DOCUMENT RESUME

ED 453 687 FL 026 749

AUTHOR	Weinberg, Jessica P., Ed.; O'Bryan, Erin L., Ed.; Moll, Laura A., Ed.; Haugan, Jason D., Ed.
TITLE	Coyote Papers: The University of Arizona Working Papers in Linguistics, Volume 11. Special Volume on Native American Languages.
INSTITUTION	Arizona Univ., Tucson. Dept. of Linguistics.
PUB DATE	2000-00-00
NOTE	99p.; Produced by the University of Arizona Linguistics Circle.
AVAILABLE FROM	Linguistics Circle, Department of Linguistics, Douglass 200E, University of Arizona, Tucson, AZ 85721 ($10.00 plus $2.00 shipping and handling).
PUB TYPE	Collected Works - General (020)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*American Indian Languages; Applied Linguistics; Child Language; Computational Linguistics; *Cross Cultural Studies; Discourse Analysis; Ethnography; Ethnology; *Language Acquisition; *Language Planning; Navajo; Phonology; Quechua; Research Methodology; Second Language Instruction; Second Language Learning; Uncommonly Taught Languages
IDENTIFIERS	Clitics; *Ethnolinguistics; *Mojave; Parsing; Yaqui

ABSTRACT

The five papers included in this volume approach the study of American Indian languages from a diverse array of methodological and theoretical approaches to linguistics. Two papers focus on approaches that come from the applied linguistics tradition, emphasizing ethnolinguistics and discourse analysis: Sonya Bird's paper "A Cross Cultural Look at Child-Stealing Witches" and Jessica P. Weinberg and Susan D. Penfield's paper "Mohave Language Planning: Where Has It Been and Where Should It Go from Here?" The other papers in the volume--"Child Acquisition of Navaho and Quechua Verb Complexes: Issues of Paradigm Learning" by Ellen Courtney and Muriel Saville-Troike; "Toward an OT Account of Yaqui Reduplication" by Jason D. Haugen; and "Critics, Scrambling and Parsing" by William Lewis--come from approaches in which the study of American Indian languages is increasingly providing important challenges to linguistic theory. References are appended at the end of each article. (KFT)

COYOTE PAPERS:
THE UNIVERSITY OF ARIZONA WORKING PAPERS IN LINGUISTICS

Volume II

2000

Special Volume on Native American Languages

edited by
Jessica P. Weinberg
Erin L. O'Bryan
Laura A. Moll
Jason D. Haugen

COYOTE PAPERS:
University of Arizona Working Papers in Linguistics

Special Volume on Native American Languages

Edited by

Jessica P. Weinberg

Erin L. O'Bryan

Laura A. Moll

Jason D. Haugen

2000

Volume 11

University of Arizona

Tucson, AZ

"Raven and Serpents" cover art by Amy Vlassia Margaris
Horned serpent motif adapted from Salado Polychrome design

Table of Contents

Introduction

The linguistic study of Native American languages includes a long tradition of ethnolinguistic and discourse analytic studies of Native American oral and written literature (e.g., Boas 1911, Tedlock 1983, Hymes 1981) as well as a more recent but rich attention to the applied sociology of Native American languages (e.g., McCarty and Zepeda 1998). Thanks in part to the efforts of Ken Hale (see Jelinek and Willie 2000), the study of Native American languages has also become increasingly important for contemporary linguistic theories. Given that the Americas are an area of extreme linguistic diversity (Nichols 1992), the linguistic study of the languages of the Americas provides a particularly important site for the cross-linguistic testing of linguistic theories.

The five papers we have included in this special volume of Coyote Papers approach the study of Native American languages from a diverse array of theoretical and methodological approaches to linguistics. Two of the papers come from approaches in which Native American languages are a familiar focus: ethnolinguistics/discourse analysis and the applied sociology of language. Sonya Bird's paper, "A cross-cultural look at child-stealing witches," makes a cross-linguistic and cross-cultural comparison among stories on a common theme in four Native American languages (Lummi, Sooke, Tohono O'odham, and Tlingit) and three non-Native American languages (Japanese, German, and Russian). Bird finds that while the stories are superficially similar in terms of content, deeper analysis reveals that they are in fact quite diverse in terms of both linguistic form and content. Jessica Weinberg and Susan Penfield's paper, "Mohave language planning: Where has it been and where should it go from here?," provides an overview of published linguistic research on Mohave, a Yuman language of Arizona and California, and an update of language planning efforts Mohaves have undertaken on behalf of their language. Weinberg and Penfield also make some recommendations for future Mohave language planning efforts that will take into account the unique situation faced by Mohave as one of four tribal languages of the Colorado River Indian Reservation.

The other three papers in the volume come from approaches in which the study of Native American languages is increasingly providing important challenges to linguistic theory: formal phonology, language acquisition, and computational linguistics. Jason Haugen's paper, "Toward an OT account of Yaqui reduplication," uses anchoring and alignment constraints within Optimality Theory to explain two reduplication patterns in Yaqui, a Uto-Aztecan language. Haugen argues that crucial to the analysis is the use of a right-edge anchoring constraint, which other theorists have argued is redundant within OT. Ellen Courtney and Muriel Saville-Troike's paper, "Child acquisition of Navajo and Quechua verb complexes: Issues of paradigm learning," uses language acquisition data from Navajo and Quechua as a test of Pinker's model of verb-inflectional paradigm learning. Courtney and Saville-Troike find in their data support for some of Pinker's predictions while their data contradicts other aspects of his model. William Lewis's paper, "Clitics, scrambling, and parsing," proposes algorithms for parsing linguistic structures that contain second position clitics, using data from three Native American languages (Quechua, Yaqui, and Lummi) and two non-Native American languages (Pakhto and Walpiri). This paper reflects the increasingly diverse areas of linguistics in which data from Native American languages are being utilized to advance linguistic theory.

References

Boas, Franz, ed. 1911. *Handbook of American Indian languages*, Bulletin 40, Part 1. Bureau of American Ethnology. Washington, DC: Government Printing Office.

Hymes, Dell. 1981. *"In vain I tried to tell you": Essays in Native American ethnopoetics.* Philadelphia: University of Pennsylvania Press.

Jelinek, Eloise and MaryAnn Willie, eds. 2000. Papers in honor of Ken Hale. *Endangered and Less Familiar Languages Working Papers* 1. MIT Working Papers in Linguistics. Cambridge, MA: MIT.

McCarty, Teresa and Ofelia Zepeda, eds. 1998. Indigenous language use and change in the Americas. *International Journal of the Sociology of Language* 132. Berlin and New York: Mouton de Gruyter.

Nichols, Johanna. 1992. *Linguistic diversity in space and time.* Chicago, IL: University of Chicago Press.

Tedlock, Dennis. 1983. *The spoken word and the work of interpretation.* Philadelphia: University of Pennsylvania Press.

A Cross-Cultural Look at Child-Stealing Witches*

Sonya Bird

1 Introduction

One of the important figures in Lummi mythology is Ch'eni, the Giant Woman (Ts'uXaelech) who comes during the night and steals children. When I first read the story of Ch'eni, I was struck by the similarity of this story to the well-known German tale by the Grimm brothers, 'Hansel and Gretel'. In fact, the story of Ch'eni is at first glance remarkably similar to several other children's tales in various cultures across the world. The goal of this paper is to explore the more subtle similarities and differences between the Lummi story and other stories in different cultures, in terms of the content of the discourse and the structure of the discourse used in the texts. We shall see that the Lummi story is in fact quite unique in its combination of elements of discourse content and structure. This makes the apparent similarity between it and other stories from around the world even more striking. Indeed, despite the numerous differences in terms of how the basic theme of the story is developed in Lummi and other cultures, the theme comes across clearly in all of the stories. This leads the reader (or listener) to mistakenly conclude that not only the main theme, but *all* aspects of the different stories are the same. The structure of the paper is as follows: in section 2, I outline the Lummi story of Ch'eni. In section 3, I discuss the content of this story, comparing it to that of /q'ɬəmáiəs/ in Sooke, Mosquito in Tlingit, Ho'ok in Tohono O'odham, Baba Yaga in Russian, Hansel and Gretel in German, and Yamamba in Japanese.[1] Finally, in section 4, I compare the discourse structure of the Lummi story to that in the other stories mentioned above.

2 Lummi's Ch'eni

In the Lummi culture, the story of Ch'eni used to be told to children so that they would keep quiet at night. This was necessary because of the constant threat of the Northern tribes, particularly the Haida and the Kwakiutl, who would come and raid the towns to acquire slaves. They would come during the night, and the villagers had to run into the woods to hide from them. Children had to remember to be very quiet, so as not to be found.[2] As in many Native American cultures, each Lummi family owns its own stories, or versions of stories. The version used for this paper was told by Mr. Al Charles, and recorded and transcribed by Richard Demers. A morpheme-by-morpheme translation was done by myself, using a dictionary by Demers (unpublished). The following is a summary of the story; for original Lummi version and loose English translation please see the Appendix.

Ch'eni is a giant witch who lives deep in the woods and steals children ("cry-babies" is the literal translation of the word used in the Lummi story) to eat. The way she does this is by luring them in with the promise of good food. She carries a basket with a big snake at the bottom of it, and when she comes across a cry-baby, she offers him or her some tasty food.

* Thanks to Richard Demers, Elizabeth Holaday, Laura Moll, and Nicole Theobald for useful discussion. This research was funded by the Social Sciences and Humanities Research Council of Canada, grant # 752-98-0274.

[1] Note: these stories are only a subset of those I found involving a female cannibal.

[2] Richard Demers, personal communication.

When the child accepts, Ch'eni stuffs him or her into her basket, and continues along her way. Eventually she arrives home with a bag full of cry-babies, and cooks them up for dinner. On the particular occasion related by the narrator, one of the children she captures is an older boy, who manages to escape by grabbing onto a tree. Then, when Ch'eni gets home, one of the cry-babies - a little girl - comes up with a plan to push Ch'eni herself into the fire. The children all agree, and thanks to careful planning they succeed in their plan. Ch'eni is consumed by the flames, and lice fly out of her, which are then transformed into a flock of birds. Thus, the children are saved and Ch'eni is never a threat again.

