to the understanding of the post-contact development of the Oregon Athabaskan languages in particular, this study sheds light on the general mechanisms by which dialects and closely related languages influence each other when they come into contact. Section §6 addresses the question of the extent to which koineization proceeds essentially deterministically, with majority variants displacing minority ones, or is instead mediated by non-linguistic social factors. Trudgill (2004) has argued for the former position in the special circumstances leading to similarities between colonial varieties of English that emerged in the southern hemisphere. Here it is argued that the observed outcomes of contact between Oregon Athabaskan varieties cannot be explained with reference only to the relative proportions of speakers of different varieties, contrary to the predictions of the deterministic theory set forth by Trudgill.

2 Theoretical Background

A well-established theory of koineization, developed over the course of many years of research by Trudgill, Kerswill, and others, is shown here (see especially Trudgill 2004, Kerswill and Trudgill 2005):

¹ Golla’s report matches that of Pierce and Ryherd (1964:142): “The amount of leveling that has taken place since these tribes were uprooted in 1856 and transported to reservations at Siletz and Grand Ronde, Oregon might tend to lead us to assume that the early nineteenth century relationships between these dialects and languages were closer than they actually were.”
Stage I  Dialect divisions are largely maintained, the adult founders of new communities having relatively little linguistic influence on one another.

Stage II  Selection of variants gets underway: extreme inter- and intra-speaker variation as children are exposed to a range of linguistic input in the home and in the community at large.

Stage III  Variation is vastly reduced as the speech community settles on a relatively stable set of linguistic norms.

According to this theory, the formation of koineized dialects involves three distinct stages, corresponding to successive generations of speakers in a mixed dialect community. In the first stage, adult speakers of diverse dialects come into contact with each other, but their speech undergoes only rudimentary leveling due to the fact that linguistic behaviors already established by adulthood are difficult (although by no means impossible) to change. In the second stage, the first generation of speakers born in the new community starts the process of selecting from the linguistic variables in their environment as speakers accommodate one another in their everyday interactions. This stage is marked by extreme variation across the community as a whole and in the speech of individual speakers. Finally, the third generation settles on a stable set of linguistic norms, typically as social cohesion among speakers develops – so-called “focusing” in the sense of Le Page and Tabouret-Keller (1985). At this point a koineized dialect may have been formed, assuming these norms are not identical to any of the original dialects that came into contact.

According to Kerswill (2002), koineization can be impeded by various factors, especially a lack of social integration in the new community – that is, a lack of focusing can mitigate convergent pressure. However, the expectation of the three-stage model is that koineization will result where focusing takes place. This is stated most clearly in Trudgill (2004:88-89), who maintains that “focusing implies koineization.” Trudgill takes this a step further in his book subtitled “the inevitability of colonial Englishes,” claiming that not only is the fact of koineization inevitable, but so too are its outcomes, in some cases at least. Considering what he calls “tabula rasa” dialect contact, in which speakers of diverse dialects migrate en masse to a new geographic location where none has pre-established claims to legitimacy, he maintains that koineization proceeds essentially deterministically:

“...given sufficient linguistic information about the dialects which contribute to a mixture, and given sufficient demographic information about the proportions of speakers of different dialects, it is possible, within certain limitations, to make predictions about what the outcomes of the mixture will be, at least in broad outline.”
(Trudgill 2004:26)

In Trudgill’s view, socially contingent factors such as “prestige,” “stigma,” “identity,” and “ideology” play little role in the outcomes of koineization in tabula rasa situations, except perhaps in the initial dialect mixing among Stage I adult speakers who first come into contact with each other.

The data presented in this paper bear crucially on these issues insofar as the resettlement of Oregon Athabaskan people has a sociolinguistic profile similar to Trudgill’s tabula rasa dialect contact. Unlike some situations in North America where reservations were established on a
particular group’s aboriginal territory, the Siletz reservation was 130 miles from the nearest Athabaskan homeland, and no single Athabaskan group had a pre-existing claim to legitimacy in the locus of contact. It is argued in §3 that the new reservation community did indeed undergo a degree of focusing in this period, leading to the expectation in the framework developed by Kerswill and Trudgill that koineization should have occurred.\(^2\) Evidence for leveling presented in §4 tends to confirm this prediction; nonetheless, cases of dialect maintenance presented in §5 show that not all differences between varieties were erased by the mid-twentieth century. Moreover, despite the sociolinguistic similarity with the tabula rasa dialect contact considered in Trudgill (2004), §6 develops an argument that the outcomes of dialect contact are not what Trudgill’s theory would predict: there is both too much and not enough retention of the idiosyncratic features of the pre-contact Oregon Athabaskan varieties, and the outcomes do not appear to be based straightforwardly on the relative proportions of speakers of each variety.

3 Overview of Oregon Athabaskan

3.1 Classification

Several varieties of Oregon Athabaskan have been identified in the linguistic literature, most associated with pre-contact villages located along watercourses in the region (Miller and Seaburg 1990). Figure 1 summarizes the classification of Golla (2011:70-75); abbreviations used when citing data are given in square brackets.

<table>
<thead>
<tr>
<th>Upper Umpqua</th>
<th>Rogue River dialect network</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upper Coquille</td>
<td>[Coq]</td>
</tr>
<tr>
<td></td>
<td>Lower Rogue</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sixes River</td>
<td>[Six]</td>
</tr>
<tr>
<td></td>
<td>Euchre Creek</td>
<td>[Euc]</td>
</tr>
<tr>
<td></td>
<td>Tututni-Joshua</td>
<td>[Tut]</td>
</tr>
<tr>
<td></td>
<td>Mikwanutni</td>
<td>[Mik]</td>
</tr>
<tr>
<td></td>
<td>Chasta Costa</td>
<td>[ChC]</td>
</tr>
<tr>
<td></td>
<td>Pistol River</td>
<td>[PiR]</td>
</tr>
<tr>
<td></td>
<td>Galice-Applegate</td>
<td>[Gal]</td>
</tr>
<tr>
<td>Chetco-Tolowa</td>
<td>Chetco</td>
<td>[Cht]</td>
</tr>
<tr>
<td></td>
<td>Tolowa</td>
<td>[Tol]</td>
</tr>
</tbody>
</table>

Figure 1: Classification of Oregon Athabaskan languages based on Golla (2011)

Figure 1 is arranged roughly north to south, with Upper Umpqua the northernmost member of the group (in the region around present-day Roseburg) and Chetco-Tolowa the southernmost

\(^2\) This is not to suggest that the languages had been previously isolated from each other, as there is ample historical and ethnographic evidence that the Athabaskan groups of Oregon had deep-rooted connections with each other and surrounding groups well before contact with Euro-American society (cf. Miller and Seaburg 1990, Hall 1992). Rather, the social and geographic dislocation in this period changed the nature and frequency of interactions.
Dialect Contact in Oregon Athabaskan

(straddling the Oregon-California border). Readers are referred to Golla (2011), Drucker (1937), and Dorsey (1890) for detailed discussion and maps of the various groups and their aboriginal territories.

According to Golla (2011), the varieties in the outermost layer in Figure 1 are distinct languages (Upper Umpqua, Rogue River, and Chetco-Tolowa, indicated in bold); the Rogue River group is a dialect network with several local varieties, some more nearly similar to one another than others. It should be pointed out, however, that both Pierce and Ryherd (1964:142) and Bright (1964:101) reported a high degree of mutual intelligibility between Chetco-Tolowa and the Rogue River dialects. Golla (2011) includes Galice-Applegate (“Galice” henceforth) as a member of the Rogue River dialect network but considers it highly distinctive within the group. Sapir (1914:274) reported that Galice “differed so much from Chasta Costa [in the Lower Rogue group] as to be but partly understood, if at all, by speakers of the latter,” and Drucker (1937:283) considered Galice to be “the most divergent of all the groups,” including Chetco-Tolowa. Krauss (1973:924) considered Galice “an isolated (but still not too distant) extreme”. These reports suggest that the Oregon Athabaskan languages were, at most, members of a shallowly-differentiated family with many speakers passively competent in other varieties; they might even have composed a single dialect continuum with a high degree of mutual intelligibility between adjacent varieties, as suggested by Krauss (1973:924).

3.2 Documentation and Orthographies

Data presented below are drawn from a number of published and unpublished sources. The earliest documentation of Tututni is found in wordlists collected by Hubbard (published in Taylor 1860) and Kautz (1855) in the mid-1850s, around the time when Oregon Athabaskan people were being forcibly removed to Siletz and Grand Ronde. The next major round of documentation was approximately one generation later in the 1880s, first by Everette (1882), and especially by Dorsey (1884a-n), who collected wordlists and paradigms for most of the known varieties of Oregon Athabaskan. Twentieth century documentation consulted in this study includes the published materials of Sapir (1914), Drucker (1937), the microfilm edition of Harrington’s (1942/1981) field notes, Bright (1964), E. Jacobs (1968, 1977), M. Jacobs (1968), Hoijer (1960, 1966, 1973), Golla (1976, 2011), Landar (1977), and Bommelyn (1989); unpublished twentieth century sources include fieldnotes collected by Barnett (1934) and Drucker

3 As noted in §1, by the mid-twentieth century “Tututni” had come to designate all of the Lower Rogue River dialects rather than what was originally a dialect spoken near the mouth of the Rogue River (Golla 1976, 2011). In this paper I will generally use the term in the narrower sense, especially in discussions of the nineteenth century documentation. Also, while some of the varieties in the Rogue River group were not spoken on the Rogue River per se, the term is convenient to designate the geographic center of the dialect group.

4 Statements from speakers interviewed by J.P. Harrington also show that the languages were mutually intelligible to some extent, one stating that Chetco people “talk half-way RR [Rogue River]” (Harrington 1942/1981, reel 25, frame 989). Barnett (1934) reported Billy Metcalf saying that the Chetco and Tolowa people talk the same and “understand but talk different” from the Rogue River dialects. See also the statement in §5.1 from speaker Hoxie Simmons on the status of Galice. These questions of mutual intelligibility matter insofar as it is not clear whether all of the Oregon Athabaskan varieties considered here qualify as “dialects” in a sense relevant to understanding koinéization. The evidence presented here suggests that they do satisfy a definition of dialect contact encompassing “contact between varieties of language that are mutually intelligible at least to some degree” (Trudgill 1986:1).
JUSTIN SPENCE

(1934), and a draft lexicon of Tututni prepared by Golla (2008). Notice that this timeline of documentation, spaced approximately one generation apart (1850s – 1880s – 1914 and forward), fits rather conveniently into the 3-stage model of koineization outlined in §2.

Given such a wide variety of documentation collected over so long a time span, data presented in this paper have been normalized for ease of comparison. Normalization has often involved simple transliteration, but no attempt has been made to correct transcriptions where a given researcher might not have recorded certain distinctions consistently. In many cases representations of the original transcriptions are also provided in angle brackets, especially for nineteenth century sources, although these are sometimes omitted where they interfere with legibility in tables. Normalized transcriptions for the most part use conventions similar to those found in Golla (1976, 2008), but with [ṣ] instead of the digraph [sʳ] for retroflex [ʂ], and [ts] rather than [c] for the voiceless coronal affricate; following standard practice in the Athabaskanist literature, aspirated stops are represented with [p t k q] (so [k] is IPA [kʰ]). In many cases it is unclear whether the parameters of variation under discussion are phonemic versus sub-phonemic. Square brackets are generally used both for particular (normalized) transcriptions, but this should not be considered a commitment to particular parameters being phonetic rather than categorical. Reconstructed Proto-Athabaskan vocabulary is taken from Krauss (2005) except where otherwise noted.

3.3 Resettlement and Focusing

As outlined in §1, the Rogue River Indian War of 1855-1856 and subsequent resettlement on the Siletz and Grand Ronde reservations were watershed events in the history of contact between indigenous people of southwestern Oregon and Euro-American colonizers (Schwartz 1997). Most Oregon Athabaskan groups wound up at Siletz and were ultimately consolidated in a small number of communities that were partially segregated from the Takelma, Shasta, and other indigenous groups living on the reservation.

Resettlement had far-ranging social consequences, among which was a partial erasure of distinctive pre-contact micro-regional identities. For example, Dorsey (1889, 1890) observed that most village groups resettled at Siletz still maintained separate burial grounds, but noted that all of the Chetco-speaking people had come to share just one, even though they had once lived in nine villages distinct enough that they could intermarry (village exogamy having been the norm in pre-contact times). The intermingling of populations is also reflected in the following state-
ment from the report of the Office of Indian Affairs’ agent at Siletz in 1887: “It would be impos-
sible to give the exact number of each tribe, on account of so much intermarrying among them” (Office of Indian Affairs 1887:189). This accords with the perception of people who were raised at Siletz in this period: Drucker (1934) reports one woman saying that “people down on Lower Farm [at Siletz] were all mixed up – Chetco, Rogue R., Euchre Cr., Sixes Cr., etc.” At the same time, however, at the time of Dorsey’s visit he found no common ethnonym in use for the Athabaskans with origins along the Rogue River (Dorsey 1890). The enduring importance of people’s ancestral villages is evident moreover in the fact that people interviewed by Barnett (1934) and Drucker (1934) in the twentieth century knew with a high degree of precision the names (if not the exact locations) of their parents’ and grandparents’ former village affiliations, even when in some cases they had never visited these places themselves (Drucker 1937:269, fn. 24).

There is very little published information about the status of Athabaskan and other indig
eous languages spoken at Siletz and Grand Ronde for nearly thirty years following resettlement, except indirectly in the annual reports of the U.S. government’s agents. The earliest of these indicate increased use of contact languages – English to some extent, but especially Chinook Wawa. J.H. Huffer, the schoolteacher at Grand Ronde in 1864, complained, “I have found considerable difficulty in getting the scholars to speak the English language properly or at all, but I am now beginning to overcome that difficulty by giving premiums to those that speak the most and the best English at the end of each month” (Office of Indian Affairs 1865:94). Two years later J.B. Clark, another teacher at Grand Ronde, reported, “I find no difficulty in making them understand the English language, and I endeavor to make them communicate their ideas in the same language, but they will use that barbarous jargon, the Chenook” (Office of Indian Affairs 1866:82). This last statement shows clearly that only eight years after resettlement, Chinook Wawa had become a regular language of communication among children, in certain contexts at least (cf. Zenk 1988): the seeds of language endangerment were already sown.

A similar ascendency of Chinook Wawa at Siletz is highlighted in comments by Dorsey (1889:55): “As more than twenty tribes, each having its own dialect, have been consolidated on this reservation, they are obliged to use a common language. So all speak Chinook jargon, and many are learning English.” In providing population figures for the Siletz, Everette (1882) counted the various Oregon Athabaskan groups separately, but noted “all these Indians talk the same language and call it (tu=tu=tene wá wá)” – that is, ‘Tututni’ plus ‘language’ in Chinook Wawa. Everette’s comments, like Dorsey’s, suggest a growing role for Chinook Wawa, and also point to an emergent pan-Athabaskan ethnolinguistic identity, with all varieties of Oregon

“they are very clannish in respect to ‘Burying their Dead’. Each tribe greatly preferring to be by itself; they sometimes bury, in a ‘Common church’ or ‘Burying Ground,’ but this occurs very rarely: this, of course, only refers to ‘Full Blood’s’ of each tribe; as the ‘Half Blood’s,’ (i.e. Indians, whose Parents and Grandparents are all of Different tribes) generally bury their Dead, at the nearest and handiest place they can find.” (Everette 1882:236)

Everette further observed of the Tututni that “these Indians are so mixed by Intermarriage, with other tribes, that they Literally have no tribal Laws or Relations of their own” (1882:236) – surely a distortion, but in keeping with the general picture of the erosion of some cultural practices associated with pre-contact village affiliations.
Ultimately, many people at Siletz also acquired some proficiency in English, and in 1889 Indian Agent Beal Gaither was able to report of people living at Siletz that “[m]ost of them speak and understand English well enough for ordinary intercourse” (Office of Indian Affairs 1889:274).

Nonetheless, the strength of indigenous languages vis-à-vis contact languages at Siletz and elsewhere in Oregon is found in the 1870 report of A.B. Meacham, Superintendent of Indian Affairs for the state, who called for the creation of boarding schools separating children from their families in order to better meet the government’s assimilationist aims. Meacham argued for this on the grounds that “so long as Indian children remain with their parents, spending all their leisure hours at home, where they use their native language only, they forget what is learned through the day” (Office of Indian Affairs 1870:53, emphasis added). A few years later, the continuing influence of Athabaskan languages at Siletz was highlighted in another comment by Everette. Expressing gratitude to Tututni cultural informant William Strong, he wrote, “though Educated in the ‘English Language,’ he still ‘thinks’ in the Indian Language and therefore cannot talk ‘Grammatically’” (1882:90).

The picture that emerges is that by the 1880s the Oregon Athabaskan languages, although still widely spoken some twenty-five years after removal to Siletz, were coming to be identified as varieties of a single language and were encountering new pressure from exogenous contact languages. This shift went hand-in-hand with the dissipation of some salient cultural practices associated with pre-reservation village affiliations, especially among younger people of mixed ancestry. This accords with Leavelle’s (1998:436) assessment of the situation on the adjacent Grand Ronde reservation, where “shared cultural practices and attitudes and the common reservation experience helped the Indians of Grand Ronde overcome the challenges of this diversity to forge an Indian identity rooted in the place and the history of their valley home.”

The importance of this social and sociolinguistic convergence lies in the link between koineization and a focused social identity: as Kerswill (2002) observes, “for a koine to form, the speakers must waive their previous allegiances and social divisions to show mutual solidarity.” The fact that some of the regional divisions separating Athabaskan groups in pre-reservation times were dissipating in the 1880s suggests that the community was becoming focused in the relevant sense and renders it more likely that leveling of erstwhile linguistic differences might have taken place in the same period.

---

8 This usage is already found in one of the earliest summaries of the Athabaskans in the region, the 1854 report of Indian Agent J.L. Parrish, who wrote, “I have found the natives all speaking one language, and from similarity of appearance, habits, and pursuits, consider them as being one nation or people, who from their language may be denominated To-to-tin, or To-to-tut-na; the latter appellation being applied to them by their early visitors” (Office of Indian Affairs 1854:286). There is no evidence that Parrish’s opinion reflects a pre-contact sentiment shared among the Athabaskans themselves, but raises the possibility that community focusing found in the 1880s had antecedents in pre-reservation times.

9 This period is often characterized as a series of injustices committed against indigenous inhabitants of the region, who are portrayed as hapless victims of circumstance. Although these injustices are undeniably real, Leavelle (1998) has emphasized the agency of indigenous groups during this period in showing extraordinary resilience and creativity in adapting to their new circumstances.
4 Leveling

The expectation of the model of new dialect formation presented in §2 is that a koineized Oregon Athabaskan should have developed in the newly-focused speech community that emerged at Siletz. To some extent, this is what the documentary record shows; two specific cases are considered in this section.

4.1 k ~ x Variation in Tututni

One case of leveling involves reflexes of Proto-Athabaskan *q (an aspirated back-velar stop) and *x (a voiceless back-velar fricative), which merge as a fricative /x/ everywhere in Oregon Athabaskan except Galice, where aspirated /k/ is found instead, as illustrated in Tables 1 and 2. This is considered a largely exceptionless categorical merger in published descriptions of Oregon Athabaskan (e.g., Sapir 1914, Bright 1964, Hoijer 1966, Golla 1976).¹⁰

<table>
<thead>
<tr>
<th>ñxeʔ</th>
<th>Coq (Harrington 1942/1981) (1SG.POSS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xʷeʔ</td>
<td>Euc (Golla 2008) (3POSS)</td>
</tr>
<tr>
<td>xʷeʔ</td>
<td>Tut (Landar 1977)</td>
</tr>
<tr>
<td>xʷe:</td>
<td>Mik (Landar 1977)</td>
</tr>
<tr>
<td>xʷe:</td>
<td>ChC (Harrington 1942/1981)</td>
</tr>
<tr>
<td>kʷeʔ</td>
<td>Gal (Landar 1977)</td>
</tr>
<tr>
<td>xʷe</td>
<td>PiR (Dorsey 1884m)</td>
</tr>
<tr>
<td>xʷeʔ</td>
<td>Tol (Bright (1964)</td>
</tr>
</tbody>
</table>

Table 1: ‘Foot’ in Oregon Athabaskan (P-A *qeʔ)

¹⁰In the Upper Coquille form in Table 1, [ñ] is a first person singular possessor agreement prefix. In the other varieties, labialization of the initial consonants indicates a third person possessor, historically derived from a prefix *mi- (Golla 1976, Collins 1985). According to Golla (1976, 2011), in most cases Oregon Athabaskan labialized velars are not direct reflexes of Proto-Athabaskan labialized velars. Instead, Proto-Athabaskan labialized velars merged with their non-labialized counterparts, with contrastive labialization re-emerging subsequently in the vicinity of round vowels and labial consonants (such as the possessed nouns shown here). Thus, while the merger of velar stops and fricatives also affected *qʷ and *xʷ, I will abstract away from details of labialization in the discussion since they are orthogonal to the issues at hand.
Table 2: ‘Canoe’ in Oregon Athabaskan (P-A *xəŋ'ə̃ts’ ‘raft’)

Despite the ubiquity of this merger, a close look at the earliest documentation of the Oregon Athabaskan languages reveals that aspirated [k] was sometimes found outside of Galice as late as the 1880s. Consider first the following selection of entries from a wordlist collected by Hubbard in the mid-1850s, one of the earliest attestations of Tututni (published in Taylor 1860). The first two columns of Table 3 show Hubbard’s original transcriptions and approximate phonetic interpretations of them; the corresponding Euchre Creek forms transcribed by Golla in the twentieth century are provided to illustrate the basis for this interpretation. As might be expected for an adaptation of English orthographic conventions to a new language, <qu> in Hubbard’s list represents a labialized velar stop, as shown by the entries for ‘arm’ and ‘knee’. The sequence <hwo> in the same entries is an indefinite/generic third person possessor prefix x(ʷ)o-, probably etymologically derived from the Proto-Athabaskan areal prefix *qo-, used as an “indefinite human object/possessor” in some Athabaskan languages (Leer 2005:292).

11 Hubbard’s <ch> in ‘woman’, ‘water’, and ‘child’ represents [x], and his <narke> is probably [na:ke], <ar> most likely a representation of vowel length by a speaker of a non-rhotic dialect of English (cf. Hinton 1979:7).

---

11 This prefix is not included in Golla’s (1976) description of Tututni but is pervasive in the documentation of Everette (1882) and Dorsey (1884i). Its use as a generic is highlighted in minimal contrasts documented by Dorsey like <qwic> (= [xʷiš]) ‘any nose’ versus <mic> (= [miš]) ‘his nose’ (1884i). It is most likely the same prefix as the indefinite human possessor ʷə- in Galice identified by Hoijer (1966), but /w/ from *q in Galice is unexpected.
The entry for ‘fire’ in Table 3, from Proto-Athabaskan *qunʔ, shows the expected velar fricative in Tututni, as does the indefinite possessor prefix (from *q) found with body part vocabulary. This is confirmed in another wordlist collected in the same period: Kautz (1855) transcribed a continuant in <whun> (= [xʷən]) ‘fire’ and several words for body parts, such as <whus-see> (= [xʷsɨ]) ‘head’, where <whu-> is the indefinite possessor prefix. In the remaining forms in Table 3, Hubbard recorded a stop rather than a fricative in ‘two’, ‘five’, and ‘ten’. This also is partly confirmed by data from Kautz (1855), who transcribed <nah-ke> (= [na:ke]) ‘two’ and <qui-es-sai> (plausibly [kwase] or [kwese]) ‘ten’ with velar stops, but <swo-lah> (= [swola:]) ‘five’ with a continuant [w]. Nothing is known about the speakers Hubbard and Kautz worked with, but these data show that in the mid-nineteenth century a velar stop was found instead of [x], for some lexical items at least, in a variety of Oregon Athabaskan identified as Tututni spoken by people raised prior to resettlement at Siletz.

A generation after resettlement, both Everette (1882) and Dorsey (1884i) transcribed [x] in Tututni [naxe] ‘two’ and [xʷ:e:se] ‘ten’. Documentation of most varieties of Oregon Athabaskan in the twentieth century show a fricative in ‘five’, contrasting with Hubbard’s stop, but Everette (1882) recorded [kw] in Tututni, matching Hubbard, where Dorsey (1884i) transcribed [w], as did Kautz:13

---

12 Leer (2005) reconstructs Proto-Athabaskan *n̥aː(-qː) for ‘two’ (in the phrase *n̥aː(-qː) dəne: ‘two persons’); Krauss and Leer (1981) reconstruct *qːe-n̥-z:yaːp for ‘ten’ (although the long vowel *ːe: makes it unclear whether or not this is the etymological source for ‘ten’ in Tututni). There is no published Proto-Athabaskan reconstruction for ‘five’ in the sources consulted for this study, but Leer reconstructs a velar stop for this word in an unpublished comparative Athabaskan lexicon.

13 Galice has /k/, as expected, and some other varieties show /w/ instead of /xʷ/. The labialization here is likely due to an original round vowel following /k/ (Golla 1976), as in the Chasta Costa forms transcribed by Sapir and Harrington.
In the 1880s, then, it appears that some speakers of Tututni still produced a velar stop rather than a fricative in ‘five’, a continuation of variation that was also present in the 1850s.

Another item of interest in Hubbard's wordlist is an entry kowlawo, glossed ‘tooth’. The syllable <la> is mysterious, but the root is probably <wo> ‘tooth’ (lenited [ɣ], transcribed as such by Harrington and others) and <ko> a possessive prefix. We might expect initial <hwo> here for the indefinite possessor, matching the other entries for body parts in his list (e.g., the first two forms in Table 3). This form is anomalous until we compare it with Dorsey’s and Everette’s documentation. Dorsey consistently transcribed a fricative for this prefix in Tututni; Everette frequently did so as well, but there are numerous examples (many in a list of body parts) where he transcribed a stop instead. A sample of these forms is given in Table 5.14 Notice that Everette transcribes ‘heart’ twice, once with a fricative [xʷ] in the prefix (in a list of human body parts) and once with [kʷ] (in a list of animal body parts).

---

14 Everette sometimes wrote <k> or <k’w> for this prefix, the apostrophe diacritic indicating that the consonant is “exploded” (Everette 1882:232); here and elsewhere, he appears to have used it for any sort of exceptional laryngeal noise, both glottalization and aspiration. Interestingly, in Dorsey’s (1884e) documentation of Chetco he notes that “kwû can be prefixed to any noun (S.3) to form the generic,” offering examples such as <kwû-ťa> ‘mouth, a mouth’ – notice that in Chetco he transcribed a stop instead of a fricative. One speaker used both variants, explicitly identifying the one with the fricative as Tututni and the one with the stop as Chetco.
Dialect Contact in Oregon Athabaskan

<table>
<thead>
<tr>
<th>Everette (1882)</th>
<th>Dorsey (1884i)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; qwōⁿ'i&gt;</td>
<td>xʷoni</td>
<td>&lt; qoni &gt; xoni</td>
</tr>
<tr>
<td>&lt; qo¹-la&gt;</td>
<td>xola</td>
<td>&lt; qwa-la'&gt; xʷala</td>
</tr>
<tr>
<td>&lt; kwûsa'-qe'&gt;</td>
<td>kʷəsaxe</td>
<td>&lt; qwo-së'-qé'&gt; xʷosexe</td>
</tr>
<tr>
<td>&lt; k'wûc&gt;</td>
<td>kʷəśkxe</td>
<td>&lt; qwic &gt; xʷiš</td>
</tr>
<tr>
<td>&lt; hwû'se'&gt;</td>
<td>xʷose</td>
<td>&lt; qwû-së'&gt; xʷose</td>
</tr>
<tr>
<td>&lt; k'wû'sä&gt;</td>
<td>kʷəsxe</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5: Indefinite possessor prefix in 1880s Tutunni

Similar variation is found in Everette’s transcription of words formed on the roots for ‘foot’ (cf. Table 1 above) and ‘fire’.

<table>
<thead>
<tr>
<th>Everette (1882)</th>
<th>Dorsey (1884i)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; ck'e'&gt;</td>
<td>ške</td>
<td>&lt; qo-qwë'&gt; xo-xʷe</td>
</tr>
<tr>
<td>&lt; tcác-k'we&gt;</td>
<td>č'aš kʷe</td>
<td>&lt; tc'ac qwë'&gt; č'aš xʷe</td>
</tr>
<tr>
<td>&lt; k'wû'g-qe&gt;</td>
<td>kʷəskxe</td>
<td>-</td>
</tr>
<tr>
<td>&lt; qo'-qe'&gt;</td>
<td>xoxe</td>
<td>&lt; qo-qwë'&gt; xoxʷe</td>
</tr>
<tr>
<td>&lt; qe'-nal.-me'&gt;</td>
<td>xe nəlmeʔ</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 6: ‘Foot’ and related vocabulary in 1880s Tutunni (< P-A *qe?)

<table>
<thead>
<tr>
<th>Everette (1882)</th>
<th>Dorsey (1884i)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; qhûn&gt;</td>
<td>xʷən</td>
<td>&lt; qwûn &gt; xʷən</td>
</tr>
<tr>
<td>&lt; qhûn-sît&gt;</td>
<td>xʷənsid</td>
<td>&lt; kwûn-sît' &gt; kʷənsid</td>
</tr>
<tr>
<td>&lt; qhûn-tse'&gt;</td>
<td>xʷəntsxe</td>
<td>&lt; qwûn-se'&gt; xʷənse</td>
</tr>
<tr>
<td>&lt; qûn'-ta'&gt;</td>
<td>xənʔda</td>
<td>&lt; qwûn-ʔa'&gt; xʷənda</td>
</tr>
<tr>
<td>&lt; qûn-ťa'-Let&gt;</td>
<td>xənda ɬed</td>
<td>&lt; qwûn-ʔa'&gt; xʷənda</td>
</tr>
<tr>
<td>&lt; k'wûn-ťa'&gt;</td>
<td>kʷənda</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7: ‘Fire’ and related vocabulary in 1880s Tutunni (< P-A *qun?)

Overall, Everette’s documentation suggests ongoing variation involving [k] and [x] in different tokens of some words and the indefinite possessor prefix.

Dorsey’s notes, collected only two years later, show much less evidence for [k], not only in Tutunni but in the other Lower Rogue dialects he documented as well (Dorsey 1884d,f-l). However, Dorsey did transcribe [kʷ] in the compound ‘charcoal’ built on the root ‘fire’ in Table 7,
where Everette had recorded [xʷ]. Moreover, for at least one lexical item Dorsey’s Tututni consultants showed considerable variation. Recall the form *scharchah* [sxa:xa] ‘child’ from the Hubbard wordlist in Table 3. Both Dorsey and Everette transcribed this word in various ways, sometimes with [k], sometimes with [x], as the normalized forms in Table 8 show.

<table>
<thead>
<tr>
<th>sxexe, txexe</th>
<th>‘child’</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>tsxeyti</td>
<td>‘she had a child’</td>
<td>Six (Dorsey 1884l)</td>
</tr>
<tr>
<td></td>
<td>(-t’i ‘proprietive’)</td>
<td></td>
</tr>
<tr>
<td>skeke yasle</td>
<td>‘boy beginning to walk’</td>
<td>Tut (Dorsey 1884i)</td>
</tr>
<tr>
<td>txexe</td>
<td>‘young man’</td>
<td></td>
</tr>
<tr>
<td>skexe, sxexe</td>
<td>‘boy’</td>
<td>Tut (Everette 1882)</td>
</tr>
<tr>
<td>ts’axe sxexe</td>
<td>‘girl baby’</td>
<td>Mik (Dorsey 1884f)</td>
</tr>
<tr>
<td>ts’axe skeke yasli</td>
<td>‘female infant’</td>
<td></td>
</tr>
<tr>
<td>sxexe</td>
<td>‘boy’</td>
<td>ChC (Dorsey 1884d)</td>
</tr>
</tbody>
</table>

Table 8: ‘Child’ and related vocabulary in Lower Rogue dialects, 1880s

Strikingly, both of Dorsey's Tututni forms in Table 8 were produced by a single speaker, Shem Lafayette, and both of his Chasta Costa forms were elicited from Jake Orton. Notice also the intermediate form [skexe] transcribed by Everette, with one velar stop and one velar fricative, perhaps a hybrid ‘interdialect form’ commonly found in dialect contact situations (cf. Trudgill 1986:62-65). Table 8 gives the impression of speakers using sometimes one, sometimes another of several variants of the same word available at Siletz in the 1880s.15

To summarize the data presented thus far, variation involving [k] and [x] is found in nineteenth century documentation of Tututni. In the 1880s especially, variants of some words and the indefinite possessor prefix are found with [k] instead of [x]. In some cases it is clear that this variation involved different tokens of the same word, sometimes as produced by a single speaker. The timing of Everette’s and Dorsey’s visits to Siletz becomes significant: one generation after the Oregon Athabaskans were consolidated on the reservation, the documentation shows precisely the inter- and intra-speaker variation predicted by Kerswill and Trudgill (2005) for Stage II speakers. The interpretation of these data is that speakers of Oregon Athabaskan who grew up at Siletz from the 1850s onwards were exposed to a wide variety of dialects, some with [k], others with [x], and some (like Tututni) most likely with both. By the 1880s, [k] was losing ground to [x] as leveling of the closely related Rogue River dialects got underway.

Twentieth century documentation of Oregon Athabaskan presents a much less complicated picture, with variation associated with [k] ~ [x] in the 1880s reduced in the last generations of first-language speakers (modulo Galice, discussed in §5.1). Although aspirated velar stops are rare in twentieth century sources, there are enough tokens recorded to warrant consideration.