In the following two sections, I compare first the content and then the structure of the Lummi story with that of 6 other apparently similar stories from diverse cultures. Interestingly, in terms of content, the Lummi story is most similar to *non*-Native American stories. The similarities to other Native American stories arise when one looks at the structure and style of the discourse.

3 Content: A comparison of Lummi's Ch'eni to similar characters across the world

As mentioned above, the stories used here for comparison are from Sooke, Tlingit, Tohono O'odham, German, Russian, and Japanese. All of these cultures have stories involving either a cannibal or a witch, generally female, who eats children (usually baking them first). The stories differ in terms of the success – or lack thereof – of the witch in actually eating the children, who are generally the heroes of the story. In this section, I start with a summary of each story used for comparison. I then briefly discuss the general content of the various stories, as well as cultural values presented in them.

One thing to keep in mind here is that I am looking at single instantiations of each story, and I cannot be absolutely sure of the representativeness of these instantiations within the various traditions. Despite this drawback, the stories constitute an interesting data set, in that they clearly show some striking similarities and differences which exist between unrelated stories. The fact that some of the stories reflect the particular narrator as much as the culture makes these similarities and differences even more striking, since individual speakers in such diverse cultures clearly have not consulted each other on their work.

3.1 Summary of Stories Used for Comparison

All of the following stories were passed on orally at some point in time. They differ in how recently they were written down. Thus for example, the Lummi text used here was transcribed in the fall of 1998 by Richard Demers, whereas the Russian story of Baba Yaga is translated by Katya Arnold from a Russian story found in a collection of folktales collected and published in serial form between 1855 and 1864.

Sooke: /q'ɬəmáiəs/

The Sooke story used here is taken from Efrat (1969). The Sooke witch is called /q'ɬəmáiəs/, and she is basically the same character as Ch'eni.[3] She too deceives children by feeding them tasty food. She then brings them to her house and fastens their eyes shut with

[3] Sooke and Lummi are both in the Salish language family, and Sooke culture and Lummi culture are closely related. It is not surprising that they share this figure and her story.

pitch, so that they cannot see. Then she dances until the rocks are hot, at which point she cooks the children that she has stolen. The Sooke story ends here, with the narrator explaining that /q'ɬəmáiəs/ was the reason children were never allowed to wander alone into the woods. Thus, in this story, the witch is not killed, and the children are not saved.

Tlingit: Mosquito ("The Cannibal Giant")[4]

Unlike the characters in Sooke and Lummi, the Tlingit character is directly referred to as a cannibal, neither male nor female. The story used here is taken from Dauenhauer and Dauenhauer (1987). In this story, a young man's older brothers disappear. As he is searching for them, he encounters the cannibal, who strikes the young man on the head, puts him in a bag and brings him back to its house. The young man escapes, kills the cannibal (by striking it with its own club) and burns it. Interestingly, the ashes of the cannibal become mosquitoes, which still represent the cannibal when they suck blood from humans. This transformation is parallel to the one in Lummi, where lice come out of Ch'eni when she is engulfed in flames.

Tohono O'odham: Ho'ok

The story of Ho'ok is quite different from that of Ch'eni in terms of content. It was included here because Ho'ok is also a female figure who steals babies and small children – always boys - to eat (she also moves on to adults when children become sparse). In this story, taken from Evers (1980), we get the history behind Ho'ok. A young woman, Ho'ok's mother, took part in a ploy to hurt a young man whose older brother was jealous of him. The young brother was turned into a man-eating eagle, and the young woman became pregnant with Ho'ok. Because she was so mean, Ho'ok had to be sent away. She found a cave in which she made her home, and took daily runs into the villages demanding to see the babies. The mothers hid their children, but inevitably one would cry, and Ho'ok would stuff him or her into her burden basket, to eat later. Eventually, the people went to see I'itoi (a deity referred to as Elder Brother) for help, and he devised a plan to kill her. All the people participated in tricking her, and she ended up burning inside her cave. Unfortunately some smoke escaped the cave which transformed itself back into Ho'ok, in the form of a hawk. The hawk is also killed in a plan devised by I'itoi and carried out by the people.[5]

Russian: Baba Yaga

Baba Yaga is another female character, although not clearly human. Many Russian fairy tales involve Baba Yaga; the one discussed here (from Arnold 1993) involves an encounter between her and a little boy, Tishka. Baba Yaga captures Tishka by pretending to be his mother. She stuffs him into a sack and takes him to her hut in the forest, where she instructs her daughter to cook him for dinner. Tishka outsmarts the daughter, and Baba Yaga ends up eating her daughter instead of Tishka. Tishka gets away thanks to a friendly gosling, and Baba Yaga never bothers his family again.

German: Witch in Hansel and Gretel

It was the similarity of the Lummi story to Hansel and Gretel, a story already familiar to me, which prompted me to explore the topic of child-stealing witches. In this story by the Grimm brothers (Browne 1981), Hansel and Gretel are sent away by their stepmother. Lost and

[4] I also found this 'cannibal' figure in Haida.

[5] I'itoi also kills the man-eating eagle at the end of the story.

extremely hungry, they are wandering in the woods when they come across a house built of bread, cake and sugar, which is owned by a witch. She gains the children's trust by offering them food, and then captures them. She intends to cook Hansel, but the children outsmart her, and she is the one who ends up being baked. The children take the witch's pearls and jewels, and return home, by which time the stepmother has died; they live happily ever after with their father.

Japanese: Yamamba

In this story (Novák 1970), the child-hero is the 'Mountain Lad', who encounters Yamamba, "the terrible forest witch", on his way back home from town. The lad stops on his way to have a bite to eat. Yamamba, temporarily transformed into a hungry boy, begs the lad for some food. The lad gives him half his lunch, at which point the boy turns back into Yamamba and chases the lad into the forest. The lad manages to escape, but seeks refuge in a cabin which turns out to be none other than Yamamba's. Yamamba comes home, sad that she has let her evening meal slip away. The lad outwits Yamamba, and burns her in the boiler. He then goes out to bury Yamamba's previous victims, digs up a box full of gold in the process, and goes home, "whistling a merry tune".

3.2 General Similarities in the Content of the Stories

An entire paper could be devoted to the content of any one of these stories. In this section I briefly discuss only a few elements of the discourse found in the texts. We shall see that although the general story of Ch'eni seems remarkably similar to other stories from diverse cultures, a number of subtle differences are revealed upon a closer look at the texts. Before going into the comparison, I would like to emphasize once again that these texts are unrelated (except perhaps in the case of Sooke and Lummi); this is what makes the apparent similarity between them so fascinating. That their details are in fact quite different from one another makes this apparent similarity even more striking.

The general outline of the above stories is the same: a witch/cannibal, generally female, makes a habit of eating children. The hero - generally a child - encounters her (usually in the woods), outsmarts her, and she ends up the victim of her own devices.[6] This common main theme is what creates the impression that the stories are so close to each other.

The notion of cannibalism is found in all of these stories. This may represent the fact that throughout the world, the idea of eating another human is one of the worst things fathomable.[7] The fact that the monster is generally female makes sense, since she is meant to be terrifying: the only thing more horrifying than a child-eating monster is a *female* child-eating monster, at least given the Western stereotypes of female attributes: nurturing, protecting, etc.[8] Since these are all told to children, it is not surprising that the heroes are generally children.[9] In all cases except the

[6] The Sooke story deviates from this plot as noted above.

[7] Note that when the heroes kill the monster, they never eat it.

[8] It would be interesting to see whether such a monster exists in cultures where the male is mainly responsible for raising the children, and if so whether it is male or female.

[9] Lüthi (1970) and Bettelheim (1976) both offer Freudian analyses of Hansel and Gretel, in which they suggest that the evil monster is the mother, from whom the children must learn to break away, so that they can become independent people. The ending, where the witch is consumed in flames, represents the children's success in becoming independent from their mother. This analysis seems a little extreme as well as culture-specific, but it is another way of explaining why the witches are generally female, and the heroes children.

Tohono O'odham and Tlingit, the witch uses deceit to capture her victims. Thus, the stories may be partly intended as a lesson for children to be wary of strangers. All stories except the Sooke end with the witch's demise, which perhaps symbolizes the destruction of evil (Bettelheim 1976; Dauenhauer and Dauenhauer 1987). In the Lummi, German and Russian stories, she is consumed by the flames that are meant for baking the children. Finally, it is interesting to note that the stories generally take place in the depths of a forest. In my experience, forests are often perceived as dark, mysterious places, perfect for the occurrence of magical events!

Of course, there are variations, even in the basic theme of the stories: in the Sooke narrative nothing happens to the witch, she is simply introduced as a dangerous figure. In the Tlingit story, the hero is in his late teens, thus probably already considered an adult. Also in the Tlingit story, the cannibal is neither male nor female, and eats people regardless of their age. In Tohono O'odham, the hero is not an innocent child, but rather a very powerful figure: I'itoi. Note that all of these differences involve other Native American narratives. It is interesting that the Lummi story is actually closer in content to the German, Russian and Japanese stories than it is to the other Native American stories. The variation found is interesting, in that it may be a reflection not only of the different cultures involved, but also of individual narrators. Particularly in the Lummi, Sooke, Tlingit, and Tohono O'odham stories, which are more recently *oral* narratives, variation may be attributable to different story tellers. In Lummi, each family owns a version of a story, and takes pride its uniqueness.[10]

Regardless of its source, this variation adds support to the idea that although the same figures seem to reappear time and time again, the precise stories around them can in fact be quite different. As an example, take the Sooke and Lummi stories. Although the Sooke and the Lummi are different cultures, they are very close; Ch'eni (Lummi) and /q'ɬəmáiəs/ (Sooke) are arguably one and the same figure, with different names in the two languages.[11] However, the narratives about this figure (outlined above) are quite different, particularly in their ending. In the Lummi story, the witch is killed and transformed into lice, and then birds. In the Sooke version, nothing happens to the witch, and she remains a threat for little children who might be wandering in the woods. Similar variation is illustrated well in Dauenhauer and Dauenhauer (1987), where two different versions of the Tlingit story "The Woman Who Married the Bear" are told by different speakers. Although the story is the same, the details are quite different in the two versions. The variation typical of oral narratives is not found in written literature, where the details are fixed. The ending of Hansel and Gretel, for instance, is always happy, even in the original German version collected by the Grimm brothers (who are known for their rather brutal, non-happy endings).

3.3 Cultural Values Presented in the Stories

When one looks at cultural values, other interesting differences and similarities appear. Indeed, although all the stories share the main values (e.g. cannibalism is bad!), they differ in terms of the specific cultural values that are presented in the story. One interesting similarity between the Lummi, Tlingit and Tohono O'odham stories is the idea of transformation of one living thing into another, which appears towards the end of all three stories. In all cases, the transformation happens upon death. In the Lummi story, Ch'eni is transformed into lice and then birds; in Tlingit, the cannibal is turned into mosquitoes (hence the name of the story). Finally, in Tohono

[10] Richard Demers, personal communication.
[11] Richard Demers, personal communication.

O'odham, Ho'ok is transformed into a hawk before being killed. These transformations are in keeping with the belief among many Native American (as well as other) cultures that there is a cycle in life, in which living beings are transformed into other living beings. In terms of material values, it is interesting that the treat with which the witch entices the children differs according to the culture of the story. In Hansel and Gretel, the children are tempted by a house made of bread, cake, and sugar. In the Lummi text, the children are offered dried salmon, and in the Sooke story, they are fed "good fruit". This is a nice reflection of the dietary habits associated with each of these three cultures, which are obviously very different!

In all of the stories, the climactic events involving the cannibal occur far away from the community, usually somewhere deep in the woods. Following Hill (1995), one can associate this geographical distance with moral distance. The great distance between the hero's home and the home of the cannibal reflects the extreme immorality of cannibalism (found in all cultures discussed here).

Finally, the Lummi story is the only one in which the children are called cry-babies. This reflects the importance in the Lummi culture of keeping quiet. Indeed, when the Lummi had to flee from their villages into the woods to avoid being trapped in a raid, it was extremely important for the children to remain silent, so as not to give away their hiding place. Presumably, this was not as critical an issue for German children, for example.[12] Although in the Tohono O'odham story the children are not called cry-babies, it is those who cry who are found and taken by Ho'ok. Evidently, in the Tohono O'odham culture, crying is also considered a breach of social rules.

In this section, we have seen that although the general story of Ch'eni seems remarkably similar to other stories from various cultures across the world in terms of content, subtle differences are revealed upon a closer look of the texts, some of which reflect cultural values. The following table illustrates the similarities and differences discussed.

[12] It is worth noting that the Sooke story does not mention cry-babies, although the Sooke people presumably had the same problem with raids as the Lummi.

Table 1
Similarities and differences in the content of the Lummi and other stories

	Lummi similar to...	Lummi different from...
Nature of the witch/cannibal	Sooke, German, Russian, Japanese (Tohono O'odham)[13] ☞ all female witches	Tlingit ☞ cannibal not female
Nature of hero	German, Russian, Japanese, (Tlingit) ☞ all children	Tohono O'odham ☞ hero: I'itoi (not a child)
Use of deceit	Sooke, German, Russian, Japanese ☞ deceit used	Tlingit, Tohono O'odham ☞ no deceit used
Ending – demise of witch	German, Russian, Japanese, Tlingit, Tohono O'odham ☞ witch killed	Sooke ☞ witch not killed
Cultural values – Transformation	Tohono O'odham, Tlingit ☞ witch transformed into other living being	Sooke, German, Russian, Japanese ☞ no transformation
Cultural values – Treats	Sooke (fruit), German (bread, cake and sugar) ☞ treats offered	Tlingit, Tohono O'odham, Russian, Japanese ☞ no treats offered
Cultural values – Crying	Tohono O'odham ☞ crying important	Sooke, Tlingit, German, Russian, Japanese ☞ crying not mentioned

This table shows that in terms of the story development, the Lummi story is closer to the German, Russian and Japanese stories than it is to other Native American stories. In terms of cultural values presented in the story, Lummi is perhaps closer to other Native American cultures. Dauenhauer and Dauenhauer (1987) relate the Tlingit story to others found, among other places, in Siberia. Although I do not believe that there is any evidence that the stories discussed here are related, it would be exciting if they did in fact have a common origin. In the next section, we take a look at the discourse structure used in the Lummi story, and see how it too differs from that in the other stories outlined in section 3.

4 Discourse Structure: The Lummi story vs. others across the world

Due to space constraints, there are many aspects of the discourse structure that will not be considered here. What follows is a brief presentation of a select few elements of the discourse used in Lummi and other texts: 1) the use of repetition, 2) the responsibility assumed by the narrator as reflected in the use of evidential particles, possessives, and direct vs. indirect speech, 3) the power attributed to the witch and the children as reflected by agentive and passive constructions, and 4) the manner of presenting cultural values in the story. Section 4 will focus

[13] Parentheses indicate that the character is not quite as similar to the Lummi character as the others are.

on the characteristics of the Lummi text, and bring in examples from the other texts only when useful for comparison. In this section we shall see the discourse structure used also contributes (along with the content) to set the Lummi story apart from others. The differences in discourse structure discussed here bring us back again to the question of how these stories can appear so remarkably similar despite the numerous differences actually displayed.

4.1 The Use of Repetition

Most texts use repetition to a certain degree, either to emphasize a certain point, or to enhance the cohesion of the story. Two main kinds of repetition are discussed here, which I call immediate repetition and delayed repetition.

4.1.1 Immediate Repetition

Immediate repetition involves repeating some syntactic unit immediately after it is first spoken. This is found primarily in the Lummi, Sooke, and Tlingit narratives (more recently oral than the others), and is used mainly for emphasis. The following are two representative examples of immediate repetition.[14]

(1) *Immediate repetition in Lummi - line 15b*[15]

a. 'i' √seng-et-i-s tse heyi √mehoy,
accom √carry-ctrans-persis-3nom arg big √basket,
'then she would sling the big basket on her back,'

b. √seng-et-i-s tse heyi √mehoy.
√carry-ctrans-persis-3nom arg big √basket.
'she would sling the big basket on her back.'

(2) *Immediate repetition in Tlingit - Lines 50-54*[16]

Á áwé	That was
a daa yoo x̲'atula.atgi nooch,	what we would tell about
yá el'kaadéi haa wulgáas'i.	when we migrated to the coast.
A daa yoo x̲'atula.átgi nuch.	What we would tell about.
Ch'u a daa yóo x̲atula.átgi áyá.	What we would still tell about.

In the Lummi example (1), the entire main clause is repeated [sengetis tse heyi mehoy], whereas in the Tlingit example, the relative clause [a daa yoo x̲'atula.atgi nooch][17] is repeated. Sarris states that the effect of repeating things is to "to underscore a theme or idea" (1993:100). Thus, in the Lummi case, the image of Ch'eni moving along with her basket heaved up on her

[14] Throughout this paper, morpheme-by-morpheme translations of the examples are given only where available.
[15] √ refers to the verb root.
[16] x̲ is a uvular fricative.
[17] Note that the spelling is slightly different from one line to the next. I'm not sure if this is a reflection of the flexible spelling system, or whether there are actually slightly different meanings involved.

back is a very important one.[18] In fact, this image is a powerful one not only in the Lummi story, but in several others as well. Baba Yaga (Russian) stuffs Tishka into a "dirty canvas sack", Mosquito (Tlingit) puts the hero into a "sack" which it carries home on its back, and Ho'ok (Tohono O'odham) puts children in her "burden basket" and carries them away on her back. In the Tlingit example (2), emphasis is on the event of telling a story, that of the cannibals that lived in the Interior (of British Columbia). The use of repetition here emphasizes that what is to follow is a story. Thus it sets up the frame of the narrative, in the sense of Goffman (1974). I shall return to the idea of setting up a frame in section 4.2.1.

Intuitively, it seems that this use of immediate repetition as a way of emphasizing certain ideas has a practical purpose for the listener, who can quickly work out what the most important aspects of the story are. It also acts as a memory cue for the narrator. In written texts (German, Russian, and Japanese), immediate repetition is generally not found. This is plausibly because reading or writing a story is not as taxing in terms of processing as telling orally or listening to it. Neither the writer nor the reader needs to be reminded as much of the important facts involved. In the English translation of the Tohono O'odham story there is also no repetition, although this is likely to be a result of a fairly loose translation from the original narrative.

4.1.2 Delayed Repetition

The second kind of repetition is *not* an immediate repetition of a fixed phrase. Rather, in delayed repetition, a unit (either a theme (as in 3) or a syntactic unit (as in 4)) occurs at least twice throughout the text, and the occurrences are separated by intervening material. This type of repetition is used in oral narratives like Lummi as well as written texts like the German Hansel and Gretel. Delayed repetition has two roles: like immediate repetition, it emphasizes the important points of a story; in addition, it contributes to the cohesion of the narration (Sarris 1993). The following example from Hansel and Gretel illustrates these two roles:

(3) *German - the theme 'famine' as a case of delayed repetition*

a. first instance:
The family was always poor, and when a terrible famine came to the land, they could find nothing to eat.

b. second instance:
Not long afterward there was widespread famine again, ...

In the German case, we are looking at a repeated event: the famine. Here, repetition emphasizes the role of the famine, which is the reason the children are sent away not once, but twice. It also contributes to the slow build-up of the events that take place before the climax at the witch's house.

In the Lummi story, the unit repeated is a syntactic unit: the evidential particle *ch'e*. The following is an example of this.

[18] Note, this repetition may also indicate an iterative: she walked and walked with a basket on her back.

(4) *Lummi –ch'e as a case of delayed repetition, lines 11b, 25b, and 37a*

11b) √yae' **ch'e** tl'ael √ch'aenel tse nets'ae-xw
√go **dis** again √move arg one-person[19]
'then she would move, **I hear**, to the next person'

'i' tl'ael 'u' √Xena-et-s tse lae' √Xwe'ong.
accom again conn[20] √grab-ctrans-3nom arg there √crybaby
'and she did the same thing to the next crybaby.'