15 The fricative is probably innovative in this word: Babine-Witsuwit’en, which retains Proto-Athabaskan *q as such (Krauss and Golla 1981), has a word sqaq ‘child’ that appears to be cognate with the Oregon Athabaskan sxexe (Hargus and Tuttle 1997:203-204). Initial [ts] vs. [s] in Table 8 is discussed in section 5.2.
Some normalized examples are given in Table 9. These data are especially difficult to interpret due to the lack of consistency across varieties and transcribers, but they do suggest that aspirated [k] was found outside of Galice in some lexical items in the twentieth century, in some cases in the same items where [k] was transcribed in the nineteenth century. Sapir (1914:284) had found a single exception to the velar stop–fricative merger in Chasta Costa, aspirated back velar [kʷ] occurring in <kʷaša/t'a/ne> [kʷəsta:ne] ‘six’. Later, Harrington (1942/1981) transcribed <k'>, an aspirated velar stop, in a handful of words in Upper Coquille, Chasta Costa, and Tututni (Upper Coquille examples given in Table 9) consistently enough that it is unlikely to be a mis-transcription; he may, however, have intended a front velar rather than back velar articulation in some cases.\(^{16}\) Similar considerations apply to a handful of words transcribed with [k] by Golla (2008). Many of these same words were transcribed with velar stops most plausibly interpreted as aspirated by Dorsey and/or Everette, suggesting that for some lexical items an aspirated velar was also found in the 1880s and transmitted as such transgenerationally throughout the reservation period.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>sakətɬ</td>
<td>-</td>
<td>ša:kkʷəɬ</td>
<td>šaʔkuɬ</td>
<td>‘beaver’</td>
</tr>
<tr>
<td>ūiːkədled</td>
<td>-</td>
<td>ūiːkkəhd</td>
<td>sikad</td>
<td>‘top of head’</td>
</tr>
<tr>
<td>kʷət-tɬ'i</td>
<td>-</td>
<td>xʷəttɬ'i?</td>
<td>kʷətɬ'i?</td>
<td>‘clay, mud’</td>
</tr>
<tr>
<td>kʷəsta:nɛ</td>
<td>kʷəsta:ne</td>
<td>gʷəsta:nne</td>
<td>gʷəstane</td>
<td>‘six’(^{17})</td>
</tr>
<tr>
<td>dəmmilke</td>
<td>dəmelʔge</td>
<td>dəmme:lkɛ?</td>
<td>-</td>
<td>‘pelican, crane’</td>
</tr>
<tr>
<td>katɬ'a</td>
<td>-</td>
<td>ka:tləʔ?</td>
<td>-</td>
<td>‘crab’</td>
</tr>
<tr>
<td>kəlli xe</td>
<td>-</td>
<td>kəlli:xxe:</td>
<td>-</td>
<td>‘moccasin’</td>
</tr>
</tbody>
</table>

Table 9: Aspirated [k] in Rogue River dialects

Variation involving [k] and [x] in particular lexical items is found only rarely in twentieth century documentation: most words have [x] and a few may have [k], but for any given word speakers seem to have converged on one or the other. Direct comparison with the nineteenth century Tututni data presented above is impossible in many cases due to the small number of forms recorded in the twentieth century with speakers identified specifically with the Tututni-Joshua dialect. However, the words in Table 10 were transcribed by Harrington with [x] in both Tututni (as spoken by Lucy Smith) and Upper Coquille (as spoken by Coquelle Thompson), where either Everette (1882) or Dorsey (1884i) had recorded [k] in Tututni.

---

\(^{16}\) Many of Harrington’s aspirated tokens are transcribed with lowercase <k>, as opposed to the much more frequent (when unaspirated) small capital <K>, perhaps indicating a front-velar vs. back-velar articulation.

\(^{17}\) In his documentation of Chasta Costa with Wolverton Orton, who Sapir had worked with thirty-five years earlier, Harrington (1942/1981) transcribed ‘six’ once with aspirated <k'> and once with glottalized <k'>.
Table 10: Variable [x] ~ [k] in nineteenth and twentieth century Rogue River dialects

<table>
<thead>
<tr>
<th>Everette (1882)</th>
<th>Dorsey (1884i)</th>
<th>Harrington (1942/1981)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>kʷəskəs</td>
<td>hʷəl hał</td>
<td>xʷəlxaːɬ</td>
<td>‘flea’</td>
</tr>
<tr>
<td>kxʷne, kʷəxǣ(ʔ)e</td>
<td>-</td>
<td>xənʔeː-</td>
<td>‘river’</td>
</tr>
<tr>
<td>skexe, sxexe</td>
<td>tsxexe, skeke</td>
<td>sxeːxee</td>
<td>‘child’</td>
</tr>
</tbody>
</table>

In each of these cases, Harrington transcribed [x] in both Tututni and Upper Coquille, suggesting that speakers of Tututni in the mid-twentieth century tended to converge on [x] for words that varied between [k] and [x] in the late nineteenth century documentation of Dorsey and especially Everette.

Compare also twentieth century tokens of words containing the root ‘fire’ in Table 11 with the nineteenth century forms in Table 7 above.

Table 11: ‘Fire’ in twentieth century Rogue River dialects (P-A *qənʔ ‘fire’)

<table>
<thead>
<tr>
<th>xʷənʔ</th>
<th>‘fire’</th>
<th>Coq (Harrington 1942/1981)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xʷənʔseʔ</td>
<td>‘ashes’</td>
<td></td>
</tr>
<tr>
<td>kʷənʔθid</td>
<td>‘charcoal’</td>
<td></td>
</tr>
<tr>
<td>xʷənʔ</td>
<td>‘fire’</td>
<td>Euc (Golla 2008)</td>
</tr>
<tr>
<td>xʷənʔsid, kʷənʔsid</td>
<td>‘charcoal’</td>
<td></td>
</tr>
<tr>
<td>xʷən</td>
<td>‘fire’</td>
<td>Tut (Harrington 1942/1981)</td>
</tr>
<tr>
<td>xʷənʔθid</td>
<td>‘charcoal’</td>
<td></td>
</tr>
<tr>
<td>xʷən</td>
<td>‘fire’</td>
<td>Mik (Landar 1977)</td>
</tr>
<tr>
<td>xʷən</td>
<td>‘fire’</td>
<td>ChC (Harrington 1942/1981)</td>
</tr>
<tr>
<td>xʷənəxənəs</td>
<td>‘steamboat’ (lit. ‘fire canoe’)</td>
<td></td>
</tr>
<tr>
<td>kʷənʔ</td>
<td>‘fire’</td>
<td>Gal (Landar 1977)</td>
</tr>
<tr>
<td>kʷənʔθid</td>
<td>‘live coals’</td>
<td></td>
</tr>
</tbody>
</table>

Here again, Tututni xʷənʔθid ‘charcoal’ as transcribed by Harrington (1942/1981) has a fricative rather than a stop. In other dialects, some residual variation between [x] and [k] is found in similar compounds, as demonstrated by the two forms meaning ‘charcoal’ recorded in both Euchre Creek and Upper Coquille.

The upshot is that even in the twentieth century, Stage III speakers (the second generation raised in the mixed dialect reservation community) retained some variation involving [x] and [k], but to a much lesser degree than was found in the 1880s. With few exceptions (such as ‘char-

---

18 Transcribed by Harrington in Tututni [xənʔeːttʃəʃə] ‘river snake’ (ttʃəʃə ‘snake’), Upper Coquille [xənʔeːkkʷət] ‘on the river’ (-kʷət ‘on’).
Dialect Contact in Oregon Athabaskan

could), most lexical items became uniformly [k] or [x], both within and across speakers. The stop variant may have been retained sporadically, but there was an overall tendency to prefer the fricative. What had been a linguistic variable in the Siletz reservation community in the 1880s and probably earlier largely ceased to be so in the twentieth century. The exception, as usual, was Galice, which maintained /k/ throughout, a point that will be considered in §5.1.

4.2 Chetco

As discussed in §3, Chetco-Tolowa is generally considered a language distinct from the Rogue River varieties. The earliest documentation of Chetco is found in Dorsey (1884e), where at least three features distinguish it from his documentation of Tututni (1884i). The first is nasalization: Proto-Athabaskan coda nasals are lost following long vowels in all of Oregon Athabaskan, but they are maintained as vowel nasalization in Chetco-Tolowa (Bright 1964, Golla 2011), transcribed by Dorsey with superscript <n>. The other is nasalization: /l/ versus /l/ in Tututni, and two words beginning with [k] in Chetco-Tolowa (Bright 1964, Golla 2011), transcribed by Dorsey with superscript <n>. The other

<table>
<thead>
<tr>
<th>Chetco (Dorsey 1884e)</th>
<th>Tututni (Dorsey 1884i)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;ts’aⁿiqe&gt;</td>
<td>ts’axe</td>
<td>&lt;tsa’-qe&gt;</td>
</tr>
<tr>
<td>&lt;qwaⁿ-kqi&gt;</td>
<td>xʷək’i</td>
<td>&lt;qwa’-kqέ&gt;</td>
</tr>
<tr>
<td>&lt;tiⁿ-teu&gt;</td>
<td>diču</td>
<td>&lt;di’-i-teu&gt;</td>
</tr>
<tr>
<td>&lt;r’i-tútčl-ta’kwē&gt;</td>
<td>yitutlagʷe</td>
<td>&lt;qwa-tul’-tu-kwē&gt;</td>
</tr>
<tr>
<td>&lt;r’i-mútčl’-kúl-lē&gt;</td>
<td>yimotīlkōle</td>
<td>&lt;qo-múl-kúl lē&gt;</td>
</tr>
<tr>
<td>&lt;natčl-mē&gt;</td>
<td>natime</td>
<td>&lt;nāl’-mē&gt;</td>
</tr>
<tr>
<td>&lt;tčus’-snē&gt;</td>
<td>ĺasne</td>
<td>&lt;tś-snē&gt;</td>
</tr>
<tr>
<td>&lt;tčis’-teu&gt;</td>
<td>ĺisču</td>
<td>&lt;tśš-teu’&gt;</td>
</tr>
</tbody>
</table>

Table 12: Features of Chetco vs. Tututni documented by Dorsey (1884e)

It is important to note that even in the 1880s, there was variation along these lines: while two of the Chetco speakers Dorsey worked with on the whole agree in the features given in Table 12, a third offered Tututni forms like <tśne> ‘man’ (with initial [d] rather than [ɮ]) and <tsa’-qe>

---

19 Tolowa is described as having allophonic [tl] for coda /l/ (Bright 1964), so the first of these differences might involve phonemic /l/ vs. /l/ with the affrication a regular allophonic rule in Chetco. However, there are many other cases where Dorsey transcribes coda <ti> (= [l]) in both Chetco and Tututni, so it is unclear precisely what conditions the difference. Dorsey (1884g) noted variation in Tututni involving [l] and [d] in proximal deictics such as <tśi> – <ti> ‘this’, which can perhaps be interpreted as due to dialect mixing in this period.

20 The possessed nouns ‘calf’ and ‘bladder’ differ in the definiteness of the pronominal element: yi- (Dorsey’s <r’i->) in the Chetco words is a proclitic demonstrative glossed as ‘that’ by Golla (1976); xo- or xʷa- (Dorsey’s <qo->) and <qwa->) in the Tututni words is the indefinite possessor prefix discussed in §4.1.
‘woman’ (with the vowel in the first syllable oral instead of nasalized).\textsuperscript{21}

In the twentieth century, Chetco people living at Siletz appear to have adopted the Tututni values for these features. The main source of information about Chetco in the twentieth century comes from Billy Metcalf, who was interviewed by Homer Barnett and Elizabeth Jacobs in the 1930s. Metcalf’s mother was Chetco and his father was from Rogue River; he was born in Tolowa country in California and lived at Smith River until the age of nine before moving to Siletz (Barnett 1934). Jacobs (1968) estimated him to be in his mid-fifties in 1935, so he was born around 1880, squarely in the generation of speakers that would be identified with the third stage of Kerswill and Trudgill’s (2005) theory of koineization outlined in §2.

In the published edition of a text dictated to her by Metcalf, Jacobs noted that he identified himself as a speaker of Chetco even though his speech lacked the vowel nasalization typical of Chetco as documented by other researchers.\textsuperscript{22} She suggested that “[s]ince Mr. Metcalf had resided at Siletz most of his life, these phonological differences may reflect an assimilation to the predominant Tututni dialect continuum spoken there” (E. Jacobs 1977). Moreover, it is evident in the texts transcribed by Jacobs (1968, 1977) that Metcalf had coda [l] rather than [tɬ] and initial [d] in ‘man’, both features associated with Tututni rather than Chetco. Other direct comparisons with Dorsey’s (1884e) documentation of Chetco are hard to find because the texts transcribed by Jacobs happen not to include vocabulary found in Dorsey’s wordlists. However, some points of comparison are shown in Table 13, and also in Table 14 if the Pistol River dialect documented in Dorsey (1884m), where coda [tɬ] is pervasive, can be used as a proxy for Chetco.\textsuperscript{23}

\textsuperscript{21} The same speaker offered body part vocabulary distinct enough from the other Chetco speakers that Dorsey (1884e) noted “may belong to some other dialect.” There were also evidently differences among speakers in their treatment of possessed noun paradigms. This is further evidence for the high degree of dialect variation at Siletz in this period, as expected in the model of dialect contact presented in §2.

\textsuperscript{22} Jacobs further noted that Metcalf’s speech lacked the prominent retroflex consonants documented in Chetco by other researchers, reflexes of a Proto-Athabaskan palatal series (cf. Bright 1964, Golla 1976, 2011). The retroflexes in Chetco-Tolowa are described by Golla (1976:219) as “strongly r-colored, lip-rounded palatal sibilants.” In most other varieties of Oregon Athabaskan, the distinction between plain versus retroflex consonants “is reduced to a subtle flat versus plain contrast, difficult even for a linguist with a trained ear to hear consistently in the speech of surviving speakers.” Retroflexion is mostly absent in Dorsey’s Chetco notes, but he did transcribe it in the few Tolowa words he recorded and probably was able to hear it accurately. Thus Tolowa <t’aⁿ-qe> ‘woman’ (Dorsey 1884n) but Chetco <ts’aⁿ-qe> (Dorsey 1884e), superscript <r> doubtless indicating retroflexion (cf. Tolowa [t̚aⁿ-e] in Bright 1964:102). Interviews with Chetco speakers conducted by Golla and Bright took place at Smith River (Golla p.c.), so the salient retroflexion detected in their speech might be a contact-induced assimilation by people who lived in close proximity to the Tolowa.

\textsuperscript{23} The Pistol River is located between the Chetco and Rogue Rivers along the Oregon coast. The dialect from this region, as documented by Dorsey (1884m), also matches Chetco with respect to vowel nasalization and the pronunciation of ‘man/male’ and ‘elk’, suggesting that it may have been linguistically intermediate between Chetco-Tolowa and the Lower Rogue dialects (contra Golla 2011:73).
Dialect Contact in Oregon Athabaskan

<table>
<thead>
<tr>
<th>Chetco (Dorsey 1884e)</th>
<th>Tututni (Dorsey 1884i)</th>
<th>Billy Metcalf</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;ts’aⁿ-qe&gt;</td>
<td>ts’āxe</td>
<td></td>
<td>‘woman’</td>
</tr>
<tr>
<td>&lt;qa¹-xūtčel-kêt&gt;</td>
<td>xayəlked</td>
<td></td>
<td>‘(animal) skin’</td>
</tr>
<tr>
<td>&lt;tćus⁻¹-sne&gt;</td>
<td>ʒəsne</td>
<td></td>
<td>‘man’</td>
</tr>
<tr>
<td>&lt;tû-tçl-qût&gt;</td>
<td>tətxəd</td>
<td></td>
<td>‘water’</td>
</tr>
<tr>
<td>&lt;ne-ţetčl¹-yu&gt;</td>
<td>nedetɬyu</td>
<td></td>
<td>‘beads’</td>
</tr>
</tbody>
</table>
twenty-first century Tututni remains to be explored, but this is probably what Trudgill’s (2004) theory of tabula rasa dialect contact would predict. Although the Chetco were a relatively large group at Siletz (see Table 18 in §6 below), they were a minority overall, so Chetco-specific linguistic features would have been at a disadvantage for selection as an Oregon Athabaskan koine developed.

5 Maintenance

While Chetco became largely indistinguishable from Tututni, other varieties of Oregon Athabaskan spoken at Siletz maintained their idiosyncrasies more or less intact into the twentieth century. We will consider two dialects for which this was true: Galice, which was relatively divergent compared to other varieties in the Rogue River dialect network, and Mikwanutni, a dialect that was otherwise thoroughly Lower Rogue in character.

5.1 Galice

Galice, as documented in the twentieth century, maintained several phonological properties that set it apart from the emergent Oregon Athabaskan koine that developed at Siletz: retention of nasalized vowels, merger of Proto-Athabaskan *q and *x as /k/ instead of /x/, denasalization of nasal stops before oral vowels (Krauss and Leer 1981), and lenition of intervocalic /ɣ/. These features are all evident in the nineteenth century documentation of Dorsey (1884b) and in twentieth century sources such as Sapir (1914), Barnett (1934), Hoijer (1966, 1973), M. Jacobs (1968), and Landar (1977). Most of the twentieth century documentation came from speaker Hoxie Simmons (b. 1870s), and his retention of these idiosyncrasies is shown in Table 15.26

<table>
<thead>
<tr>
<th>Galice (Dorsey 1884b)</th>
<th>Galice (Hoijer 1973)</th>
<th>Euchre Creek (Golla 2008)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;kwûn&gt;</td>
<td>kʷan?</td>
<td>xʷən?</td>
<td>‘fire’</td>
</tr>
<tr>
<td>&lt;kwe’&gt;</td>
<td>-ke?</td>
<td>-xe?</td>
<td>‘foot’</td>
</tr>
<tr>
<td>&lt;ţšañ-kê&gt;</td>
<td>ts’á:ke:</td>
<td>tšaxe</td>
<td>‘woman’</td>
</tr>
<tr>
<td>&lt;tčli’-iⁿ-tcū &gt;</td>
<td>liʔ-čoh</td>
<td>liʔ-čuh</td>
<td>‘horse’</td>
</tr>
<tr>
<td>&lt;tû’-dē &gt;</td>
<td>dōde:</td>
<td>dōne</td>
<td>‘person’</td>
</tr>
<tr>
<td>&lt;da’-č &gt;</td>
<td>-da:i</td>
<td>-naye?</td>
<td>‘eye’</td>
</tr>
</tbody>
</table>

Table 15: Dialect features retained in twentieth century Galice

26 Barnett (1934) did not transcribe nasalized vowels with Simmons or with Nettie West, the other Galice speaker he worked with, but his transcriptions are most likely in error since they disagree with all other researchers who worked with Simmons. West, whose speech has the same features as Simmons’, did not move to Siletz until the age of ten, so she might have been unlikely to acquire the Oregon Athabaskan koine in any case, given that people exposed to new dialects in late childhood are less likely to acquire them accurately, if at all (cf. Labov 2007:349-350). Note, however, that Chetco speaker Billy Metcalf moved to Siletz in late childhood as well but apparently did acquire competence in the reservation koine.
The point is simply that the Galice dialect retained the most salient features setting it apart from other members of the Rogue River group even among Stage III speakers raised in the mixed dialect community at Siletz. Moreover, this lack of convergence does not seem to be due to a lack of awareness of differences or of competence in Tututni (i.e. the Oregon Athabaskan koine known by that name), as comments by Simmons paraphrased in Harrington’s fieldnotes suggest:

“It is funny – these coast Inds. can’t talk our Gal. lang, but they can understand what we say. Hox can carry on a conversation in the coast language & can give the equivalent of almost every Gal. word in the coast language, but the coast Indians rarely indulge in trying to say a few Gal. words and then say them poorly.” (Harrington 1942/1981, reel 28, frame 42)

This statement indicates that Simmons had awareness and command of the language of the ‘Coast Indians’ (Tututni), and moreover that there was an asymmetry in accommodation, with the speaker of Galice competent in Tututni but not vice-versa. Nonetheless, this does not seem to have had a significant impact on how Galice was spoken.