25b) **ch'e** √nilh s-u √tl'ew-n-onget tse s√wi'q-o'elh.
dis √it is then √notice-cmid-refl arg nom√man-dim
'**I hear**, he was lucky to notice (the growth), the boy.'

37a) nilh **ch'e** s-u√kwel-eng-s tse
it is **dis** then√fly out-cmid-3nom arg

√ngessens e se ts'u'Xaelech
√lice-3poss obl fem-arg Ts'uXaelech

'Then, **I hear**, the lice from Ts'uXaelech flew out'

As we shall see in the next section, this particle is used by the narrator to distance himself from the story. However, its occurrence throughout the text also adds to the coherence of the story. Throughout all the developments in the story, one thing remains common: *ch'e*. I shall not go into a detailed analysis of the use of *ch'e* here, except to say that given where it occurs in the text, its importance seems to be related more to the rhythm of the story than it is to the specific semantic information involved. Other particles which also contribute to the rhythm of the narrative are the connectors *i* ('and') and *su'* ('and then'). Like immediate repetition, delayed repetition involving a syntactic element is not generally found in written texts. Again, this is probably related to the processing difficulties associated with oral narratives, which require extra help to keep the story going. It is possibly also related to the need to distinguish oral storytelling from other oral events – i.e. to set up the "story" frame (Goffman 1974). This is not needed as much in written texts because there are other visual cues which set up the frame (pictures, for example) as well as linguistic constructions like "once upon a time".

To summarize this section, whereas delayed repetition of ideas and themes is common to all texts discussed here, delayed repetition of syntactic elements and immediate repetition only occur in the Native American texts. Thus, in terms of repetition, the Lummi discourse is more similar to other Native American stories than it is to the German, Russian and Japanese ones.

The use of *ch'e* in Lummi not only adds to the cohesion of the story (as seen above), it also reflects the responsibility assumed by the narrator with respect to the narrative. The responsibility of the narrator, and how it is expressed in different stories, is addressed in the next section.

[19] This word translates as 'the next one'.
[20] Together, *tl'ael 'u'* mean 'also'.

4.2 Responsibility of the Narrator with Respect to the Story

An interesting aspect of the Lummi story, when compared to the other texts discussed here, involves the responsibility assumed by the narrator in telling the story. This is reflected in the use of the evidential *ch'e* (4.2.1), possessives (4.2.2), and reported speech (4.2.3).

4.2.1 Use of the Evidential Particle ch'e

Throughout the Lummi text, the narrator uses the evidential particle *ch'e*, which is translated as 'reportedly' or 'so I hear'. The following example illustrates this:

(5) *The use of the evidential particle* ch'e - *line 12b*

'i'	heyi	**ch'e**	s√olhqe	e	tse	s√newe-lh
accom	big	**dis**	nom√snake	obl	arg	nom√to be inside-dur

e	tse	√mehoy-s	--	heyi	s√olhqe.
obl	arg	√basket-3sgposs	--	big	nom√snake.

'and, **I hear**, there was a big snake at the bottom of the basket – a big snake.'

The evidential particle indicates that Al Charles, the narrator, has no first hand experience with the story he is telling. This evidential sets up a particular frame (Goffman 1974), that of a story. The first point worth noting is that the only other story in which an evidential is found is the Tohono O'odham one. In Tohono O'odham, the evidential particle is *ṣ*. The following is an example of its use.

(6) *The use of the evidential particle in Tohono O'odham - line 1*

Ṣ 'am ki: g wiapoi kc s-melidag kc s-ke:g wuḍ 'o'odham kc 'am hab masma maṣ hab cum si 'i: ha-taicud g 'o'odham g 'e-'a'alga.

'There lived a youth, swift, handsome, who was all that parents want their children to be (the evidential particle is not translated into English).'

One could argue that the particle does not show up in any of the other languages simply because no such particle exists. However, all languages have other options in terms of expressing the meaning of the evidential particle. In English, for example, one could start sentences with "I hear" or "they say". The fact that none of the other languages uses these options is significant, as it represents different styles of narration. The most interesting case to look at is Sooke, which has the same evidential particle as in Lummi: *ch'*. The following is an example not taken from the narrative, from Efrat (1969)

(7) *The evidential particle in Sooke - from Efrat (1969)*

sq'^{w} á? - šn **č'** sə? n
companion - foot **evidential** future 1st sg. subject (sn)
'I'm going to be a walking companion, **I hear**.'

This particle is not used in the Sooke story. Given the similarity between Sooke and Lummi, and the existence of an existential marker (the same one) in both languages, the fact that it is used in the Lummi narrative but not in the Sooke version is quite striking. It is not clear whether this difference is a reflection of the individual narrators telling the stories, or whether it is a more general difference between cultures. The fact that *ch'* is absent in the Sooke story indicates that the story is more "real" in the eyes of the narrator. Indeed, no "story" frame is set up; as a result the narrator seems more closely involved in the recounted story. In fact, by Western genre conventions, the Sooke narrative is not a "story" in the same way the Lummi version is. There is no development of events leading to a climax, no happy ending in which the witch is killed, etc. Instead, the narrator simply introduces the witch as an explanation of why children were never allowed to walk alone in the forest. It would be worth investigating genre conventions in Sooke, and perhaps in Salish languages more generally, to see whether or not the Sooke narrative would fit into the same genre category as the Lummi narrative. One thing to note is that Al Charles, who told the Lummi story, was unlike some Lummi in that he did not believe in the existence of supernatural beings like the Lake Spirits, which many Lummi believe in. Thus, the fact that he distances himself from the story much more than the Sooke narrator does is perhaps a reflection of his personal beliefs in the magical, or lack thereof.[21]

As mentioned above, no equivalents of the evidential particle of the type "I hear" are used in languages which do not have such a particle. However, a more in depth analysis of the use of *ch'e* indicates that it does have an equivalent in other languages. English fairy tales for example traditionally start with the sentence "Once upon a time...". Thus, the English translation of the Russian story starts off with the following sentence:

(8) *English translation of Russian story - line 1*

Once upon a time there lived an old man and an old woman.

In Russian fairy tales, the traditional beginning is "Zhili bɨli..." (literally "there lived there were") which has the same connotation as "Once upon a time". Such expressions arguably have the same role as evidential markers in Native American stories; they too establish the "story" frame (Goffman 1974). Bauman (1992) talks about formulae like "Once upon a time" as elements which key the performance frame. Indeed, when an English speaker hears "Once upon a time" or a Russian speaker hears "Zhili bɨli...", he or she knows that what follows is a fairy tale. Thus, although evidential particles are not used in English or Russian, these languages also have specific resources that have the same role as evidentials. In Lummi as well as in English and Russian, the narrator is decreasing his or her responsibility for the story by distancing

[21] Richard Demers, personal communication. Demers also told me that Charles' sister believes in the supernatural much more than Charles did. It would be interesting to see whether *ch'e* would appear or not if she were to narrate the story.

himself or herself from it by using a formula which sets up the narration as a story. Thus, with respect to evidentials, Lummi is actually quite similar to other international stories as well as Native American ones. In the following sections, we see how the responsibility of the narrator is further decreased by the use of possessives and reported speech.

4.2.2 Responsibility of the Story as Reflected in the use of Possessives

Another way the narrator steps back from the story is by using a possessive (explicitly *not* referring to himself) when introducing the story. In the Lummi story, the narrator refers to the story as *Ch'eni*'s. This is shown in line 3 of the text:

(9) *Lummi - possessive used in introduction of the story - line 3*

√nilh tse s√t'aew'kw-lh xw√eyem-o-s-t-ong
it is arg nom√children-offspring loc√tell-res-purp-ctrans-pass
'It was to children that this story was told –'

e s-i'e s-xw√eyaem-s.
obl fem-part nom-loc√story-3poss
'it is **her** story.'

Saying that the story belongs to Ch'eni reinforces the fact that the narrator is simply transmitting a story he has heard; it is not his own story. Interestingly, in the Sooke narrative, the story is possessed not by the witch, but by the 'old people', as illustrated by line 1 of the story:

(10) *Sooke story - line 1*

/sx̲^{w}iém?s cə si?éiəx^{w}/
'This is a story of the old people.'

Whereas Lummi, in a sense, assigns the responsibility of the story to the witch herself, Sooke gives the responsibility of the story to the 'old people'. One could say that for the Sooke, the legitimacy of the story stands on collective authority. One thing to note here is that in the Lummi (but not in the Sooke) the possessive form used reflects who the main character in the story is. The fact that it is *Ch'eni*'s story does not mean that she owns the story, but rather that she is the main player (agent) in the events that take place. Thus, she is responsible in that she is the source of the story.

In terms of assigning responsibility of the story through the use of possessives, the Lummi and Sooke narratives are the only ones which explicitly say whose story it is. The use of a possessive is another means of distancing the narrator from the story. In the Lummi case, Al Charles is not responsible for the story, Ch'eni is.

4.2.3 The Use of Reported Speech

Besnier (1992) discusses the use of direct vs. indirect speech in the narration of stories. He follows Bakhtin (1981) in claiming that direct and indirect quotes differ in that direct quotes show "concern for the integrity and authenticity of the quoted message" (Besnier 1992:162), and keep the narrator and the quote maximally distinct from each other. Indirect quotes on the other hand allow for greater manipulation of meaning, since the voice of the narrator is present as well as the voice of the person who originally uttered the quote (Bakhtin 1981). In the Lummi story, there is an interesting dichotomy between the narration as a whole, which is an instance of indirect reported speech, and the few cases of direct reported speech used to recount what certain characters said. The text as a whole is indirect reported speech, which means that the narrator has a certain amount of freedom in choosing which details to include and how to tell them. The fact that the narrator can manipulate the text shows a closer involvement with the story. The use of indirect reported speech contrasts with the use of evidentials discussed in 4.2.1. While evidentials distance the narrator from the story, indirect reported speech brings him closer to it.