5.2 Mikwanutni

As noted in §1, Golla (1976) found no significant non-lexical differences between the Lower Rogue dialects spoken at Siletz in the early 1960s. There is, however, evidence for one feature that set the Mikwanutni dialect apart from the emergent Athabaskan koine in the twentieth century, perhaps a continuation of an older dialect difference. Dorsey (1884f) transcribed <ç> in a number of words before onset <s>, <ç> (= [ʃ]), and <ts> in Mikwanutni but not in Tutuntni. This is shown in the first six forms in Table 16; the remaining forms in the table show that this was not the case for all onset coronal fricatives, even ‘child’ (where Dorsey had transcribed both initial [s] and [ts] in Lower Rogue dialects in the 1880s – cf. Table 8 above). Given that Dorsey transcribed only a few dozen Mikwanutni words, the density of this exceptional <ç> before <s> is striking; it is found almost nowhere else in his documentation of Oregon Athabaskan.

<table>
<thead>
<tr>
<th>Mikwanutni (Dorsey 1884f)</th>
<th>Tutuntni (Dorsey 1884i)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;çsi&gt;</td>
<td>(ç)si</td>
<td>ði’</td>
</tr>
<tr>
<td>&lt;çsú-gâl’&gt;</td>
<td>(ç)səgə</td>
<td>ərga</td>
</tr>
<tr>
<td>&lt;çsú-rxêl’&gt;</td>
<td>(ç)səyə</td>
<td>soyə</td>
</tr>
<tr>
<td>&lt;çsêl-rxêl’&gt;</td>
<td>(ç)seyə</td>
<td>xosexe</td>
</tr>
<tr>
<td>&lt;ts’a-çci&gt;</td>
<td>ts’a-(ç)ʃi</td>
<td>ts’asi</td>
</tr>
<tr>
<td>&lt;çtsê&gt;</td>
<td>(ç)ts’e</td>
<td>xʷəts’e</td>
</tr>
<tr>
<td>&lt;sa’-çlu&gt;</td>
<td>satu</td>
<td>satu</td>
</tr>
<tr>
<td>&lt;sqêl-qê&gt;</td>
<td>sxexe</td>
<td>skeke</td>
</tr>
</tbody>
</table>

Table 16: Exceptional <ç> transcribed by Dorsey in 1880s Mikwanutni
The phonetic interpretation of Dorsey’s <ç> here is uncertain. Its usual value is [θ] in the Bureau of American Ethnology alphabet Dorsey was using, as in the Tututni forms for ‘head’ and ‘hair’ in Table 16. However, Dorsey also used <ç> in ad hoc digraphs he invented to represent sounds not accounted for in the BAE alphabet, such as <çl> for [ɬ]. That he might have intended some sort of quasi-affricated quality is suggested by its appearance in transcriptions of ‘navel’: Mik. <çtsˈɛ> (Dorsey 1884f), PiR <çsˈɛ> (Dorsey 1884m, practically the only example of <çs> outside of Mikwanutni) alongside Tut. <tsˈɛ> (Dorsey 1884i), cf. Euchre Creek -tsˈɛʔ in Golla (2008); the appeal of this interpretation will be apparent shortly.27 While Dorsey’s intent for this letter is uncertain, it is clear enough that in the 1880s Dorsey sometimes detected a difference between fricatives in Mikwanutni versus Tututni and other Lower Rogue dialects that warranted a special representation.

Mikwanutni as spoken by Miller Collins (b. 1880s) in the twentieth century – tape-recorded by Morris Swadesh in the 1950s, transcribed and published by Landar (1977) – has a number of words with initial affricates where other dialects have a plain fricative, shown in Table 17. Merger of Proto-Athabaskan *ts, *dz, *s, and *z as /s/ (/[θ]/ in some dialects) is one of the hallmark features of Oregon Athabaskan (Sapir 1914, Golla 1976). Remarkably, Collins produced [ts] or [tθ] in nearly all of the words for which Proto-Athabaskan reconstructions with *ts are available, the main exception being [sn] ‘star’ (reconstructed by Krauss 2005 variably with or without an initial affricate). Collins also tended to produce [ts] in ‘child’, a word that exhibited a great deal of variation in the 1880s (cf. Table 8). In many cases Collins’ affricate is found alongside [s] or [θ] as a variant, e.g., the root meaning ‘head/hair’ in the compound [θiːdogeʔ] ‘grey-haired’, or the possessed noun [čəθən] ‘some animal’s meat’. The last four forms in Table 17, ‘star’, ‘skin’, ‘liver’, and ‘ear’ show that Collins did not always produce onset coronal fricatives with affrication; in one such case (‘ear’), Dorsey (1884f) had transcribed <çs>.28

27 Dorsey’s Mikwanutni vocabulary also includes <nû-rxultˈ-sûs> [nəɣətsəs] ‘upper eyelid’ alongside <nû-rxeʃə-ən səs> ‘lower eyelid’. <nû-rxû> ~ <nû-rxə> here is equivalent to Tututni naye ‘eye’, and <səs> is səs ‘skin’ (<P-A *zətsə>). The <t> in <nû-rxultˈ-sûs> ‘upper eyelid’, then, can be considered another affricated onset in Mikwanutni (cf. Tut. <qoˈ-naːrəʔ sus> [xonaye səs] ‘someone’s upper eyelid’). But notice that <t> is not transcribed in ‘lower eyelid’, which has plain <səs>.

28 Landar (1977) also transcribed Collins producing [ʰθeː] and [ʰtseː] ‘dry’ (P-A *tsəːy) and [ʰθoː] ‘yellow’ (P-A *tsux”), the former with an affricated onset and the latter without. The affrication in ‘dry’ could be due to an independent process, identified by Harrington (1942/1981) in Upper Coquille, of affrication separating clusters of fricatives like /s/ created at morpheme boundaries (e.g., [ʰθuːwvi] ‘yellow’, [ʰθeː] ‘dry’, both beginning with a prefix ʰ). The process applies more often than not in Harrington’s notes, but apparently not exceptionlessly, as would be the case here with Collins’ ‘yellow’.
**Dialect Contact in Oregon Athabaskan**

<table>
<thead>
<tr>
<th>Mikwanutni (Landar 1977)</th>
<th>Euchre Ck. (Golla 2008)</th>
<th>gloss</th>
<th>Proto-Athabaskan</th>
</tr>
</thead>
<tbody>
<tr>
<td>tse:(ʔ), tθe:, tsi:(ʔ)</td>
<td>siʔ</td>
<td>‘head, hair’</td>
<td>*tsiʔ</td>
</tr>
<tr>
<td>θi:doge?</td>
<td>si-təlge</td>
<td>‘grey-haired’</td>
<td></td>
</tr>
<tr>
<td>tse:γoɭ</td>
<td>siγʷος</td>
<td>‘brain’</td>
<td>*tsi:γa:ŋ’?</td>
</tr>
<tr>
<td>tθən, tsən</td>
<td>sənʔ</td>
<td>‘meat’</td>
<td>*tsəŋ’</td>
</tr>
<tr>
<td>ɕəθən</td>
<td>sənʔ</td>
<td>‘some animal’s meat’</td>
<td></td>
</tr>
<tr>
<td>tse:, θe:</td>
<td>se</td>
<td>‘stone’</td>
<td>*tse:</td>
</tr>
<tr>
<td>tsa:lo:, tθa:lo:</td>
<td>saluʔ</td>
<td>‘tongue’</td>
<td>*tsu: (loʔ)</td>
</tr>
<tr>
<td>txseï, tseï, sxeï</td>
<td>sxe</td>
<td>‘child’</td>
<td></td>
</tr>
<tr>
<td>sən</td>
<td>sin</td>
<td>‘star’</td>
<td>*tsəw, *səw</td>
</tr>
<tr>
<td>səθ</td>
<td>səʔs</td>
<td>‘skin’</td>
<td>*-zəts’</td>
</tr>
<tr>
<td>ɕ’əθəd</td>
<td>-</td>
<td>‘its liver’</td>
<td>*-zət’</td>
</tr>
<tr>
<td>soye:</td>
<td>srəyeʔ</td>
<td>‘ear’</td>
<td>*dʒəγəʔ</td>
</tr>
</tbody>
</table>

Table 17: Onset coronal fricatives in Miller Collins’ Mikwanutni (1950s)

Although the Euchre Creek equivalents in Table 17 are on the whole typical of other Rogue River dialects, affricated [ts] and [tθ] were occasionally produced by speakers identified with varieties other than Mikwanutni as well. Landar (1977), for example, transcribed in Hoxie Simmons’ Galice both [tθe:] and [ɕe:] ‘stone’, [tsi:] ‘hair’, and [wastsʔ] ‘its meat’. Sapir (1914) and Harrington (1942/1981) transcribed [tsxa:xe] and [tsxe:xe] respectively for ‘child’ in Wolverton Orton’s Chasta Costa. Thus, affricated coronal fricatives were found outside of Mikwanutni on occasion, although Mikwanutni appears to have had them most consistently. The variation associated with [ts] and [s] in the twentieth century is reminiscent of that found with velar obstruents in Tututni in the 1880s: There is a lack of consistency both across varieties and in the speech of individual speakers. The significant fact is that whereas the [k] ~ [x] variation had largely leveled out by the mid-twentieth century, the [ts] ~ [s] variation seems to have remained in a state of flux, speakers not having converged on a single value. Moreover, although it remains unclear whether the affrication of Miller Collins in the twentieth century can be fully equated with the exceptional <ç> of Dorsey (1884f), such an interpretation is plausible. If correct, Mikwanutni appears to have maintained from the 1880s into the mid-twentieth century a stronger tendency to affricate onset coronal fricatives than other dialects did.

**6 Discussion and conclusions**

The main findings presented above can be summarized as follows. First, certain differences between closely related varieties of Oregon Athabaskan spoken at Siletz were the targets of leveling in the reservation period, with [k] ~ [x] variation becoming more or less uniformly [x] in all but Galice in the twentieth century. Second, differences between more distantly related Oregon Athabaskan varieties underwent leveling: Chetco in the twentieth century had no nasal-
ized vowels, coda [ɬ] instead of [t̪ɬ], and initial [d] instead of [ǯ] in ‘man/male’ (but [ǯ] still found in ‘elk’). Third, some dialect differences resisted leveling: Galice as spoken by Hoxie Simmons and Nettie West in the twentieth century retained a number of idiosyncrasies, including /k/ instead of /x/, denasalized *n before oral vowels, nasalized vowels, and lenition of intervocalic /ɣ/. Even some differences between closely related Lower Rogue dialects were apparently maintained: Mikwanutni as spoken by Miller Collins had initial affricated [ts] instead of [s] in many words, plausibly a continuation of a dialect difference involving coronal fricatives also manifest in the documentation of Dorsey (1884f). These data bear on issues of current interest in the study of contact-induced language change, especially the degree to which the outcomes of dialect contact are deterministic. Trudgill has argued that they are, in gross outline at least, in cases of tabula rasa contact where “there is no prior-existing population speaking the language in question, either in the location in question or nearby” (2004:26). He maintains that koineization in the development of southern hemisphere dialects of English proceeded according to a principle of majority rule: variants that were common to most dialects survived in the koine and minority variants perished, such that “the minority quite simply and mechanistically accommodated to the minority” (2004:148). There can be no doubt that the relative proportion of speakers in a population (or, more precisely, of particular variants in a mixed dialect population) is an important consideration in many, perhaps most, language contact situations: as Thomason (2001:66) puts it, “numbers count.” The question is whether they always count equally, or whether entrenched social attitudes can alter the outcomes in unpredictable ways, even in the special tabula rasa circumstances that Trudgill (2004) considers.

Although Trudgill initially takes pains to limit his arguments to tabula rasa scenarios (2004:27), he situates his claims in a broader debate about the extent to which people’s attitudes concerning variation matter in determining the course of language change (2004, ch. 7). For Trudgill and others working in this framework, long-term outcomes of dialect contact are the cumulative result of accommodation between speakers in micro-interactional settings. The extent to which people actively use their knowledge of the social world in tuning their speech to interlocutors is an area of current research (cf. Babel 2009 for a recent example). Trudgill adopts a mechanistic view attributed to Labov (2001) whereby metalinguistic evaluations of variants (or people who employ them) play only a minor role in delimiting the diffusion of linguistic changes. Such evaluations might, of course, influence how frequently people interact with each other, and thereby indirectly affect the outcomes of language change, but under this view they have little or no direct impact on linguistic behavior when interaction occurs. This is at odds with practice-based theories grounded in ethnographic observation in which conversational interactions take place in a dynamic social-indexical field and local meanings and related acts of identity are crucial to understanding the diachronic trajectories of sociolinguistic variables (cf. Eckert 2008).

As noted in §2, the Oregon Athabaskan case considered in this paper satisfies Trudgill’s definition of tabula rasa dialect contact. The Siletz reservation was established in northwestern Oregon near Corvallis, and the nearest homeland of any of the Athabaskan groups residing there was at least 130 miles to the south. While the distance between place of origin and locus of resettlement is not as extreme as in the cases of overseas colonial resettlement studied by Trudgill, upon arrival in Siletz no Athabaskan group had any special claim to legitimacy there.

Nonetheless, the outcomes of contact between Athabaskan varieties that underwent mixing at Siletz are not predictable based strictly on the relative proportions of dialect features in the
Dialect Contact in Oregon Athabaskan

population as a whole. Table 18 presents the census figures for the Oregon Athabaskans living at Siletz from available government sources.²⁹

<table>
<thead>
<tr>
<th>Athabaskan</th>
<th>1855</th>
<th>1858</th>
<th>1861</th>
<th>1865</th>
<th>1878</th>
<th>1882</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coquille</td>
<td>105</td>
<td>[313]</td>
<td>[221]</td>
<td>[142]</td>
<td>[84]</td>
<td>[114]</td>
</tr>
<tr>
<td>Sixes</td>
<td>143</td>
<td>242</td>
<td>218</td>
<td>320</td>
<td>74</td>
<td>53</td>
</tr>
<tr>
<td>Euchre Creek</td>
<td>102</td>
<td>84</td>
<td>177</td>
<td>150</td>
<td>59</td>
<td>40</td>
</tr>
<tr>
<td>Tututni</td>
<td>120</td>
<td>202</td>
<td>168</td>
<td>227</td>
<td>137</td>
<td>83</td>
</tr>
<tr>
<td>Joshua</td>
<td>120</td>
<td>179</td>
<td>188</td>
<td>250</td>
<td>84</td>
<td>44</td>
</tr>
<tr>
<td>Mikwanutni</td>
<td>135</td>
<td>129</td>
<td>247</td>
<td>348</td>
<td>-</td>
<td>40</td>
</tr>
<tr>
<td>Chasta Costa</td>
<td>[132]</td>
<td>110</td>
<td>214</td>
<td>162</td>
<td>47</td>
<td>55</td>
</tr>
<tr>
<td>Pistol River</td>
<td>51</td>
<td>-</td>
<td>71</td>
<td>161</td>
<td>57</td>
<td>33</td>
</tr>
<tr>
<td>Galice</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>18</td>
<td>37</td>
</tr>
<tr>
<td>Chetco</td>
<td>241</td>
<td>215</td>
<td>262</td>
<td>221</td>
<td>63</td>
<td>69</td>
</tr>
<tr>
<td>Other</td>
<td>-</td>
<td>575</td>
<td>259</td>
<td>331</td>
<td>466</td>
<td>430</td>
</tr>
<tr>
<td>Total</td>
<td>1149</td>
<td>2049</td>
<td>2025</td>
<td>2312</td>
<td>1089</td>
<td>998</td>
</tr>
</tbody>
</table>

Table 18: Population of Siletz, 1850s-1880s

These numbers must be used with some caution, since it is unclear how the ethnic determinations were made, and the extreme fluctuations in some cases may be due more to administrative whim than to how the people being counted might have chosen to self-identify.³⁰ Nonetheless, Table 18 offers a rough sense of how many speakers of each variety resided at Siletz in the years after the reservation was established.

The key things to notice in Table 18 are that the Chetco, Tututni-Joshua, and Mikwanutni groups were among the most numerous on the reservation while the Galice, for the years in which data are available, were a minority.³¹ It has already been noted in §4.2 that the outcomes for Chetco, whose distinctive dialect features did not survive the leveling associated with koineization, are probably as predicted by Trudgill’s deterministic theory. Although the Chetco were one of the largest groups at Siletz, they were outnumbered by speakers of the Rogue River dialects that lacked nasalized vowels, coda [t], and onset [ʃ] in the relevant vocabulary items. These features were at a disadvantage for selection in the emergent koine and for the most part were not retained, even though the ethnolinguistic label “Chetco” survived.