All stories discussed here involve reported speech. However, there is a clear difference between the Native American stories which are more recently oral, and the international stories, which have been written down for a longer time. In the Native American stories, the voice of the narrator in the reported speech is much more present, with each narrator telling his or her own version of the story (as discussed in section 3.2). That the German, Russian and Japanese stories have been passed on in written form for so long means that they are much more fixed, such that the use of reported speech is not as significant an issue. The narrator does not manipulate the story each time he tells it. Only the reader can manipulate the story to a certain extent with suprasegmentals like intonation and pauses.

Although the Lummi story is overall an instance of indirect reported speech, the narrator does use direct reported speech a few times. The following is an example:

(11) *Direct reported speech in Lummi - lines 9-10*

9) 'i' √kwen-t-i-s e 'es√Xaech s√chaenuxw --
accom √hold-ctrans-persis-3sgnom obl compl√dry nom√salmon
'She held some dried salmon --

s√q'ile -- su' √Xenae-t-s tse √Xwe'ong
nom√dried salmon nom-conn[22] √say-ctrans-3nom arg √crybaby
sq'ile -- and said to the crying child'

10) "√kwen e t-i'e 'ey' s√'ilhen."
√take obl exis good nom√eat
"Here – take this! It is good to eat."

The fact that the narrator is using direct speech indicates that although he holds a lot of responsibility in terms of the narration as a whole, he is not as strongly responsible for the actual

[22] These two morphemes together mean 'then'. *s-* is a the nominalizer, and *u* is a connector. The word sometimes appears with a glottal stop, and sometimes without. Throughout the rest of the text, this will be translated as 'then'.

utterances of the characters of the story. The use of direct reported speech is found in all of the stories discussed here. However, where it is most interesting is where it is in conflict with the more general narrative style. Thus, in Native American stories where the narrator is very much responsible for the story he is telling, the use of direct reported speech in reporting what the different characters say is quite striking. It reflects the fact that there is indeed some degree of independence between the behavior of the characters and the narrator.

In this section, we have discussed the responsibility of the narrator with respect to the story. We have seen that in all of the stories discussed here, specific formulae are used to distance the narrators from the stories they are telling. Other methods are used to diminish the responsibility of the narrator particularly in the Native American stories. In the next section we turn to the responsibility not of the narrator, but of the different characters in the story.

4.3 The Responsibility of Different Characters as Reflected in Syntactic Structures

Duranti discusses the relation between agency and responsibility, maintaining that "the attribution of agency to a party typically coincides with an implicit or explicit assignment of responsibility" (1990:644). In this section, I look at the responsibility assigned to different characters in the stories as reflected by the use of specific syntactic structures.

4.3.1 Agent vs. Patient Structures

In his chapter *Grammatical Parallelism and Thought*, Urban (1991) looks at Agent and Patient structures used in Shokleng and Bella Coola stories, and discusses how they reflect the power of the mythical entity involved in these stories. Urban defines the terms "agent-centric" and "patient-centric", where the former refers to stories where the agent is always the same but the patients differ, and the latter refers to stories where the patient is always the same but the agents differ. Whether the story is agent- or patient-centric determines which character of the story is the most important, and what his or her role is in terms of power. For the purposes of this analysis, I have focussed on the occurrence of the "control transitive" morpheme, which is used in Lummi to imply a patient object and an agent subject which exerts conscious control over the activity expressed in the predicate.[23]

In terms of content, the Lummi story is split into four sections. The first section (lines 1-4) introduces the story, the second section (lines 5-16) introduce the habitual actions of Ch'eni, the third section (lines 17-36) presents the particular events leading up to Ch'eni's demise, and finally the fourth section (lines 37-39) presents the consequences of Ch'eni's demise. In the first section, two Control Transitives (CTs) are used, both in constructions where the patient is the subject. Thus one could say that for this brief introduction, the story is patient-centered, the patient being the story. This makes sense given that it is the *story* that is being introduced. The second and third sections are more interesting. In the second section, 8 CTs are used, 7 of which have Ch'eni as the agent.[24] The patients of 4 of these CT clauses are different cry-babies. The fact that 7 of the 8 CTs have Ch'eni as the agent reflects her importance in this section of the

[23] See Montler (1986:163)

[24] The 8th is in an imperative construction, used by Ch'eni and directed at a crybaby: 'take this, it's good to eat'.

story.[25] She is the powerful character here, who is responsible for all the events that are taking place.

In the third section, Ch'eni's power - and her responsibility for the events recounted - is diminished as other agents are introduced. In this section there are 10 CTs, only one of which has the witch as its agent. Two of the CTs are passive constructions, in which there is no agent,[26] and one is an imperative. The remaining 6 CTs have children as their agents. Two of the 6 involve the older boy who gets away, one has as its agent the little girl who comes up with the plan to burn Ch'eni. The other 4 are used in the directives the little girl gives the other children; the agents of these 4 are the thus the other children (cry-babies). The third section can no longer be considered agent-centric, since there are several different agents; it is also not patient-centric since there is not one constant patient. What is crucial here though is that the story moves from agent-centric in the second section to non-agent-centric in the third section. This reflects the diminishing power of the witch in the third section of the story, which ends in her demise, the epitome of powerlessness. In contrast, the power of the children increases, as does their responsibility in the development of the story. This is reflected by their increased role as agents. In the fourth section, no CTs are used at all.

The Sooke story is agent-centric throughout (except in the introduction and the conclusion), with the focus – and all the power – assigned to the witch. All of the other stories are somewhat different from the Lummi, in that the focus is not so much on the witch as on the hero. This is reflected in the lack of a section in which the main agent is the witch. The fact that the witch in Lummi is the focus of attention, at least for part of the story, may reflect the "educational" nature of the story. Given that it was told to children to keep them quiet during raids, more emphasis on the witch was perhaps required to scare the children into obedience.

4.3.2 Passives

Another construction which involves agency - or lack thereof - is the passive construction. In the Lummi story *Ts'uXaelech*, the narrator uses passives in an interesting manner, reflecting the relative responsibility given the different characters in the story. There are two types of situation where Mr. Charles uses passives: 1) when introducing the story, and 2) when talking about cooking plans. In line 3, a passive is used in introducing the story:

(12) *Passives in the Lummi story - introduction (line 3)*

√nilh	tse	s√t'aew'kw-lh	xw√eyem-o-s-t-**ong**
it is	arg	nom√children-offspring	loc√tell-res-purp-ctrans-**pass**

'It was to children that this story was told...'

Here, the passive is used to place the emphasis on the children, who are the important players in terms of the purpose of the story.

[25] This case is different from those discussed by Urban (1991) in that here the patients are not all of the same type (they are not all cry-babies, they include salmon, basket, etc). In the Shokleng story of *The Giant Falcon* (Urban 1991:36), for example, the patients are all different falcons.

[26] The implied agent will be discussed in the next section (4.3.2).

In talking about cooking plans, the use of passives is somewhat more interesting.[27] In lines 29, 30a, and 35b passives are used:

(13) *Passives in the Lummi story – cooking, lines 29, 30a, and 30b*

29) √nilh s-u-xw√chqwe-t-eng-s tse s√tae'kw-lh,
√it is then-loc√bake-ctrans-pass-3nom arg nom√children-offspring
'Then the children were to be baked.'

30a) √qw'ele-t-eng,
√roast-ctrans-pass
'They were to be roasted,'

35b) 'i' √'enae √qw'eyelesh se ts'uXaelech
accom √come √dance fem-arg Ts'uXaelech
'and when Ts'uXaelech came in dancing,

su'√sXot-eng-s ch'e s-u'.
then√push-res-pass-3nom dis then
she was pushed (into the fire), I hear.'

It is interesting that in both 29 and 30a, the responsibility of cooking the children is taken away from the witch by the use of a passive. One can infer that she is the agent of the action, but this is not explicitly stated. In 35b, responsibility is again taken away from the agents - the children who are presumably responsible for pushing the witch into the fire. The diminished responsibility assigned the agents here is perhaps a result of the narrator not wanting to assign full responsibility for a deed as horrifying as roasting someone else. Even the witch is saved from such a responsibility. When one looks at other texts, they do not all show the same reluctance to assign agency. In the Russian story, "Tishka instantly shoved her onto the hot coals and slammed the oven doors together". In Hansel and Gretel, "(...) Gretel gave her a great push which sent her right in, banged shut the iron door, and locked it". In the Tlingit story, the hero also takes full responsibility, to the point where he asks himself "What more can I do to make it feel more pain?" In the Japanese and Tohono O'odham stories the agents also take full responsibility for their actions. Thus, it seems that if the Lummi narrator has any qualms about assigning responsibility for such horrific acts as roasting someone, other narrators do not share these qualms.

In summary, the Lummi text differs from others in terms of how responsibility is assigned to its characters. The Lummi and Sooke stories are the only ones in which the witch is presented as the most powerful figure for a good part of the story. However, the Lummi text is unique in using passives when talking both about the witch baking the children and the children baking the witch. This seems to reflect the hesitation of the narrator to assign responsibility for such a terrible act.

[27] Note that it is possible to use agentive constructions with cooking verbs in Lummi, thus the use of the passive construction is not simply a result of grammatical constraints in the language.

4.4 The Presentation of Cultural Values

Again for reasons of space, I will limit myself in this section to a brief discussion of the presentation of values in Lummi, Sooke, and Tohono O'odham. These three stories indicate that there are at least three different degrees of explicitness used when transmitting cultural values in narratives. Sooke illustrates the most explicit of these. The narrator overtly says that children should not walk alone in the woods. Lummi and Tohono O'odham illustrate less explicit means of transmitting values. In both cultures, refraining from crying is an important cultural value. In the Lummi story, this is expressed by calling Ch'eni's victims *cry-babies.* Although it is never explicitly said that crying is bad, the use of this term makes it quite clear that silence is an important virtue in the Lummi culture. The same fact is expressed in Tohono O'odham even less overtly. Here, it is done by saying that Ho'ok would come to villages, and those children who were found and taken away were those who cried. Again, one can infer from this that crying is frowned upon in Tohono O'odham culture. The reason I consider the Lummi presentation more explicit is the repeated use of the word "cry-baby", which has the same negative connotations in Lummi and English (Demers, personal communication). The use of this term is as effective, if not more, as the explicit explanation used in the Sooke narrative.