²⁹ Data are drawn from Office of Indian Affairs reports for each of the years listed. The 1855 report gives the population of each group in situ, before relocation to Siletz; the number of Chasta Costa adult males is missing for that year and is here estimated at 32 (the remainder after subtracting all other groups from the total number of adult males). Population tables for all years except 1855 list simply “Coquill”, so it is unclear whether Upper Coquille (Athabaskan) or Lower Coquille (Milluk Coos) is intended; comments in Everette (1882) suggest that the 1882 figure is for Milluk rather than Upper Coquille. The sharp decline in population in the 1870s is likely due to the fact that people started leaving the reservation to seek their livelihoods elsewhere (Kent 1973:23, Leavelle 1998:449).

³⁰ See Hall (1992) for discussion of the problem of assigning people to discrete ethnic categories where identities were probably considerably more fluid.

³¹ Although it is known from government reports that some Galice people were present at Siletz at least as early as the 1860s, they were not counted separately until the late 1870s, 20 years after the establishment of the reservation.
Trudgill’s deterministic theory, does not, however, explain how idiosyncratic features were maintained in Galice and Mikwanutni. It is surprising that Galice’s distinctive properties survived into the twentieth century at all, given that the Galice contingent was one of the smallest at Siletz. And while it is true that the Mikwanutni were a large group, in most years for which data are available they were no more numerous than the Chetco. Nonetheless, if the interpretation of the data presented in §5.2 is correct, Mikwanutni speakers in both the 1880s and 1950s had a tendency to affricate coronal fricatives. In both cases (Mikwanutni and Galice vis-à-vis Chetco), population asymmetries do not explain the outcomes of contact. If a group as small as Galice maintained its distinctiveness, Chetco should have too; if a group as large as Chetco succumbed to leveling pressure, it is surprising that Mikwanutni speakers did not do so as well. One cannot simply count the number of speakers of Galice, Tututni, Chetco, and Mikwanutni and anticipate the outcomes of contact – something needs to be said about the social configurations in which the contact took place.\[^{32}\]

It becomes important that Galice speakers living at Siletz, in addition to their linguistic distinctiveness, also had a number of cultural practices that were significantly different from other Oregon Athabaskan groups. Their homeland on Galice Creek was adjacent to that of the non-Athabaskan Takelma. According to Drucker’s ethnographic analysis, Galice people had in pre-contact times adopted many aspects of Takelma material and intellectual culture: “Probably never a large group, these people had become through daily contact and intermarriage almost wholly Takelman in culture ... [The Galice Creek culture] is so permeated with Takelman elements as to be scarcely distinguishable from the culture of these alien people” (Drucker 1937:283; cf. Gray 1987). The affinity between Galice-speaking people and the Takelma was a community-internal perception as early as the 1880s: Dorsey (1884b), in a wordlist collected with Galice Creek Jim, paraphrased a Tututni speaker saying that “these are not real Galice Creek words, but Jim’s pronunciation of Tutu words. The G.C. lang resembles the Takelma.” Thus, while from a structural standpoint the Galice language is unmistakably Athabaskan, its speakers were considered more closely affiliated with Takelma. Their ongoing distinctiveness from other Athabaskan-speaking groups in the twentieth century is suggested by Hoxie Simmons’ use of the term “coast Indians” to refer collectively to the non-Galice Athabaskan groups at Siletz (§5.1), and two of Harrington’s consultants suggested that Simmons most likely knew “quite a little” of the Takelma language because “he was raised by the old people” (Harrington 1942/1981, reel 27, frame 568).

This returns to Kerswill’s observations about the relationship between koineization and social integration in a new community, mentioned in §2. While many Athabaskan groups at Siletz began to forge a new, focused identity whose linguistic consequence was koineization, the Galice and perhaps other groups may have maintained a quasi-exogenous status afforded to non-

\[^{32}\] It is also worth pointing out that structural distance also does not seem to be an adequate explanation for the outcomes of dialect contact at Siletz. As often noted in the language contact literature, structural/typological distance between varieties in contact can impede (if not prevent entirely) some kinds of contact effects (cf. Thomason and Kaufman 1988, Winford 2003). Perhaps Galice’s idiosyncracies were maintained because, as discussed in §3.1, it was structurally divergent from other Rogue River varieties spoken at Siletz. However, this would fail to account for the fact that Chetco’s distinctive traits did not survive koineization, even though Chetco was divergent enough to be considered a language distinct from Rogue River varieties (unlike Galice in the classification of Golla 2011). Moreover, for the most part Mikwanutni is otherwise indistinguishable from other Lower Rogue dialects, providing no structural explanation for its retention of unusual onset coronal fricatives.
Athabaskans. Under this view, Galice maintained its distinctive properties because Galice-speaking people did not face the same pressure to assimilate linguistically as did other groups who had more pre-existing social affinities with one another. Whether this explanation can be extended to account for the maintenance of Mikwanutni distinctiveness remains to be seen – the Mikwanutni homeland was located squarely among the other Lower Rogue dialect groups and there is no evidence that the Mikwanutni were culturally distinct in the same way as the Galice were. However, it seems a promising line of explanation to suppose that the linguistic outcomes of Galice’s contact with Tututni and other Athabaskan varieties are due at least in part to the fact that they had to overcome a greater social distance from the outset.

Overcoming social distance, however, is not to be equated with infrequent interaction. In the current context, Trudgill’s determinism could perhaps be salvaged by maintaining that Galice retained its distinctiveness because Galice speakers did not interact regularly with other Athabaskan groups. However, while there was some degree of segregation according to ethnolinguistic divisions at Siletz, this mainly affected non-Athabaskan groups such as the Takelma. The historical record and reports from people who lived there suggest that the Athabaskan groups at Siletz were closely intermingled, sharing resources, intermarrying, and so on, making it improbable that Galice speakers would have interacted with Tututni speakers less frequently than, say, the Chetco did. Instead, Galice remained distinct because its speakers did not simply and mechanistically accommodate to the majority variants in the sea of Athabaskan dialects engulfing them (even if they had some degree of command of those variants, as Hoxie Simmons apparently had). More generally, social factors are very much relevant to understanding the outcomes of linguistic contact, and not only in constraining the frequency of interaction between speakers – even in the tabula rasa cases considered by Trudgill.

Acknowledgments

This research was supported in part by a Graduate Student Fellowship at the Smithsonian Institution in the spring of 2011. Thanks are due to audiences at the 2009 and 2010 Dene/Athabaskan Languages Conference and the 2011 Winter Meeting of the Society for the Study of Indigenous Languages of the Americas for their helpful comments and suggestions, and especially to Victor Golla, whose vivid description of J. Owen Dorsey’s Oregon Athabaskan documentation first sparked my interest in this project.

References


Dorsey, J. Owen. 1884a. *Dakube tede Vocabulary from the Athapascans Formerly Living on Applegate Creek, Oregon October 1, 1884*. National Anthropological Archives: NAA MS 4800 (4.1.1) [372].

Dorsey, J. Owen. 1884b. *Galice Creek (Talt uct un tude) Vocabulary and Grammatical Notes, Formerly Spoken by the Athapascans in Josephine County, Oregon (30 miles north of Kerby) September 18-October 9, 1884*. National Anthropological Archives: NAA MS 4800 (4.1.2) [373].

Dorsey, J. Owen. 1884c. *Upper Coquille (Miciqwutme tunne) and Grammatical Notes August, October 1884*. National Anthropological Archives: NAA MS 4800 (4.1.4) [375]. 112 pages.

Dorsey, J. Owen. 1884d. *Chasta Costa (Cistakawusta) Vocabulary from the Athapascans Formerly Living on Rogue River, Oregon September-October 1884*. National Anthropological Archives: NAA MS 4800 (4.1.3) [374].

Dorsey, J. Owen. 1884e. *Chetco (Teeti tunne) Vocabulary and Grammatical Notes, Formerly Spoken by the Athapascans Living on the Chetco River, Oregon September 20, 25; October 4-21, 1884*. National Anthropological Archives: NAA MS 4800 (4.1.5) [376].

Dorsey, J. Owen. 1884f. *Mikwunu tunne Vocabulary and Grammatical Notes, Spoken by the Tututni Athapascans Formerly Living on Lower Rogue River, Oregon October 15, 1884*. National Anthropological Archives: NAA MS 4800 (4.1.6) [377].

Dorsey, J. Owen. 1884g. *Tutu tunne or Tutu and Joshua Vocabulary, as Spoken in Several Athapascan Villages, with Grammatical and Miscellaneous Notes August 20-October 28, 1884*. National Anthropological Archives: NAA MS 4800 (4.1.7) [378].

Dorsey, J. Owen. 1884h. *Tceme tunni or Joshua Vocabulary and Grammatical Notes August 1884*. National Anthropological Archives: NAA MS 4800 (4.1.7) [379].

Dorsey, J. Owen. 1884i. *Tutu Vocabulary and Grammatical Notes August 20-?, 1884*. National Anthropological Archives: NAA MS 4800 (4.1.7) [380].


Dorsey, J. Owen. 1884k. *Yukitce Vocabulary and Grammatical Notes, Formerly Spoken by the Athapascans Living on Euchre Creek, Oregon September 2, 1884*. National Anthropological Archives: NAA MS 4800 (4.1.8) [386].

Dorsey, J. Owen. 1884l. *Kwatami Vocabulary and Grammatical Notes, Spoken by the Athapascans Formerly living on Sixes Creek, Oregon August 29, 1884*. National Anthropological Archives: NAA MS 4800 (4.1.9) [387].

Dorsey, J. Owen. 1884m. *Naltunne Vocabulary and Grammatical Notes, as Spoken by the Athapascans Formerly Living on Naltunne River, Oregon October 1884*. National Anthropological Archives: NAA MS 4800 (4.1.10) [388].
Dorsey, J. Owen. 1884n. *Quauwate Vocabulary from the Athapascans Formerly Living on Smith River, California September 20, 1884*. National Anthropological Archives: NAA MS 4800 (4.1.11) [389].


Justin Spence
University of California, Davis
Department of Native American Studies
2401 Hart Hall
One Shields Avenue
Davis, CA 95616
jspence@ucdavis.edu
Affix Ordering in Imbabura Quichua*

JOHN SYLAK-GLASSMAN
University of California, Berkeley

1 Introduction

Imbabura Quichua is a Quechua II-B language spoken by over 300,000 people in the Imbabura province of Ecuador (Lewis 2009). It is a prototypical example of an agglutinating language, and rarely blends two distinct concepts in a single morpheme. Imbabura Quichua (henceforth IQ) possesses no prefixes, and verbs containing up to nine distinct suffixes have been recorded, as shown in the elicited example in (1).

(1) Pay-kuna ishkay sumax ushi-gu-kuna-ta papa-kuna-ta
    3P-PL two pretty daughter-DIM-PL-ACC potato-PL-ACC
    yangu-chi1-riya-gri-naya-chi2-naxu-ra-pa-nga
    cook-CAUS-DUR-PROSP-DESID-PERSZR-JACT-STILL-DULC-3P.FUT

“They will still want to make the two pretty little daughters cook potatoes together for a while.”

When so many suffixes are present and able to co-occur, one wonders what principles underlie the formation of such complex words. While one can imagine extralinguistic factors, such as mental processing, playing a significant role, this work will concentrate on formal linguistic factors due to the absence of experimental evidence. Formal linguistic factors underlying the formation of complex, polyaffixal words may be phonological, morphological, syntactic, and semantic in nature. This work will focus especially on the interaction of morphological and semantic factors in constraining possible affix orderings.

This study will restrict its investigation to affix ordering in the IQ verb, specifically between the root and person marking, which are both required in every form and represent, respectively, the leftmost and rightmost peripheries of each verbal form. This study does not consider clitics, such as evidential markers, because they invariantly appear to the right of person marking suffixes. This study also does not consider part-of-speech changing morphology, such as nominalizers, and does not consider subordinating morphology. These elements seem to limit the number of possible morphemes that can occur and tend to occur as far rightward as possible.

*Acknowledgments: I would like to express my sincere thanks to Mariana and Augusto for sharing their language and culture with me and the other members of the UC Berkeley Department of Linguistics 2009-2010 Field Methods class. I would also like to thank the other members of that class for their work in uncovering the structure of the language and especially the precise meaning of the morphemes referenced in this work. Thanks are also due to Lev Michael and Keren Rice for suggestions on how to improve this paper. I am grateful to the Department of Linguistics, Johanna Nichols, and the Beinecke Scholarship Program (part of the Sperry Fund) for support during the time this study was conducted. Any errors are solely my own.

1 There is syncretism only in the suffixes sha, shun, and nga, which encode future tense and, respectively, first person singular, first person plural, and third person (unspecified for number). These suffixes are grouped together for the purposes of affix ordering and are symbolized by the abbreviation FUT.PERS.

2 See the Appendix for a key to abbreviations. In IQ, stress almost always falls on the penultimate syllable. Deviations from IPA: ch = /ʃ/ , sh = /ʃ/ , y = /j/ , r = [z] word-initially, otherwise [r].
The first section reviews the main affix ordering tendencies of all fifteen verbal affixes or affix classes. The second section discusses how the affix ordering data are problematic for a variety of approaches to affix ordering, including the Mirror Principle (Baker 1985), semantic scope (Rice 2000), position classes (Inkelas 1993), and templates (Hyman 2002). The third section presents an alternative analysis that uses Construction Morphology (Booij 2009) to define formal constructions that each apply to single morphemes. These constructions are then brought together under a construction hierarchy, which helps to highlight the commonalities between the constructions.

2 Data and Pre-Theoretical Generalizations

The cited data for this study come from targeted elicitation sessions with a single speaker, Mariana, who served as a consultant for the UC Berkeley Department of Linguistics 2009-2010 Field Methods class. As such, this study must be considered preliminary (but is, the author hopes, indicative of overall tendencies). Another speaker, Augusto, also served as a consultant for the Field Methods class and provided valuable insights, especially into the meaning of the morphemes described here. Both consultants are middle-aged, from the Imbabura Province of Ecuador, native speakers of IQ, and have used the language continuously since childhood. Both are also fluent in Spanish and English.

Several pre-theoretical generalizations limit the variability in affix ordering and help to structure the data. First, the root must be the leftmost element in the verb, and person marking must be the rightmost element. Clitics, such as evidential markers, invariably appear to the right of person marking. Two phonologically and semantically identical morphemes cannot occur in a single form. Finally, there is an absence of morphologically-conditioned phonology. The following subsections review the affix ordering tendencies of the fifteen verbal affixes or affix classes in IQ.

2.1 Causative \( \chi_1 \) (CAUS)

Causative \( \chi_1 \) preferentially appears as root-adjacent (i.e., as far left) as possible. In 12/17 acceptable sentences with \( \chi_1 \) present, \( \chi_1 \) was root-adjacent. Of the 5 acceptable sentences in which \( \chi_1 \) was not root-adjacent, it was never at a distance of more than one morpheme from the root.

(2) Pishku-gu pay-ta kalpa-chi-pa-n
bird-DIM 3SG-ACC run-CAUS-DULC-3P
“The little bird makes him run.”

Causative \( \chi_1 \) is able to alternate in position with the morphemes \( ri \) REFLECTRECIP and \( riya \) DUR without producing semantic differences. Positional alternations between \( \chi_1 \) and \( ri \) might be expected to produce different meanings attributable to one affix taking semantic scope over the other, as is attested in Chichewa (Hyman 2002:248-9) and other Bantu languages. However, such differences in meaning do not occur. In the following example, both sentences have the same

---

3 The person marking morphemes are considered to be a single affix class, since there is no differentiation in ordering between the individual person marking morphemes. It seems reasonable, then, to treat all the person marking morphemes as a single morpheme for the purposes of affix ordering generalizations.

4 As will be explained in more detail, a bigram morphotactic analysis as in Ryan (2010) is promising, but infeasible here for lack of a suitably large corpus.

5 Cole (1982:190) reports not being able to elicit any acceptable utterances with the two morphemes co-occurring.
meaning.

(3) a. Kwitsa-kuna papa-kuna-ta yanu-chi-ri-xu-n
girl-PL potato-PL-ACC cook-CAUS-REFL.RECIP-IMPFV-3
b. Kwitsa-kuna papa-kuna-ta yanu-ri-chi-xu-n
girl-PL potato-PL-ACC cook-REFL-CAUS-IMPFV-3
“The girls make each other cook potatoes.”

Causative $\text{chi}_1$ is able to alternate in position with $\text{mu}$ TRANSLOC, $\text{xu}$ IMPFV, and $\text{wa}$ OBJ, but such alternations lead to differences in meaning, as shown in (4) and (5).\(^6\)

(4) a. Kan kalpa-chi-wa-ni
2SG run-CAUS-1OBJ-2SG
“You make me run.”
b. Kan kalpa-wa-chi-ni
2SG run-1OBJ-CAUS-2SG
“You make me run (= You’re turning me out / kicking me out / making me leave.).”

(5) a. Nyuka pay-ta kalpa-chi-xu-ni
1SG 3SG-ACC run-CAUS-IMPFV-1SG
“I am making him run / am kicking him out / am making him run (out of the house).”
b. Nyuka pay-ta kalpa-xu-chi-ni
1SG 3SG-ACC run-IMPFV-CAUS-1SG
“I am making him run / keep running.”

The additional meaning in (4b) might arise in two ways. First, $\text{chi}_1$ may be taking semantic scope over $\text{wa}$ and in effect reglossing the utterance as, “You make it so that I run,” where the additional meaning arises from a closer association between the root and the first person object. This points to a hypothesis whereby being root-adjacent is being in a position of prominence. A second possibility is that because $\text{chi}_1$ occurs more peripherally and outside its normal position in (4b), it is able to take scope over the affixes occurring left of it, and the causativity or forcing of the action is thereby brought into focus. I follow Rice (2000) in calling this the semantic scope hypothesis.

The following table summarizes the ordering of $\text{chi}_1$ relative to the other verbal affixes and the consequences of ordering alternations.

(6) Affix Positions Relative to $\text{chi}_1$

<table>
<thead>
<tr>
<th>LEFT</th>
<th>$\text{chi}_1$</th>
<th>Freely Varying</th>
<th>Varying with Semantic Consequences</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\text{chi}_1$</td>
<td>$\text{ri}$, $\text{riya}$</td>
<td>$\text{mu}$, $\text{wa}$, $\text{xu}$</td>
<td>$\text{gri}$, $\text{naxu}$, $\text{naya}$, $\text{pa}$, $\text{rka}$, $\text{shka}$, FUT.PERS</td>
</tr>
</tbody>
</table>

2.2 Prospective Aspect $\text{gri}$ (PROSP)

In contrast to $\text{chi}_1$ CAUS, the prospective aspect marker $\text{gri}$, which can be translated as “going to,” does not seem to be root-adjacent by default. It occurs root-adjacent in only 11 of the 26 (42%)
acceptable sentences containing it. The morphemes \( \chi_1 \) and \( \tau \) seem to precede \( \text{gri} \).

(7) a. Kan papa-kuna-ta yaru-\( \chi_1 \)-\( \tau \)-\( \text{gri} \)-wa-ngi
2SG potato-PL-ACC cook-CAUS-DUR-PROSP-1OBJ-2SG

b. Kan papa-kuna-ta yaru-\( \tau \)-\( \chi_1 \)-wa-ngi
2SG potato-PL-ACC cook-DUR-CAUS-PROSP-1OBJ-2SG

“You’re going to make me cook potatoes for a while.”

While Cole (1982:193) finds that \( \text{gri} \) precedes \( \text{naxu} \) JT.ACT, it was \( \text{naxu} \) that preceded \( \text{gri} \) in two out of the three acceptable sentences elicited. The only difference between these data (shown in (8)) and those discussed by Cole (1982:193) is that in these data, one other morpheme, \( \mu \) TRANSLOC, is present.

(8) a. Nyukanchi chaya-mu-\( \text{naxu} \)-\( \text{gri} \)-nchi
1PL arrive TRANSLOC-JACT-PROSP-1PL

b. Nyukanchi chaya-\( \text{naxu} \)-mu-\( \text{gri} \)-nchi
1PL arrive-JACT-TRANSLOC-PROSP-1PL

c. Nyukanchi chaya-mu-\( \text{gri} \)-\( \text{naxu} \)-nchi
1PL arrive-TRANSLOC-PROSP-JACT-1PL

“We’re going to arrive together.”

These sentences show the general tendency for \( \mu \) TRANSLOC to precede \( \text{gri} \). This appears to be the preferred order, especially given the data in (9), which show that \( \mu \) precedes \( \text{gri} \) in isolation.\(^7\) However, this has to be called a tendency rather than an absolute rule, given the data below in (10a).

(9) Nyuka chaya-\( \text{mu} \)-\( \text{gri} \)-ni
1SG arrive TRANSLOC-PROSP-1SG

“I am going to arrive here.”

(10) a. Kanguna chaya-\( \text{gri} \)-ra-mu-ngichi
2PL arrive-PROSP-STILL-TRANSLOC-2PL

b. Kanguna chaya-mu-ra-\( \text{gri} \)-ngichi
2PL arrive-TRANSLOC-STILL-PROSP-2PL

c. Kanguna chaya-ra-mu-\( \text{gri} \)-ngichi
2PL arrive-STILL-TRANSLOC-PROSP-2PL

“You all are still going to arrive here.”

The following data show that \( \text{gri} \) alternates with \( \text{wa} \) 1OBJ and \( \text{pa} \) DULC. Another interesting aspect of these data is that they seem to lend support to the idea that root-adjacent position is the most prominent position for affixes.

(11) a. Pishku-gu kashtu-\( \text{gri} \)-wa-pa-n
bird-DIM bite-PROSP-1OBJ-DULC-3

\(^7\) Because the root and person marking are obligatory, “in isolation” means that the morphemes occur in the position of the ellipsis in ROOT-...-PERSON.NUMBER.
b. Pishku-gu kashtu-gri-pa-wa-n
bird-DIM bite-PROSP-DULC-1OBJ-3
“The little bird is going to bite me (right now).”

(12) a. Pishku-gu kashtu-pa-gri-wa-n
bird-DIM bite-DULC-PROSP-1OBJ-3
b. Pishkugu kashtu-wa-gri-pa-n
bird-DIM bite-1OBJ-PROSP-DULC-3
“The little bird is going to bite me (soon).”

(13) Pishku-gu kashtu-wa-pa-gri-n
bird-DIM bite-1DULC-PROSP-3
“The little bird is going to bite (sometime, now or later, it’s unknown).”

In (11a) and (11b), we see that gri is root-adjacent and that the meaning of the sentence is that the bird will bite right now, this instant. However, in (12a) and (12b), gri occurs one morpheme away from the root, and the change in meaning is that the bird will bite soon, but not immediately. In effect, the immediateness of the action is slightly decreased when gri is moved away from root-adjacent position. In (13), when gri is as far as possible from the root, the immediateness of the action is decreased to null, and gri functions only to assure that the action definitely will happen at some point in the future. These data show that the intensity of gri correlates directly with its adjacency to the root, not with its adjacency to the right periphery, where one might expect it to take scope over other affixes. The degree to which the morpheme is root-adjacent correlates with its intensity of meaning while the degree to which it is able to take scope over other morphemes seems to be irrelevant.

Finally, although it seems to precede it in isolation, gri is able to alternate with xu without producing any differences in meaning.

(14) a. Nyuka papa-kuna-ta yanu-gri-xu-sha
1SG potato-PL-ACC cook-PROSP-IMPFV-1SG.FUT
b. Nyuka papa-kuna-ta yanu-xu-gri-sha
1SG potato-PL-ACC cook-IMPFV-PROSP-1SG.FUT
“I am going to/will be cooking the potatoes.”

The following table summarizes the ordering of verbal affixes relative to gri.

<table>
<thead>
<tr>
<th>LEFT</th>
<th>gri</th>
<th>Freely Varying</th>
<th>Varying with</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>chi₁, ri, riya</td>
<td>gri</td>
<td>mu, naxu, ra, xu</td>
<td>pa, wa</td>
<td>naya, rka, shka, FUT.PERS</td>
</tr>
</tbody>
</table>

2.3 Translocative mu (TRANSLOC)

The translocative mu, roughly meaning “to here,” strongly favors root-adjacent order. Out of the 21 occurrences of mu in acceptable sentences, 15 were root-adjacent. Of the remaining 6 non-root-adjacent occurrences, only one had mu occurring at a distance greater than one morpheme from
the root. This one example, in (16) below, is also the only acceptable sentence elicited in which *gri* does not follow *mu*.

(16) Kanguna chaya-gri-ra-mu-ngichi
    2PL  arrive-PROSP-STILL-TRANSLOC-2PL
    “You all are still going to arrive here.”

The translocative *mu* does not seem to occur with *naya DESID* in isolation, and never with *ra STILL* or *riya DUR*. The affixes *mu* and *wa* alternate, apparently without any semantic effects, as shown in (17) and (18).

(17) a. Kan waxta-mu-wa-ngi
    2SG hit-TRANSLOC-1OBJ-2SG
    “You come and hit me.”
b. Kan waxta-wa-mu-ngi
    2SG hit-1OBJ-TRANSLOC-2SG

c. Kan waxta-wa-mu-xu-ngi
    2SG hit-1OBJ-TRANSLOC-IMPFV-2SG
    “You’re coming and hitting me.”

The following table summarizes the order of affixes relative to *mu*.

(19) Affix Positions Relative to *mu*

<table>
<thead>
<tr>
<th>LEFT</th>
<th><em>mu</em></th>
<th>Freely Varying</th>
<th>Varying with Semantic Consequences</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ri</em></td>
<td><em>mu</em></td>
<td><em>chi₁, gri, naxu, ra, wa</em></td>
<td>naya, pa, rka, shka, xu, FUT.PERS</td>
<td></td>
</tr>
</tbody>
</table>

### 2.4 Joint Action *naxu* (*JT.ACT*)

The morpheme *naxu* signifies joint action and is translated as “together.” Out of the 9 occurrences of *naxu* in acceptable sentences, 6 were in root-adjacent position. The morpheme *naxu* does not combine in isolation with *ra STILL* or *naya DESID*, but these two morphemes do occur with many others in the example in (1). The following chart shows the possible alternations in the ordering of the ten affixes in (1).

(20) Possible Affix Reorderings in *yanu-čhi₁-riya-gri-naya-čhi₂-naxu-ra-pa-nga* from (1)

```
yanu  čhi₁  riya  gri  naya  čhi₂  naxu  ra  pa  nga
    naxu  čhi₂  naya
        naxu  čhi₂
            naya  gri
               gri  riya
```
The morpheme *naxu* appears to alternate with *gri* and *mu* without any semantic consequence.

(21) a. Nyukanchi chaya-mu-naxu-gri-nchi
    1PL    arrive-TRANSLOC-JACT-PROSP-1PL

b. Nyukanchi chaya-naxu-mu-gri-nchi
    1PL    arrive-JACT-TRANSLOC-PROSP-1PL

c. Nyukanchi chaya-mu-gri-naxu-nchi
    1PL    arrive-TRANSLOC-PROSP-JACT-1PL

“We’re going to arrive together.”

The full ordering of affixes relative to *naxu*, then, is as follows in (22).

(22) Affix Positions Relative to *naxu*

<table>
<thead>
<tr>
<th>LEFT</th>
<th><em>naxu</em> Freely Varying</th>
<th>Varying with Semantic Consequences</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>chi₁</em>, <em>naya</em>, <em>ri</em>, <em>xu</em></td>
<td><em>naxu</em> <em>chi₂</em>, <em>gri</em>, <em>mu</em>, <em>riya</em>, <em>ra</em></td>
<td><em>pa</em>, <em>rka</em>, <em>shka</em>, <em>wa</em>, <em>FUT.PERS</em></td>
<td></td>
</tr>
</tbody>
</table>

2.5 Desiderative *naya* (DESID)

The desiderative marker *naya* appears in root-adjacent position in only 4 out of 12 acceptable sentences containing it. Cole (1982:195) claims that *naya* alternates with *xu* IMPFV and that their ordering shows semantic scope effects. I was not able to confirm this, but *naya* is able to alternate with *wa* and *pa* similarly to how *gri* is able to alternate with those same two affixes. The data below show the alternation of *naya* with *wa* and *pa*, and show how this alternation seems not to have any semantic effects.

(23) a. Pishku-gu kashtu-naya-pa-wa-nga
    bird-DIM bite-DESID-DULC-1OBJ-3SG.FUT

b. Pishku-gu kashtu-naya-wa-pa-nga
    bird-DIM bite-DESID-1OBJ-DULC-3SG.FUT

c. Pishku-gu kashtu-wa-pa-naya-nga
    bird-DIM bite-1OBJ-DULC-DESID-3SG.FUT

“The little bird will want to bite me.”

The morphemes *naya* and *ra* do not combine in isolation. When they occur in the presence of other suffixes, however, *ra* is able to alternate freely in position with respect to *naya*.

(24) Nyuka chaya-ra-mu-naya-ni
    1SG    arrive-STILL-TRANSLOC-DESID-1SG

“I still want to arrive here.”

The overall ordering of affixes relative to *naya* is as follows.

The overall ordering of affixes is as follows.
2.5.1 Personalizer \textit{chi}_2 (PERSZR)

Cole (1982:192) discusses a “personalizing” construction involving \textit{naya} and \textit{chi}_2 in which the otherwise accusative-marked subject of an impersonal construction is instead marked as nominative. The verb then overtly agrees with this subject rather than showing the default third person agreement characteristic in impersonal constructions in IQ. Impersonal constructions are used with certain roots that denote experiential or sensory notions. In the pair of sentences below, (26a) is an example of a standard impersonal construction while (26b) is an example of the “personalized” construction.

(26) Examples of the personalizing construction with \textit{naya} and \textit{chi}_2 (Cole 1982:192)
   a. Nyuka-ta miku-naya-n \\
      1SG-ACC eat-DESID-3 \\
      “I want to eat.” / “I feel like eating.”
   b. Nyuka miku-naya-chi\textsubscript{2}-ni \\
      1SG.NOM eat-DESID-PERSZR-1SG \\
      “I want to eat.”

This personalizing \textit{chi}_2 most often directly follows \textit{naya}. It does not require that \textit{naya} be present as long as the root calls for an accusative experiencer subject (when \textit{chi}_2 is not present). The data shown in (20) demonstrate that \textit{chi}_2 may occur before \textit{naya} or after it, and may alternate with \textit{naxu}.\footnote{The consultant for this study noted that ordering \textit{chi}_2 before \textit{naya} is something that an elderly speaker would do.}

The overall ordering of affixes relative to \textit{chi}_2 is as follows.

(27) Affix Positions Relative to \textit{chi}_2

<table>
<thead>
<tr>
<th>LEFT</th>
<th>\textit{naya}</th>
<th>Freely Varying</th>
<th>\textit{chi}_2</th>
<th>Varying with Semantic Consequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{chi}_1, gri, mu, ra, ri</td>
<td>\textit{naya}</td>
<td>\textit{chi}_2, pa, wa, xu</td>
<td>\textit{naxu}</td>
<td>riya, rka, shka, FUT.PERS</td>
</tr>
</tbody>
</table>

2.6 Dulcitive \textit{pa} (DULC)

The morpheme \textit{pa} encodes a “dulcitive” meaning on verbs.\footnote{The term “dulcitive” was coined by Tom Recht.} Cole (1982:185) calls this affix “honorable,” but our consultants described it as indicating that the situation described by the verb was viewed “with affection” by the speaker.
With regard to ordering, *pa* tends to be positioned toward the right periphery of the word. Only 11 of 43 acceptable sentences which contained *pa* featured it in root-adjacent position. *pa* is able to alternate freely with *gri, naya, ra, riya, shka, wa*, and *xu*. The total affix ordering relative to *pa* appears to be as follows.

(28) Affix Positions Relative to *pa*

<table>
<thead>
<tr>
<th>LEFT</th>
<th>Freely Varying</th>
<th>Varying with Semantic Consequences</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>chi₁, mu, naxu, ri</em></td>
<td><em>pa</em> <em>gri, naya, ra, riya, shka, wa, xu</em></td>
<td><em>ra</em>, FUT, PERS</td>
<td></td>
</tr>
</tbody>
</table>

2.7 "Still" *ra* (STILL)

The morpheme *ra* implies a continuity of the action of the verb and is rendered as “still” in English. The affix *ra* seems to have stricter co-occurrence restrictions than many of the other verbal morphemes in Imbabura Quichua and may be of limited productivity. In isolation, *ra* may only combine with *gri* PROSP, *riya* DUR, and *xu* IMPFV, which appear to be the dedicated aspectual markers in Imbabura Quichua.¹⁰

While the reason why *ra* should combine only with aspectual morphemes is unclear, its ordering with respect to them in isolation is apparent. The morpheme *ra* precedes *gri* and *riya* in isolation, but follows *xu*.

(29) a. Nyuka kalpa-ra-gri-ni
    1SG run-STILL-PROSP-1SG
    “I am still going to run.”

b. Warmi papa-kuna-ta yanu-ra-riya-n
   woman potato-PL-ACC cook-STILL-DUR-3
   “The woman still cooks potatoes (for a while).”

c. Nyuka chalwa-ta miku-xu-ra-ni
   1SG fish-ACC eat-IMPFV-STILL-1SG
   “I am still eating the fish.”

In the presence of additional morphemes, these ordering statements remain true for *riya* and *xu*, but *gri* may alternate in position with *ra* without semantic effects. In addition, the morphemes *mu, naxu, naya, pa*, and *ri* were able to alternate in position with *ra* without semantic effects.

Semantic scope effects were present in the data when *ra* and the FUT, PERS affixes co-occurred. These data were the only cases in which person marking was able to be re-ordered with another morpheme.¹¹

---

¹⁰ The morpheme *riya* is the subject of debate. Cole (1982:184) calls it durative, but recent investigation points to interpreting it as a non-recent past marker. In addition, *shka* seems not to function as an aspectual marker (Cleary-Kemp 2013).

¹¹ For elicitation, I situated these phrases in the hypothetical context that a person had recently gotten drunk to the point of sickness on *chicha* (a traditional Andean homebrew made from masticated and fermented maize), and yet was determined to drink it again. (30a) was said to be “more aggressive” while (30b) implied continuity, and a
(30)  a. Nyuka-ka aswa-ta ufy-a-ra-sha  
1SG-TOP chicha-ACC drink-STILL-1SG.FUT  
“I will still drink chicha.”  
b. Nyuka-ka aswa-ta ufy-a-sha-ra  
1SG-TOP chicha-ACC drink-1SG.FUT-STILL  
“I still will drink chicha.”  

The overall ordering of affixes relative to ra is as follows in (31).