In the discussion on discourse structure, it was shown that Lummi shares certain characteristics with various stories, but is different from every one of the stories in terms of the overall structure. The following table summarizes the similarities and differences between the Lummi story and the others discussed here.

Table 2
Similarities and differences in the structure of the Lummi and other stories

	Lummi similar to...	**Lummi different from...**
Use of repetition	Sooke, Tlingit ☞ immediate repetition, delayed repetition	German, Russian, Japanese, ☞ no immediate repetition, delayed repetition only of ideas (not syntactic units)
Responsibility of narrator – evidential markers	Russian, Tohono O'odham, Tlingit ☞ use of *ch'e* or equivalent	Sooke ☞ no use of *ch'e* or equivalent
Responsibility of narrator – possessives	Sooke ☞ story is assigned "ownership"	
Responsibility of characters – Agent vs. Patient constructions	Sooke ☞ section where witch has the main role	
Responsibility of characters – Passives		All other stories ☞ no use of passives in the cooking acts
Presentation of moral values		Sooke, Tohono O'odham ☞ different degrees of explicitness

This shows that the Lummi story is slightly more similar to the other Native American stories than to the German, Russian, and Japanese ones. However, it is not as similar as one might expect, given that the Native American stories are all instances of oral literature (vs. the other stories which are part of a written literature). Indeed, one might think that stories can be split into two categories, corresponding to oral vs. written styles. According to the analysis presented above, it is not possible to make such a clear-cut distinction.

5 Conclusion

The goal of this paper was to compare and contrast the Lummi story of Ch'eni to other stories from around the world, stories which at first glance all seem very similar. The similarity in itself is astonishing given how different cultures are from one another. From the discussion in sections 3 and 4, one can say that the Lummi story, while sharing its basic theme with the other stories, is quite different from them. This uniqueness is created by the combination of specific discourse structure and content elements, a combination not found in any other one story.

Despite the big differences in structure and content between the Lummi text and the others discussed in this paper, the basic theme remains transparent, giving the lay-person the impression that all of these stories are „pretty much the same". Why is it that despite the differences these stories appear so similar? Clearly, one simple salient element suffices to give this impression. Thus, that all the stories are about a witch/cannibal who steals and eats children is enough to ensure that they are perceived as the same. The additional information presented through the use of culture-specific discourse elements is obviously quite subtle, and is revealed only upon a closer inspection of the texts.

References

Arnold, Katya. 1993. *Baba Yaga – A Russian folktale*. New York, NY: North-South Books Inc.

Bakhtin, Mikhail M. [1934] 1981. *The dialogic imagination*, ed. and trans. Caryl Emerson and Michael Hoquist. Austin: University of Texas Press.

Bauman, Richard. 1992. Disclaimers of performance. In *Responsibility and evidence in oral discourse*, ed J.H. Hill and J.T. Irvine, 182-196. Cambridge, MA: Cambridge University Press.

Besnier, Niko. 1992. Reported speech and affect on Nukulaelae Atoll. In *Responsibility and evidence in oral discourse*, ed J.H. Hill and J.T. Irvine, 161-181. Cambridge, MA: Cambridge University Press.

Bettelheim, Bruno. 1976. *The uses of enchantment – the meaning and importance of fairy tales*. New York: Alfred A. Knopf, Inc.

Browne, Anthony, illustrator. 1981. *Hansel and Gretel*. New York: Alfred A Knopf, Inc.

Dauenhauer, Nora, and Richard Dauenhauer. 1987. *Haa Shuká, Our Ancestors: Tlingit oral narratives*. Seattle: University of Washington Press.

Demers, Richard. 1984. *Dictionary of the Lummi language*. Ms., University of Arizona.

Duranti, Alessandro. 1990. Politics and grammar: Agency in Samoan political discourse. *American Ethnologist* 17:646-666.

Efrat, Barbara. 1969. *A grammar of non-particles in Sooke, a dialect of Straits Coast Salish.* Doctoral dissertation, University of Pennsylvania, Philadelphia.

Evers, Larry, ed. 1980. *The south corner of time*. Tucson: University of Arizona Press.

Goffman, Erving. 1974. *Frame analysis: An essay on the organization of experience.* New York: Harper and Row.

Hill, Jane. 1995. The voices of don Gabriel: Responsibility and self in a modern Mexicano narrative. In *The dialogic emergence of culture*, ed. Dennis Tedlock and Bruce Mannheim, 97-147. Urbana: University of Illinois Press.

Lüthi, Max. 1970. *Once upon a time – on the nature of fairy tales.* New York: Frederick Ungar Publishing Co.

Montler, Timothy. 1986. An outline of the morphology and phonology of Saanich, North Straits Salish. *Occasional Papers in Linguistics No 4*. University of Montana.

Novák, Miroslav. 1970. *Fairy tales from Japan.* Translated from Czech by Alice Denešová. London, England: The Hamlyn Publishing Group Ltd.

Sarris, Greg. 1993. *Keeping Slug Woman alive: A holistic approach to American Indian texts*. Berkeley: University of California Press.

Urban, Greg. 1991. *A discourse centered approach to culture: Native South American myths and rituals*. Austin: University of Texas Press.

Department of Linguistics
University of Arizona
200 E. Douglass Bldg.
Tucson, AZ 85721

sbird@u.arizona.edu

Appendix

Ts'uXaelech
Original Lummi text

The following text is split into numbered sentences. Each number corresponds to one full sentence[28] in the Lummi version of the text. Sometimes, one Lummi sentence includes more than one idea or thought. Where this is the case, each idea is represented by a letter. This story was transcribed by Dr. R. Demers from a tape. The transcription was done using the practical orthography adopted by the Lummi.

1. Yesostsense ti'e sxweyaem.
2. Nilh sxweyaems e kw kwel his sxweyaems, xweyaem.
3. Nilh tse st'aew'kwlh xweyemostong e si'e sxweyaems.
4. 'I' nilh snaes e tse slhaeni' kwse chenis.
5. Checholeqw su nilh.
6. Ho' kwel taengen 'I' e hae'saew' 'I' tl'aeng e Xwo'ong st'aew'kwlh.
7. Heyi tse mehoys
8. 'I' ho' Xwe'ong tse stae'kwlh 'I' en'ae nuelaeng.
9. 'I' kwentis e 'esXaech schaenuxw -- sq'ile – su' Xenaets tse Xwe'ong
10. "kwen e ti'e 'ey' s'ilhen."
11. a) Nilh e sukwenets 'I' newaes e tse mehoy tse Xwe'ong
    b) yae' ch'e tl'ael ch'aenel tse nets'aexw 'I' tl'ael 'u' Xenaets tse lae' Xwe'ong.
12. a) Nilh e su'newaes e tse mehoys
    b) 'I' heyi ch'e solhqe e tse snewelh e tse mehoys -- heyi s√olhqe.
13. a) 'I' tl'ael chaenel yae' haewel e tse netsae'xw
    b) 'I' tl'ael 'u' Xenaets tse lae' Xwe'ong.
14. "Kwenet tl'ae'. 'ey' s'ilhen."
15. a) Nilhch'e sukwensis 'I' nuwaes e tse mehoys
    b) 'I' sengetis tse heyi mehoy, sengetis tse heyi mehoy.
16. 'I' tl'ael ch'aenel yae' 'u' tu 'aeXeng ch'e, lets tse mehoys, lets e tse stae'kwlh, Xwe'ong stae'kwlh,
17. 'I' notse ch'e 'u' kwel 'esaeluxw 'I' 'u' Xchits kw enae'senae' e ts'uXaelech, ts'uXaelech.
18. nilh e su' kwenets tse smonech.
19. kw'ae'les tse smonech.
20. nilh suXenengs se ts'uXaelech,
21. " 'enae'. "
22. "kwenet tl'ae' 'ey' s'ilhen."
23. nilh e suyae's.
24. a) nilh sunuwaengs e tse mehoy,
    b) 'u' yae' shteng, 'I' shetengch'e yae' 'I' t'et'ekw'
    c) 'I' 'en'ae swi'qo'elh notse tse swi'qo'elh.
25. a) nilh e sutl'chilewaeleng tse ch'ae'seng 'I' kwentaeles
    b) ch'e nilh su tl'ewnonget tse swi'qo'elh.

[28] Sentences were defined in terms of punctuation. Each sentence ends in a full stop.

26. 'I' 'ew' s'I' Xchits se ts'u'Xaelech kwe Xenengs tse swi'qo'elh.
27. nilh suyaesyae tes e tse 'aelengs.
28. nilh sunewelengs, suheyi choqwos – heyi cheqwos.
29. nilh suxwchqwetengs tse stae'kwlh,
30. a) qw'eleteng,
    b) s'ilhens tse ts'uXaelech tse st'ae'kwlh.
31. heyich'e tse cheqwos.
32. 'I' nilh e suXenengs se slhincho'elh, yesost ch'e tse stae'kwlh.
33. a) ho' se kwel lo'st kw sqw'elengs se ts'uXaelech
    b) 'I' 'eswaechstsxw meqweits
    c) 'I' kw enae's'enae' 'I' qw'eyelesh se.
34. 'I' hose toXw 'u' kw enae' taechel e tl'e 'I' sXetlhelh, cheqwetlhelh su nilh kw enae' taechel.
35. a) 'I' kwel 'eslhael e tse st'ae'w'kwlh
    b) 'I' 'enae qw'eyelesh se ts'uXaelech su'sXotengs ch'e su'.
36. nilh ch'e su' cho'qws
37. a) nilh ch'e sukwelengs tse ngessens e se ts'u'Xaelech
    b) 'I' Xweneng 'I' tu Xweneng e ti'e qeyes.
38. ho' sxw shaeteng 'I' lenitsxw tse ngessen se ts'uXaelech.
39. txwts'its'ets'eng ch'e tse ngessensle.