(31) Affix Positions Relative to ra  

<table>
<thead>
<tr>
<th>LEFT</th>
<th>ra</th>
<th>Freely Varying</th>
<th>Varying with Semantic Consequences</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>chi1</td>
<td>ra</td>
<td>gri, mu, naxu, naya, pa, ri, xu</td>
<td>FUT.PERS</td>
<td>riya, shka, rka</td>
</tr>
</tbody>
</table>

2.8 Reflexive or Reciprocal ri (REFL.RECIP)

The REFL.RECIP morpheme ri is strongly root-adjacent and occurs in that position in 12/17 sentences which were deemed acceptable. As explained in §2.1, ri can be freely ordered with respect to chi1 CAUS. The affix ri also seems to be able to freely vary in position with ra STILL, riya DUR, and xu IMPFV. This ordering freedom is illustrated by the pairs in (32), (33), and (34).

(32)  a. Nyukanchi riku-ra-xu-nchi  
1PL see-STILL-REFL.RECIP-IMPFV-1PL  
b. Nyukanchi riku-ri-xu-nchi  
1PL see-REFL.RECIP-STILL-IMPFV-1PL  
“We are still looking at each other.”

(33)  a. Pishku-gu-kuna riku-riya-ri-pa-rka  
bird-DIM-PL see-DUR-REFL.RECIP-DULC-PA  
b. Pishku-gu-kuna riku-riya-ri-pa-rka  
bird-DIM-PL see-REFL.RECIP-DUR-DULC-PA  
“The little birds looked at each other (for a while).”

(34)  a. Nyukanchi riku-ri-xu-nchi  
1PL see-REFL.RECIP-IMPFV-1PL  
b. Nyukanchi riku-xu-ri-nchi  
1PL see-IMPFV-REFL.RECIP-1PL  
“We are looking at each other.”

The total affix ordering relative to ri is as below in (35).
Affix Ordering in Imbabura Quichua

(35) Affix Positions Relative to $ri$

<table>
<thead>
<tr>
<th>LEFT</th>
<th>$ri$</th>
<th>Freely Varying</th>
<th>Varying with Semantic Consequences</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ri$</td>
<td>$chi_1$, $ra$, $riya$, $xu$</td>
<td></td>
<td></td>
<td>gri, mu, naxu, naya, pa, shka, rka, FUT.PERS</td>
</tr>
</tbody>
</table>

2.9 Durative $riya$ (DUR)

The morpheme $riya$ occurs in root-adjacent position in 11/19 acceptable sentences that contain it. The affix $riya$ may co-occur with every verbal morpheme examined here except for $mu$, with which it may not co-occur in isolation or with other morphemes present. The ordering of affixes relative to $riya$ is as follows.

(36) Affix Positions Relative to $riya$

<table>
<thead>
<tr>
<th>LEFT</th>
<th>$riya$</th>
<th>Freely Varying</th>
<th>Varying with Semantic Consequences</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>gri, naya, ra</td>
<td>$riya$</td>
<td>$chi_1$, naxu, pa, $ri$, $shka$</td>
<td></td>
<td>rka, wa, FUT.PERS</td>
</tr>
</tbody>
</table>

2.10 Eyewitness Past Tense $rka$ (PA)

The morpheme $rka$ indicates not only past tense, but that the action of the verb was witnessed (see Cleary-Kemp 2013, this volume). The affix $rka$ is strongly biased towards occurring on the right periphery. Every affix must appear to the left of $rka$ except for $shka$. The affixes $shka$ and $rka$ together form the pluperfect, or past perfect (Cleary-Kemp 2013). In this instance, $shka$ is sometimes able to vary in order with $rka$, as shown below.12

(37) a. Nyuka chaya-mu-shka-rka-ni
   1SG arrive-TRANSLOC-PERF-PA-1SG
b. Nyuka chaya-mu-rka-shka-ni
   1SG arrive-TRANSLOC-PA-PERF-1SG
   “I had arrived here.”

(38) a. Kanguna papa-kuna-ta yanu-naxu-shka-rka-ngichi
   2PL potato-PL-ACC cook-JACT-PERF-PA-2PL
b. Kanguna papa-kuna-ta yanu-naxu-rka-shka-ngichi
   2PL potato-PL-ACC cook-JACT-PA-PERF-2PL
   “You all had cooked potatoes together.”

As is clear from the glosses, the re-ordering of $shka$ and $rka$ has no semantic impact. The more prevalent order and the one that occurs in isolation is $shka$-$rka$. The ordering of all verbal morphemes with respect to $rka$ is as follows.

12 It should be noted that Jessica Cleary-Kemp, in specific elicitation of sentences with $rka$ and $shka$ was unable to elicit an acceptable verbal complex with the order $rka$-$shka$. This prompts doubt that the variation in the ordering of $rka$ and $shka$ is unmarked or common.
2.11 Non-Eyewitness Past Tense shka (PA.NWIT)

The morpheme shka is heavily biased toward appearing at the right periphery. It is able to vary in ordering to a very limited degree with rka, and is able to vary in ordering more robustly with pa DULC and riya DUR. The morpheme rka appears preferentially to the right of shka, and all other morphemes must appear to the left. The following examples show the free ordering of shka with pa and riya.

(40) a. Uchila ushi-kuna papa-kuna-ta yanu-riya-pa-shka
   small daughter-PL potato-PL-ACC cook-DUR-DULC-PA.NWIT

   b. Uchila ushi-kuna papa-kuna-ta yanu-shka-pa-riya-n
      small daughter-PL potato-PL-ACC cook-DULC-PA.NWIT-DUR-3

   c. Uchila ushi-kuna papa-kuna-ta yanu-shka-pa-riya-n
      small daughter-PL potato-PL-ACC cook-PA.NWIT-DULC-DUR-3

   “The little daughters have cooked the potatoes for a while.”

The ordering of all verbal morphemes with respect to shka is as follows. Cleary-Kemp (2013) reports that shka seems not to be able to co-occur with the FUT.PERS morphemes for our speakers, even though this is attested in Cole (1982). For detailed information on shka, see Cleary-Kemp (2013).

(41) Affix Positions Relative to shka

<table>
<thead>
<tr>
<th>LEFT</th>
<th>shka Freely Varying</th>
<th>Varying with Semantic Consequences</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>chi₁, gri, mu, naxu, naya, ra, ri, wa, xu</td>
<td>shka</td>
<td>pa, riya, rka</td>
<td></td>
</tr>
</tbody>
</table>

2.12 First Person Object Marker wa (1OBJ)

As mentioned in §2.1, wa exhibits semantic scope effects with chi₁. The first person object marker wa is able to vary freely with the morphemes gri, mu, naya, pa, and xu. (18) shows the alternation of wa with mu and xu. The ordering of all morphemes with respect to wa is as follows.
2.13 Imperfective xu (IMPFV)

The imperfective $xu$ can be re-ordered with $chi_1$ to produce semantic scope effects, as shown in (5). The ordering of all affixes with respect to $xu$ is as follows.

2.14 Future Tense Person Markers (FUT.PERS)

One means of marking future tense (use of $gri$ is the major alternative) is to use the syncretic suffixes $sha$, $shun$, and $nga$, for 1SG, 1PL, and 3rd person, respectively. These suffixes are much like $rka$ in that they vary only with one morpheme, $ra$. Every other morpheme appears to the left of FUT.PERS, and none must obligatorily appear to the right.

2.15 Summary of Affix Ordering Data

The following table is a compilation of the tables provided for each morpheme showing which morphemes must appear to its left, which can be freely ordered with respect to it, which can be re-ordered with it but with semantic scope complications, and which morphemes must be to its right.
(45) Summary of Affix Ordering Data

<table>
<thead>
<tr>
<th>LEFT</th>
<th>M Freely Varying</th>
<th>Varying with Semantic Consequences</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi_1$, ri, riya</td>
<td>$\mu$, wa, xu</td>
<td>gri, naxu, naya, pa, rka, shka, FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>$\chi_1$, gri, mu, ra, ri</td>
<td>$\chi_2$, naya, pa, wa, xu, naxu</td>
<td>riya, rka, shka, FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>$\chi_1$, ri, riya</td>
<td>gri, mu, naxu, ra, xu</td>
<td>pa, wa</td>
<td>naya, rka, shka, FUT.PERS</td>
</tr>
<tr>
<td>ri</td>
<td>$\mu$, $\chi_1$, gri, naxu, ra, wa</td>
<td>naya, pa, rka, shka, xu, FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>$\chi_1$, naya, ri, xu</td>
<td>naxu, $\chi_2$, gri, mu, riya, ra</td>
<td>pa, rka, shka, wa, FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>$\chi_1$, gri, mu, ra, ri</td>
<td>naya, $\chi_2$, pa, wa, xu</td>
<td>naxu, riya, rka, shka, FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>$\chi_1$, mu, naxu, ri</td>
<td>pa, gri, naya, riya, shka, wa, xu</td>
<td>rka, FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>$\chi_1$, ra, riya, xu</td>
<td>$\mu$, $\chi_1$, naxu, pa, ri, xu</td>
<td>FUT.PERS</td>
<td>riya, shka, rka</td>
</tr>
<tr>
<td>gri, naya, ra</td>
<td>riya, $\chi_1$, naxu, pa, ri, shka</td>
<td>rka, wa, FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>$\chi_1$, gri, mu, naxu, naya, pa, ra, ri, riya, wa, xu</td>
<td>rka, shka</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi_1$, gri, mu, naxu, naya, ra, ri, wa, xu</td>
<td>shka, pa, riya, rka</td>
<td>FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>naxu, riya</td>
<td>wa, gri, mu, naya, pa, xu</td>
<td>$\chi_1$, rka, shka, FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>mu, ri</td>
<td>xu, gri, naya, pa, ra, $\chi_1$, wa</td>
<td>naxu, rka, shka, FUT.PERS</td>
<td></td>
</tr>
<tr>
<td>$\chi_1$, mu, naxu, naya, pa, ri, riya, shka, wa, xu</td>
<td>FUT.PERS</td>
<td>ra</td>
<td></td>
</tr>
</tbody>
</table>

3 Problems for Previous Accounts of Affix Ordering

Five main kinds of analyses have figured prominently in explaining affix ordering. Those five kinds of analyses are: 1) the Mirror Principle (Baker 1985), 2) relevance (Bybee 1985), 3) position
classes (Inkelas 1997), 4) semantic scope (Rice 2000), and 5) templates (Hyman 2002). Some of these analyses are able to offer insight into affix ordering in IQ, but all face problems with the data.

3.1 Mirror Principle

Baker’s (1985:374) Mirror Principle states informally that “morphological derivations must directly reflect syntactic derivations.” The idea that morphology should reflect syntax is powerful, and has been developed most clearly in Distributed Morphology (Halle and Marantz 1993 inter alia). Both the Mirror Principle and Distributed Morphology assume that a universal syntactic structure, provided by the theoretical frameworks of Government and Binding or Minimalism, is what is reflected in affix ordering.

However, an explanation based on the Mirror Principle must grapple with the fact that affixation and morphological constructions play such a large role in IQ that there are relatively few inter-word syntactic relationships that can be brought to bear on affix ordering data. This is not a formal problem since the syntactic derivation is assumed to proceed according to principles inherent in Universal Grammar. However, it is an empirical problem in that there is very little language-internal evidence that the syntactic derivation inherent in the morphological derivation is correct, which weakens the reliability of a Mirror Principle explanation.

3.2 Relevance

Muysken (1986:639) discusses affix ordering in Bolivian Cochabamba Quechua, and in doing so provides a synopsis of how Bybee’s (1985) principle of relevance as a determiner of affix ordering operates in a Quechua language. The idea of relevance determining affix ordering is that affixes which are more directly relevant to the root’s lexical content will be positioned closer to the root while those that are relevant either mostly or exclusively to the syntax will be expressed farther from the root. Bybee’s theory predicts the order: STEM-VALENCE-CHANGING-VOICE-ASPECT-TENSE-MOOD-NUMBER-PERSON-GENDER.

Muysken (1986:639) shows that this order does not match the order attested in Cochabamba Quechua. Similarly, one can confirm through examining (45) that Bybee’s predicted order is not attested as a general tendency in IQ. Bybee’s proposal wins some support in that _ci1 CAUS_ occurs root-adjacently by default. However, it is able to freely alternate with an aspectual marker, _rika DUR_. This may not pose a problem since aspect is also very close to the root in Bybee’s theory. However, the first person object marker _wa_, predicted to be far to the right, can appear before _ci1 CAUS_. Moreover, the derivational TRANSLOC morpheme _mu_ seems to be able to vary freely with _wa_. Upon consulting the table in (45), it becomes clear that there are many such alternations that straddle the informal position classes embedded in Bybee’s theory. The theory predicts that this variation should not occur since presumably the semantic content of the morphemes does not change from use to use, and thus one might expect them to stand in the same relevance relations and thus to be positioned in a fixed order. Because of this variation, Bybee’s theory is unable to explain all the affix ordering tendencies in IQ.
3.3 Position Class Morphology

The data in (47) present an interesting problem for a Position Class analysis (à la Inkelas 1993) given the data in (46). The data in (46) (repeated from (9)) seem to show unambiguously that \textit{mu} precedes \textit{gri} in isolation.

(46) Nyuka chaya-mu-gri-ni
    1SG arrive-TRANSLOC-PROSP-1SG
    “I am going to arrive here.”

Under a position class analysis, one might take this as evidence for two position classes. However, the data show that this division cannot be fully maintained. First, (47b) (repeated from (10b)) shows \textit{mu} preceding \textit{ra} and \textit{gri}. One might suppose that \textit{ra} and \textit{gri} should be put together into one position class. However, consider the data in (47c) (repeated from (10c)), which show that \textit{mu} may intrude between \textit{ra} and \textit{gri}.

(47) a. Kanguna chaya-gri-ra-mu-ngichi
    2PL arrive-PROSP-STILL-TRANSLOC-2PL
b. Kanguna chaya-mu-ra-gri-ngichi
    2PL arrive-TRANSLOC-STILL-PROSP-2PL
c. Kanguna chaya-ra-mu-gri-ngichi
    2PL arrive-STILL-TRANSLOC-PROSP-2PL
    “You all are still going to arrive here.”

This destroys the validity of the claim that \textit{ra} and \textit{gri} form a position class to the exclusion of \textit{mu}. The problem of a position class analysis can be shown as below in (48).

(48) \[ \text{[mu]} \ [ \text{[ra]} \ [\text{gri}]] \]
    \[ \text{[ra]} \ [\text{mu}] \ [\text{gri}] \]
    \[ \text{[gri]} \ [\text{ra}] \ [\text{mu}] \]

This series of data shows that a bracketing paradox arises when attempting to formulate a position class analysis because it shows that \textit{mu} and \textit{ra} cannot be considered one position class to the exclusion of \textit{gri}, nor can \textit{gri} and \textit{ra} form a position class to the exclusion of \textit{mu}. The only solution is to unify all three into one position class where variation is possible. The data from (8) show a bracketing paradox of exactly the same kind except that \textit{ra} is replaced by \textit{naxu} there. The only solution to these two bracketing paradoxes together is to fuse the position classes further so that we have a position class composed of \textit{gri}, \textit{mu}, \textit{ra}, and \textit{naxu}.

Examination of additional data through the table in (45) shows that attempting to maintain such a position class analysis will result in very few position classes. The reason for this is that there are many alternations involving multiple suffixes. If one assumes a certain order for the morphemes in a position class, this becomes apparent. For example, if one assumes the order \[ \text{[mu gri naxu]} \], then an alternation involving \textit{mu} and a morpheme outside of the current position class, for example \textit{ra} will necessitate broadening the position class to include \textit{ra} such that the position class is now \[ \text{[mu ra gri naxu]} \]. This is a persistent problem that reduces the number of position classes and in turn reduces their explanatory and restrictive power. Further examination of alternations in this way shows that a single position class will encompass the majority of morphemes, and thus the divides
like those made between affixes in Nimboran (Inkelas 1993) do not exist in IQ. A position class analysis, then, will not be able to distinguish between morphemes in such a way as to elucidate their interactions.

3.4 Semantic Scope

Rice (2000) makes the claim that semantic scope determines affix order in the Athabaskan verb and that the verb root in languages in that family moves to its surface position. She puts forth the idea that semantic scope is instrumental in determining affix order in other languages. Effects of semantic scope are attested in IQ. For example, the morphemes $chi_1$, $gri$, and $ra$ all participate in alternations that seem to show semantic scope effects. These effects can be seen for $chi_1$ in the examples below, repeated from (4) and (5).

(49) a. Kan kalpa-chi-wa-nga
   2SG run-CAUS-1OBJ-2SG
   "You make me run."

b. Kan kalpa-wa-chi-nga
   2SG run-1OBJ-CAUS-2SG
   "You make me run (= You’re turning me out / kicking me out / making me leave.)."

(50) a. Nyuka pay-ta kalpa-chi-xu-nga
   1SG 3SG-ACC run-CAUS-IMPFV-1SG
   "I am making him run / am kicking him out / am making him run (out of the house)."

b. Nyuka pay-ta kalpa-xu-chi-nga
   1SG 3SG-ACC run-IMPFV-CAUS-1SG
   "I am making him run / keep running."

In these examples, semantic scope seems to extend from the rightmost suffix leftward to the root. However, scope does not seem to operate uniformly this way, as shown below (repeated from (11), (12), and (13)), where instead root-adjacency seems to correlate directly with semantic prominence.

(51) a. Pishku-gu kashtu-gri-wa-nga
   bird-DIM bite-PROSP-1OBJ-DULC-3

b. Pishku-gu kashtu-gri-pa-nga
   bird-DIM bite-PROSP-DULC-1OBJ-3
   "The little bird is going to bite me (right now)."

(52) a. Pishku-gu kashtu-pa-gri-nga
   bird-DIM bite-DULC-PROSP-1OBJ-3

b. Pishkugu kashtu-wa-gri-pa-nga
   bird-DIM bite-1OBJ-PROSP-DULC-3
   "The little bird is going to bite me (soon)."

(53) Pishkugu kashtu-wa-pa-gri-nga
   bird-DIM bite-1OBJ-DULC-PROSP-3
   "The little bird is going to bite (sometime, now or later, it’s unknown)."
As explained in §2.2, semantic prominence (here, immediateness of the impending action) seems to be directly correlated with root-adjacency.

This means that semantic considerations influence affix ordering according to at least two independent principles: 1) semantic scope proper, where one affix seems to “scope over” another, influencing its interpretation, and 2) prominence by root-adjacency, where the closer an affix is to the root, the more influential it is as part of the verb complex’s overall meaning. Semantic scope proper, then, is only half of what must be explained when considering semantically-influenced affix ordering.

IQ, however, seems to exhibit non-semantically-driven affix ordering in which variation occurs without semantic effects. Where one might expect scope considerations to come into play, for example with the causative and reflexive/reciprocal morphemes chi and ri, one instead finds variation without semantic effects. Given the different types of semantic considerations that may come into play and the fact that semantics does not obligatorily influence affix ordering, it is clear that semantic scope, and even a broader notion of semantically-driven ordering, cannot fully explain affix ordering in IQ, even though semantic factors do explain some ordering tendencies.

### 3.5 Templates

Hyman (2002) shows that morphological templates are what determine affix ordering for at least a portion of the affixes in many Bantu languages (such as Chichewa), and that these templates must be strictly adhered to even in the face of semantic scope considerations that would seem to require a different affix ordering. Templates do play a role in IQ affix ordering. The template \texttt{ROOT}+...+\texttt{PERSON} must be adhered to in every form. Any additional (non-clitic) verbal morphemes that occur must be placed in the area indicated by the ellipsis.

However, strict templates do not seem to play a substantial role in IQ affix ordering beyond establishing the peripheries of verbal forms (that is, verbal forms not including clitics, which are invariably placed to the right). Affix ordering that varies according to semantic scope shows that templates do not govern those combinations of affixes. Affix ordering that is able to vary independently of semantic considerations also speaks against the presence of templates for the affected affix combinations. Templates, then, do play some role in IQ affix ordering, but templatic restrictions are unable to account for all the affix ordering phenomena shown in this study.

### 3.6 Bigram Morphotactics

Due to the high variability of affix ordering in IQ, the method of discovering bigram morphotactics described in Ryan (2010) is especially promising. This approach allows for variation to be described in a principled, quantitative manner that compares very favorably to data generated by plausible artificial learning paradigms, such as gradual learning algorithms. Bigram morphotactics would also directly capture the variation in ordering between pairs of affixes, which is a major source of variation in IQ affix ordering. While a bigram morphotactic analysis is extremely promising for IQ and would quantitatively capture the variation that is described here only in a qualitative manner, it requires a sufficiently large corpus of verb forms that can be parsed. While the Field Methods class built a corpus of over 15 narratives in IQ that were subsequently parsed, this is too small of a corpus for conducting a bigram morphotactic analysis.
4 A Construction Morphology Analysis of Affix Ordering in IQ

Construction Morphology (Booij 2009) offers a means of beginning to formally describe and relate verbal morphemes and their similarities with respect to affix ordering. The abstract schema below is the representation of a morphological construction in Construction Morphology.

\[(a_x \ [b]_y^\alpha)\]

This construction (adapted from Booij 2009:3(3')) is the representation of element \(y\), which is defined as \(y_i\) (represented by the morpheme \(b\)) with relation \(R\) to \(x\) (represented by the morpheme \(a\)). \(y_i\) is the head of \(y\). \(R\) is the relation between the elements in the construction, which “is left unspecified in the schema, since it is not predictable on structural grounds” (Booij 2009:1). \(\alpha\) is a relevant set of subclass features. Subclass features can be associated both with individual morphemes (like \(y_i\)) and constructions as a whole (like \(y\)). Constructions can be arranged into a hierarchy such that commonalities between them are inherited. The intent in using constructions is to formally indicate affix ordering restrictions and to group morphemes whose ordering restrictions are similar. I will first discuss the general shape that constructions take for IQ verbal morphemes, and will then go on to explore the arrangement of the constructions for each of the verbal morphemes into a hierarchy.

4.1 The Structure of IQ Morphological Constructions

Constructions for verbal morphemes in IQ take the following shape.

\[
[\text{ROOT}]_{V_1} [X, Y^{sem}]_a [\ldots]_b [Z]_c [\text{PERS}]_d \]

This construction is the representation of a verbal complex \(V\) for the morpheme \(X\), which is defined as a \(\text{ROOT}\) with relation \(R\), modified by \(sem\), to \(a, b, c,\) and \(d\). For IQ, the relation \(R\) is the theoretical mechanism responsible for producing a semantically interpretable structure. The subscripts \(a, b, c,\) and \(d\) are simply notational devices used to divorce the bracketed group from the phonological content inside. The \(\text{ROOT}\) is a lexical root, such as \(yanu-\) “cook.” \(X, Y,\) and \(Z\) are the verbal morphemes that this study has focused on. \(\text{PERS}\) represents person agreement morphemes. The position represented by an ellipsis (\(\ldots\)) instead of a letter is a position in the construction in which any member of the verbal morphemes examined in this study may appear. This position may appear in its own bracketed group or in another bracketed group. It is assumed that a very strict version of the repeated morph constraint (Menn and MacWhinney 1984) acts as a filter on this position such that morphemes which are already present in the construction are not inserted again. The diacritic feature \(sem\) signals that semantic interpretation is sensitive to the ordering of the morpheme to which it is attached. This feature is used to formally indicate that scope effects may occur.

For the purposes of this analysis, the modification of meaning which is accomplished by the ordering of a \(sem\)-bearing morpheme happens in the semantics of the language, which is represented here as the relation \(R\). This places the burden of accomplishing either a semantic scope or root-adjacency analysis on the semantics of the language. While this analysis does not provide an

---

13 While deserving a great deal of attention, the mechanism for producing a semantically interpretable structure will not be further explored here.
explanation of specific theoretical machinery, it does provide a structured input for such machinery to operate on.

In IQ morphological constructions as in (55), the ordering of elements matters. The order of the bracketed groups from left to right within the construction reflects the first to last ordering of those elements in spoken surface forms. Variability in ordering between affixes is modeled by putting the elements which vary in position together in the same bracketed group and separating the morphemes themselves with commas. These bracketed groups are position classes within the construction, but not across the language as a whole.

Each construction is meant to capture the affix ordering generalizations particular to a given morpheme. This given morpheme is positioned leftmost in one of the bracketed groups. In order to indicate which morpheme the construction pertains to, the morpheme is listed in the position of X in the V–X subscript on the entire construction. Here, this is strictly a notational device, but it could be used as a formal device to identify the construction uniquely, should the grammar require it.14 Note that here, the ROOT is always assumed to be the head of the construction and the element that bears the syntactic category feature V.

The construction indicates important facts that must be adhered to when the morpheme X is employed in a verbal construction. For example, the construction in (55) indicates the following about the morpheme X. First, the root must appear to the left of X. X may alternate in position with Y, but because of Y’s sem diacritic, alternations with Y will produce differences in meaning. If Z co-occurs with X, it must be immediately before person marking and must be to the right of X. Any other morphemes employed will go to the right of X, but to the left of Z. Person marking will be the rightmost element in the construction. The struck-out morpheme in braces to the right of the construction may not appear if X appears in the construction.

4.2 Building a Hierarchy of Constructions

Arranging morphological constructions into a hierarchy allows one to show that certain affix orderings are shared between constructions. This is important because it provides a degree of explanatory insight to the constructions that keeps them from just being convenient presentational devices. A construction hierarchy shows that constructions are related by the degree to which they share structure and by which structure they share. The construction hierarchy formalizes such similarity, and this formalization makes falsifiable predictions about which affix orderings are possible and which are not.15

A construction hierarchy for IQ morphemes can be established using certain principles. The highest node of the hierarchy contains the least information about affix ordering while the lowest nodes of the hierarchy contain the most information. In the construction hierarchy for IQ, the construction for the node X immediately above a given node Y is imposed on that node Y. The construction in node Y may add or delete material in bracketed groups or may add or delete bracketed groups themselves. When adding material, however, the construction in node Y may add only material that is not already present in the construction. That is, when adding material, the construction in node Y may only add material that was a part of the position represented by the ellipsis

14 For example, if a construction was associated with a particular phonological subgrammar, such as a different stress pattern, the label would be a formal device used to co-index the construction and its cophonology.

15 The motivation for a construction hierarchy is similar to the motivation for Partial Ordering (Anttila 2002), which helps model and explain relationships between cophonologies, which may share most of their constraints or rules.
The extent to which material is deleted is kept to a minimum because, ideally, affix ordering would be strictly more specified at lower levels by stipulating that material from a dominating construction could not be deleted.

4.3 The Hierarchy of Constructions for IQ Verbal Morphemes

Using those principles, it is possible to build a construction hierarchy for IQ in which constructions represented at lower levels in the hierarchy have more strict affix ordering requirements. I represent the hierarchy of constructions for IQ below in a tree structure in which each construction is represented by its namesake morpheme.

(56) Set of Verbal Morphemes

```
Template

FUT.PERS  rka
  |    |
ra  shka  ri  wa  xu  mu
  riya  pa  chi_1  naxu  chi_2  gri
          naya
```

The set of verbal morphemes is unordered and shown in (57).

(57) \[ \ldots = [chi_1, chi_2, gri, mu, naxu, naya, pa, pers, ra, ri, riya, rka, shka, \text{ROOT}_V, wa, xu, \text{FUT.PERS}] \]

The next level in the hierarchy is represented by the template common to all IQ verbal complexes. The category feature of \text{ROOT}_V percolates up and establishes that morpheme as the head of the construction and formally indicates that the construction characterizes an element of syntactic category V (verb).

(58) \[ \left[ \text{ROOT}_V \right] \ldots \left[ \text{PERS} \right] \]

Two branches split from the template construction. The first branch is that of the FUT.PERS construction. This construction is as follows.

(59) \[ \left[ \text{ROOT}_V \right] \ldots \left[ \text{FUT.PERS}^{sem}, ra \right] \]

The construction for ra can be derived from that in (59) using the principles in §4.2. This is shown below.

(60) \[ \left[ \text{ROOT}_V \right] \ldots \left[ \text{FUT.PERS}^{sem} \right] \]

\[ \left[ [chi_1, ra, \ldots, rka, shka] \right] \]

\[ \text{FUT} \left[ -ra \right] \]

\[ \text{PERS} \]

\[ \text{FUT} \left[ ra \right] \]

I assume that the position of the ellipsis may be moved freely as needed. This may need to be restricted later, but it seems necessary here.
The second branch from the template is that of the rka construction. This construction is the basis for all remaining constructions and takes the following form.

(61) \[
\text{[ROOT}_V\text{] }\ldots\text{[rka, shka] [PERS]}\]_{v^{-rka}}
\]

The construction for shka represents the first subbranch from rka, and takes the following form.

(62) \[
\text{[ROOT}_V\text{] }\ldots\text{[shka, pa, riya, rka] [PERS]}\]_{v^{-\text{shka}}} \{\text{FUT, PERS}\}
\]

This subbranch in turn gives rise to those for riya and pa given below in that order.

(63) \[
\text{[ROOT}_V\text{] }\text{[gri, naya, ra] }\ldots\text{[riya, shka, pa] [wa] [rka] [PERS, FUT.PERS]}\]_{v^{-\text{riya}}} \{\text{FUT, PERS}\}
\]

(64) \[
\text{[ROOT}_V\text{] }\text{[chi}, \text{mu, naxu, ri} \text{] [pa, \ldots, riya, shka] [rka] [PERS, FUT.PERS]}\]_{v^{-\text{pa}}}
\]

The second subbranch from rka is occupied by the construction for ri and its daughter construction for chi\textsubscript{1}, presented below in that order.

(65) \[
\text{[ROOT}_V\text{] }\text{[ri, chi}_1\text{, ra, riya, xu] }\ldots\text{[rka, shka] [PERS, FUT.PERS]}\]_{v^{-\text{ri}}}
\]

(66) \[
\text{[ROOT}_V\text{] }\text{[chi}_1\text{, ri, riya, mu}^{\text{sem}}, \text{wa}^{\text{sem}}, \text{xu}^{\text{sem}}] \ldots\text{[rka, shka] [PERS, FUT.PERS]}\]_{v^{-\text{chi}_1}}
\]

The third subbranch from rka is occupied by the construction for wa and its daughter construction for naxu, presented below in that order.

(67) \[
\text{[ROOT}_V\text{] }\text{[naxu, riya] [wa, \ldots, chi}_1^{\text{sem}] [rka, shka] [PERS, FUT.PERS]}\]_{v^{-\text{wa}}}
\]

(68) \[
\text{[ROOT}_V\text{] }\ldots\text{[naxu, riya, chi}_2\text{, gri, mu, ra] [wa, pa] [rka, shka]} [\text{PERS, FUT.PERS}]\]_{v^{-\text{naxu}}}
\]

The construction for xu, its daughter construction for chi\textsubscript{2}, and finally chi\textsubscript{2}’s daughter construction for naya, are all represented in the fourth subbranch of the construction for rka. The forms of these new constructions are as follows.

(69) \[
\text{[ROOT}_V\text{] }\text{[mu, ri] [xu, \ldots, chi}_1^{\text{sem}] [naxu] [rka, shka] [PERS, FUT.PERS]}\]_{v^{-\text{xu}}}
\]

(70) \[
\text{[ROOT}_V\text{] }\text{[mu, ri, \ldots] [chi}_2\text{, naya, pa, wa, xu, naxu] [riya] [rka, shka]} [\text{PERS, FUT.PERS}]\]_{v^{-\text{chi}_2}}
\]

(71) \[
\text{[ROOT}_V\text{] }\text{[mu, ri, \ldots] [naya, chi}_2\text{, pa, wa, xu] [naxu, riya] [rka, shka]} [\text{PERS, FUT.PERS}]\]_{v^{-\text{naya}}}
\]

The fifth and final subbranch of the construction for rka is occupied by the construction for mu and its daughter construction for gri.

(72) \[
\text{[ROOT}_V\text{] }\text{[ri] [mu, \ldots] [naya, pa, xu] [rka, shka] [PERS, FUT.PERS]}\]_{v^{-\text{mu}}} \{\text{FUT, PERS}\}
\]
5 Summary

This paper has sought to provide a descriptively adequate and formally-framed account of affix ordering in Imbabura Quichua. First, the data on affix ordering in IQ were summarized and synthesized. Second, this paper attempted to present the main arguments of prominent theories of affix ordering and show that although they are insightful, no single theory is entirely adequate for accounting for all the IQ data. Third, this paper presented an analysis of IQ affix ordering that was framed in terms of Construction Morphology (Booij 2009). That analysis defined one construction for each verbal morpheme and then grouped those constructions according to a set of principles into a construction hierarchy. This construction hierarchy was built on commonalities and subregularities shared among sets of affixes.

Appendix: Abbreviations & Affix Gloses

The Leipzig glossing rules were taken as the standard for the abbreviations used here. In the event that an abbreviation is not glossed, it is to be assumed that it follows those conventions.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset$</td>
<td>Unpronounced third person marking (unspecified for number or tense)</td>
</tr>
<tr>
<td>$chi_1$</td>
<td>CAUS</td>
</tr>
<tr>
<td>$chi_2$</td>
<td>PERSZR</td>
</tr>
<tr>
<td>gri</td>
<td>PROSP</td>
</tr>
<tr>
<td>gu</td>
<td>DIM</td>
</tr>
<tr>
<td>mu</td>
<td>TRANSLOC</td>
</tr>
<tr>
<td>$n$</td>
<td>3</td>
</tr>
<tr>
<td>naxu</td>
<td>JT.ACT</td>
</tr>
<tr>
<td>naya</td>
<td>DESID</td>
</tr>
<tr>
<td>nchi</td>
<td>1PL</td>
</tr>
<tr>
<td>nga</td>
<td>3.FUT</td>
</tr>
<tr>
<td>ngi</td>
<td>2SG</td>
</tr>
<tr>
<td>ngichi</td>
<td>2PL</td>
</tr>
<tr>
<td>ni</td>
<td>1SG</td>
</tr>
<tr>
<td>pa</td>
<td>DULC</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>Tag</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>ra</td>
<td>STILL Translates to English “still”; may not be fully productive</td>
</tr>
<tr>
<td>ri</td>
<td>REFL.RECIP Reflexive or Reciprocal; indicates that action of the verb is directed back at the subject</td>
</tr>
<tr>
<td>riya</td>
<td>DUR Durative (Cole 1982:184)</td>
</tr>
<tr>
<td>rka</td>
<td>PA Past tense, witnessed (Jessica Cleary-Kemp, p.c.)</td>
</tr>
<tr>
<td>sha</td>
<td>1SG.FUT 1st person singular future tense</td>
</tr>
<tr>
<td>shka</td>
<td>PA.NWIT Past tense, non-witnessed (Jessica Cleary-Kemp, p.c.)</td>
</tr>
<tr>
<td>shun</td>
<td>1PL.FUT 1st person plural future tense</td>
</tr>
<tr>
<td>ta</td>
<td>ACC Accusative</td>
</tr>
<tr>
<td>ka</td>
<td>TOP Topic marker</td>
</tr>
<tr>
<td>wa</td>
<td>1OBJ First person singular object marker; “me”</td>
</tr>
<tr>
<td>xu</td>
<td>IMPFV Imperfective aspect</td>
</tr>
</tbody>
</table>

**References**


Affix Ordering in Imbabura Quichua


John Sylak-Glassman
University of California, Berkeley
Department of Linguistics
1203 Dwinelle Hall
Berkeley, CA 94720-2650
sylak@berkeley.edu