The witch who stole cry-babies
English Translation from Lummi

I am going to tell a story. This story is an old one that I am telling. It was to children that this story was told – it is her story. Her name is Ch'eni, and she came from way back in the woods. She would emerge from the woods when it started to get dark, and she would come looking for children who were crying. She had a big basket. If the children were crying she would come into their house. She would hold some dried salmon – sq'ile – and say to the crying child,

"Here – take this! It is good to eat."

Then she would grab the crying child and put him in her basket. She would then move, I hear, to the next house and do the same thing again to the next cry-baby. She would put him into her basket. And, I hear, there was a big snake at the bottom of basket – a big snake. And she would move on and continue her trip to the next house. And she did the same thing to the crybaby there.

"Take this. It is good to eat."

I hear she would then grab the child's hand and put him in her basket. Then she would sling the big basket up on her back. And she kept moving along doing this, I hear, until the basket was full, full of children, crying children. And there was one, I hear, a bit older, and he knew Ts'uXaelech was coming. And then she took a pitch. She heated the pitch (to seal the children's eyes shut). The giant woman said to the older child:

"Come. Take this. It's good to eat."

She went on. Then the children were inside the basket, and she was walking on her way home, and the boy was coming along, the other boy (the older one). He took hold of a growth and got away. I hear he was lucky to notice the growth, the boy. Tsu'Xaelech did not know

what the (older) boy had done. She was walking along, and arrived at her home. She went inside. There was a big fire – a big fire. The children were to be baked. They were to be roasted, the food of Ts'uXaelech, the children. There was, I hear, a big fire. A little girl said, I hear she told the other children:

"When Ts'uXaelech is ready, such that she will cook, you will watch and wait, and when she is coming she will be dancing. When she gets here, we will push her and we will bake her, when she arrives."

The children got ready and when Ts'uXaelech came in dancing, I hear she was pushed into the fire. I hear she was burnt. Then, I hear, the lice from Ts'uXaelech flew out. That is how it was, and that is how it still is today. When you're out walking, you will see Ts'uXaelech's lice. The birds, I hear, are transformed from her former lice.

Child Acquisition of Navajo and Quechua Verb Complexes: Issues of Paradigm-Learning*

Ellen Courtney *Muriel Saville-Troike*

1 Introduction

1.1 Acquisition of Verb Morphology in Contrasting Indigenous Languages

Navajo and Quechua, both morphologically rich languages, present an interesting testing ground for proposals regarding the acquisition of inflectional systems. Of particular interest for these languages is the development of the verb, which encodes not only tense, aspect, and number-/person-of-subject, but also such grammatical notions as transitivity, causation, modification and internal arguments. In fact, the complex verb forms which characterize these languages often constitute the entire VP or, indeed, the whole sentence. The languages are all the more fascinating because, typologically, the structure of the verbs in Navajo is roughly the mirror image of Quechua verbs: while complex Navajo verbs are formed by appending prefixes to the root or stem, Quechua verbs are formed entirely through suffixation. This is illustrated in the Navajo and Quechua equivalents of the English sentence, 'They were feeding it to me, too':

(1) a. Navajo: Shí – áłdo' - shá - da - 'í - ø - ɬ - tsóód
me also for Pl obj subj CL impf stem:FEED

b. Quechua: Mikhu - chi - sha - wa - rqa - n - ku - pis
EAT Caus Prog 1 obj Past 3 subj Pl Add

In the Navajo verb form, the disjunct prefixes (those furthest from the stem) include oblique object, adverbial, postposition, and plural, and the conjunct prefixes (those closest to the stem), the direct object, subject, and transitive classifier. The stem 'feed' occurs in final position and its form indicates imperfective aspect. By contrast, in the Quechua verb shown in (1b), the stem created by affixing the causative suffix to the root means 'cause to eat' or 'feed', and the final element, the Additive suffix, is an independent enclitic meaning 'also.' The morphemes occurring between the causative stem and the final enclitic are part of the inflectional set.

In both languages, a verb must minimally consist of a root and a person-of-subject affix; that is, adult speakers do not produce bare verb roots or stems. However, in Navajo, the ordering of the prefixes in relation to the verb stem is quite rigid, whereas, in Quechua, some of the suffixes attached to the verb stem may occur in varied order. While the Quechua suffixes have unique and identifiable meanings, the rules for their combination often have no basis in semantics: they are idiosyncratic, including ordering restrictions which must be formulated as negative filters (Muysken 1986, Muysken 1988).

Recent work by several researchers has yielded relevant, isolated proposals, many intended to enlighten the nativist-empiricist debate. For example, Hyams proposed the stem

*Appreciation is extended to the Spencer Foundation and to P. Bloom, recipient of the grant, for partial funding of the 1996 fieldwork undertaken for the collection of the Quechua child language data.

parameter, with two possible settings: a verbal stem is/is not a well-formed word (Hyams 1986a, Hyams 1994, Pizzuto and Caselli 1994). She further claimed that children set this parameter very early on. Assuming that Hyams' parameter is a valid constraint on verb formation, we would expect children acquiring morphologically rich languages never to produce bare verb roots or stems. In a number of studies on the acquisition of synthetic languages, the perceptual salience of particular syllables is cited as a vital cue in segmentation (Pye 1983, Aksu-Koç and Slobin 1985, Mithun 1989), with the phonologically most salient morphemes occurring at the periphery of words. We would therefore expect children to produce the inflectional morphemes first, before the derivational affixes found closer to the verb stem (Peters 1995). On the way to meaningful productivity, children acquiring morphologically rich languages may at first produce frozen chunks or amalgams of affixes, even splicing together different unanalyzed strings to form novel, sometimes ungrammatical combinations (Peters 1985 and elsewhere, Franco and Landa 1998, Rubino and Pine 1998). Finally, it has often been noted that children learning a variety of agglutinative languages may insert novel filler syllables into the affixal string, (e.g., Aksu-Koç and Slobin 1985, Saville-Troike 1996). Peters (1985 and later work) views the insertion of such "placeholder" syllables as further evidence of unanalyzed amalgams in child production.

As illuminating as these studies may be, none presents a model for the acquisition of morphology as comprehensive as Pinker's (1984) proposal regarding the acquisition of inflectional systems through paradigm-learning. In the Introduction to the second edition (1996), Pinker later cites a number of studies confirming his original proposal. In developing the proposal, Pinker draws on observations of inflectional learning in a variety of languages, both his own and those of other researchers such as Slobin. This brings us to the aim of the present study: an exploration of the acquisition of verb morphology by children learning Navajo and Quechua in the context of Pinker's theory.

1.2 Pinker's Proposal

On Pinker's approach, children initiate paradigm-learning on the basis of whole words they have extracted from the speech stream. They infer the meanings of these verbs, and they exploit innate syntax-semantics mappings to assign them the category of VERB (contra Olguin and Tomasello 1993). Children then isolate the verb root/stem by matching meaningful stored verb forms which share the same roots or stems, so as to separate the shared portion from the inflectional "residue." They label the shared segment as ROOT, presumably an innate primitive which is available to the child as part of UG. Children's hypotheses as to the function and meaning of the inflectional residue are also innately constrained, since the closed set of grammatical features which are potential candidates for mapping onto the inflections are part of UG. The set of features may be ordered according to a grammaticizability hierarchy (Pinker 1996, Slobin 1985).

Pinker further proposes that children hypothesize one of these innately given "dimensions" (e.g., tense, person, number, aspect), and they construct whole-word paradigms for different "levels" of the hypothesized dimension. Eventually, they use the whole-word paradigm to construct the corresponding inflectional paradigm. An important additional outcome of inflectional paradigm construction is the creation of word-structure templates which specify the order of affixes. Throughout the development of paradigms, children adhere to the Unique Entry

Principle, which further restricts their hypotheses by disallowing multiple entries for a single paradigm cell (e.g., hypothesizing two different forms for first-person singular).

We present naturalistic child language data which shed light on the following aspects of Pinker's theory, either stipulated or inferrable from his discussion of paradigm-learning:

a. Children first isolate the verb root or stem. Children build their verb templates outward from the root/stem, since the more "accessible" notions tend to be affixed closer to the root/stem.
b. Children are unlikely to hypothesize certain dimensions at first, that is, those which are low on the accessibility (grammaticizability) hierarchy.
c. The Unique Entry Principle disallows multiple entries for a single paradigm cell. Therefore, paradigm-splitting, with the formation of new cells, is motivated by this principle.
d. The acquisition of an affix may be delayed if (1) the affix is homophonous with other affixes or (2) the affix occurs infrequently in the input.
e. Children's production of affixes is meaningful, even though children often overgeneralize the use of individual affixes to words that do not allow them in the adult language.
f. Paradigm-building proceeds from word-specific to generalized patterns of inflection.
g. The template developed by the child is rigidly ordered.

While our data largely support Pinker's theory, some of these proposals do not hold up in the light of Navajo and Quechua data.

2 Verb Morphology in the Two Languages: Highlights

2.1 Navajo

We first present a chart (Table 1) showing the classes of prefixes which may occur in Navajo complex verb forms. The Roman numerals indicate the positions of these prefixes with respect to the stem, which appears in Position X. The inflected verb is thus typically stem-final, except for a few enclitics which may follow the stem. Since there is little allowable variation in the ordering of these affixes, the chart represents a relatively rigid template. Sentences (2-4) illustrate some of this complexity, with different combinations of prefixes. The stem, while linearly indivisible, may vary according to aspect and mode, for instance. Morphophonemic changes sometimes obscure the transparency of the morphological composition of forms, raising questions regarding the isolability of the stem. This presents a complicating factor for Pinker's proposal insofar as the Navajo stem, unlike the Quechua stem, may vary, encoding more than the semantic reference. (The reader is referred to Young (1987) for further information about the Navajo verb complex.)

Table 1

The complex verb in Navajo

DISJUNCT PREFIXES	
0	Object of postposition
I	Postposition; Adverbial-Thematic
II	Iterative mode
III	Distributive plural
CONJUNCT PREFIXES	
IV	Direct object pronoun
V	Subject pronoun (usually fourth person; impersonal; indefinite)
VI	Thematic and adverbial elements
VII	Modal Conjugation marker
VIII	Subject pronoun
IX	Classifier
X	Stem
Enc	Enclitics (negative, relativizer, 'usually', etc.)

(2) akǫ - dii - sh - yeed 'I will go over there.'
I VI VIII stem
there inceptive I RUN

(3) sh - aa - d - oo - t - 'ááɬ '(It) will be given to me.'
0 I VI VII IX stem
me to incept prog classifier GIVE

(4) da - nighi - sh - chí 'She[1] gave birth to us.'
III IV VII stem
dist pl us perf GAVE BIRTH

The canonical order of major constituents in Navajo is SOV, or OSV with the object in focus or otherwise "outranking" the subject. (See Creamer 1974). The verb complex is almost always in sentence-final position.

2.2 Quechua

While the canonical order of major constituents in Quechua is also SOV, the word order in matrix clauses is flexible, with all six possible orders commonly produced by both adults and

[1] Neither Navajo nor Quechua indicates gender on pronominal affixes; gender in the English glosses is inferred from context. We use 'he' as the unmarked referent.

children. Quechua verbs differ from Navajo forms in exhibiting variation in the ordering of some of the affixes.[2] In (5a), from Bolivian Quechua (Cerrón-Palomino 1987:284), we find stylistic variation of the Assistive suffix with respect to the Inchoative. In (5b), the ordering of Causative *-chi-* with respect to Desiderative *-naya-* produces semantic effects, as revealed in the English glosses. Finally, some pairs of suffixes are rigidly ordered with respect to each other. As shown in (5c), for example, the causative always precedes the object-marking inflections. This combination is ambiguous, with the object marker representing either the lower SUBJECT, as in 'He makes me wash someone', or the lower OBJECT, as in 'He makes someone wash me.'

(5) a. Apa - **ri** - **ysi** - wa - y Apa - **ysi** - **ri** - wa - y 'Help me take it.'
TAKE Inch Asst 1 obj Imper TAKE Asst Inch 1 obj Imper

b. Yacha - **chi** - **naya** - wa - n 'I feel like causing (someone) to learn.'[3]
LEARN Caus Desid 1 obj 3 subj

Yacha - **naya** - **chi** - wa - n '(Something) makes me feel like learning.'
LEARN Desid Caus 1 obj 3 subj

c. Maqchhi - **chi** - **wa** - n. 'He makes me wash someone.'
WASH Caus 1 obj 3 subj 'He makes someone wash me.'

There are instances of homophonous affixes in both languages that are important in the following discussion of child data. As illustrated in (6a), in Chalhuancan Quechua, an allomorph of the Augmentative suffix, one of several modifier suffixes, is homophonous with the reflexive suffix. (The Augmentative allomorphs appear to occur in free variation.) In (6b), we find that allomorphs of the Exhortative suffix, another modifier, are homophonous with allomorphs of the Past Tense morpheme in some environments.

(6) a. Maqchhi - **ku** - sha - n 'He's washing himself.'
WASH Reflex Prog 3 subj

Tusu - **yku** - sha - n 'He's dancing intensely/decisively.'
DANCE Aug Prog 3 subj

Tusu - **ku** - sha - n 'He's dancing intensely/decisively.'
DANCE Aug/Reflex Prog 3 subj OR 'He's dancing by himself.'

b. Maqa - **r(q)a** - chi - n. 'He makes (someone) hit (someone)
HIT Exhort Caus 3 subj quickly/violently/suddenly.'

Maqa - chi - **r(q)a** - n. 'He made (someone) hit (someone).'
HIT Caus Past 3 subj

c. Tusu - na – **yki**. 'You are to dance.' Maqa - **yki**. 'I hit you.'
DANCE Pot 2 poss HIT 1 subj> 2 obj

[2] Abbreviations used in interlinear glosses are presented in the Appendix.
[3] Since the subject is third-person singular, a more literal gloss might be '(Something) produces in me the desire to cause (someone) to learn.' However, a literal English rendition is elusive here.

d. Maqa - wa - **nki**. 'You hit me.' Maqa - su - **nki**. 'He hits you.'
HIT 1 obj 2 subj HIT 3 subj>2 obj

The verbs in (6c) show that the suffix *-yki-*, which denotes first-person subject/second-person object, is homophonous with the second-person subject on nominalized verb forms. Finally, in (6d), we find that the final suffix, *-nki,* in *maqa-wa-nki* and *maqa-su-nki* represents the SUBJECT in one instance and the OBJECT in the second.

Now that we have reviewed the relevant facts about the Navajo and Quechua languages, we provide a brief description of our procedures for collecting naturalistic child language data in Navajo and Quechua.

3 Procedures for Data Collection

3.1 Navajo

The Navajo corpus for this study totaled 762 nonduplicate verbal utterances produced by five children from the Kayenta and Shonto regions of Arizona in the period from 1990 to 1992. The five children were recorded three or four times each during that period. In Table 2, the children, identified by pseudonyms, are listed in ascending order of proficiency as judged from (1) their Mean Length of Verb (MLV), or average number of morphemes produced in the verbal complex (a range of 1.0 to 3.9), (2) their total production of elements in each constituent position, and (3) their percentage of accurate production in each constituent position (as reported in Saville-Troike 1996). The table also shows the ages at which each child was recorded, ranging from 1;1 to 4;7.

Table 2

Navajo children by age at recording and in ascending order of proficiency

Nora	1;1 1;3 1;5			
Alice		2;11 3;1	3;6	4;0
Rose			3;6 3;8	4;0 4;7
Albert		2;10 2;11 3;0		
Lucy			3;5 3;6 3;7	

A total of approximately fifteen hours of audiotape recording was collected and transcribed. All recording was done by a female relative or friend of the family in a familiar setting with no non-Navajo or stranger present. The audiotapes were transcribed by three native speakers of Navajo who were from the same geographical region as the subjects. The corpus of

762 verbal utterances breaks down as follows: Nora, 25; Alice, 134; Rose, 204; Albert, 215; and Lucy, 184.

3.2 Quechua

Five children from Chalhuanca, an Andean community in the Caylloma Province of Arequipa, Peru, were each recorded in Chalhuanca for an approximate total of five to six hours from June through the beginning of October, 1996. A sixth child, Ana, was recorded for eleven hours from June through December, 1996. The age range for the children, as shown on Table 3, is 2;0 to 3;9. All the recordings were carried out in the homes of the children or in a one-room home/daycare facility, with the researchers sometimes present and usually with older siblings or cousins. The audiotapes were transcribed by native speakers of the Cuzco-Collao variety.

The child corpus for the present analysis included 1062 verb forms: Max, 81; Ana, 314; Hilda, 184; Ines, 253; Juan, 116 (only verbs with two or more internal suffixes); Tomas, 114 (only verbs with two or more internal suffixes). In addition, a total of 279 adult verb forms were taken from child-directed speech as a basis of comparison and for the analysis of input frequency. To facilitate the analysis, the corpus of 314 verbs produced by Ana was divided into three sub-groups according to age: Ana: 2;5-2;6 (89 verbs), Ana: 2;6-2;7 (109 verbs), and Ana: 2;9-2;10 (116 verbs).

Table 3

Quechua-learning children, by pseudonym and age at recording

Max	2;0 2;1 2;2						
Ana		2;5 2;6	2;7 2;8	2;9 2;10			
Hilda				2;10	2;11 3;0 3;1		
Juan					3;0	3;2	
Ines						3;2 3;3 3;4 3;5	
Tomas							3;7 3;8 3;9

We now turn, one by one, to the previously mentioned aspects of Pinker's proposal. In the ensuing discussion of each aspect, we present relevant data from the Navajo and Quechua child corpora and discuss whether the data support or challenge the predictions.

4 Discussion of Data in the Context of Pinker's Proposal

4.1 Children first isolate the verb root or stem.

While Pinker does not propose that children actually PRODUCE the verb root or stem first, we find for both Navajo and Quechua that children's early verb forms often consist of uninflected roots or stems (contra Hyams). This is surprising, since these children never hear such forms in the input.

The youngest Navajo child, Nora, produced the bare stems shown in (7-8) below at ages 1;1 and 1;3. Target adult forms are included in parentheses. Although vowel length, tone, and nasalization are phonemic in adult Navajo, no differentiation of these features was detectable in Nora's speech. Indeed, they were not systematically realized even in the production of the older children in this sample. The bare stem *teeh* was used for homophonous utterances in (8), meaning both 'Lie down' and 'Pick me up'. (Interpretation of the meanings in context was provided by Nora's mother.) The third utterance in (8), a two-word, four-syllable utterance occurring at the same age as consistently bare stems in verb constructions, indicates that Nora's productive constraints in the verb were not simply due to length of utterance.

(7)	Nora (1;1)	Da.	'Sit.'	(nídaah.)
		Go.	'Come here.'	(hágo.)
(8)	Nora (1;3)	Teeh.	'Lie down.'	(níteeh.)
		Teeh.	'Pick me up.'	(náshidiilteeh.)
		Haagi gaagi?	'Where is the crow?'	(háájí gáagi?)

The youngest Quechua speakers also produced bare roots and stems. The examples in (9) show Max, at ages 2;0 to 2;1, producing a partial root in the first utterance, and an uninflected reflexive stem in the second. At ages 2;5 to 2;6, Ana's sentences are three or four words in length, and yet she persists in often producing bare roots. In (10), the bare root is *muna-* 'want' for both second-person *qan* and first-person *noqa.*

(9)	Max (2;0-2;1)		
	Mu__ chicha.	'I want chichasara.'	(munani chichasarata)
	Noqa laqaku__.	'Mine fell down.'	(noqaq laq'akun)
(10)	Ana (2;5-2;6)		
	Chay muna__ qan.	'You want that.'	(chayta munanki qan)
	Carruta muna__ noqa.	'I want the car.'	(noqa carruta munani)
	Noqa carru muna__.		
	Mana muna__ noqa carrupi.	'I don't want it in the car.'	
			(noqa carrupi mana munanichu)

One might speculate that these children are simply dropping the final consonants as children often do universally (Hoff-Ginsberg 1997). However, in the Navajo examples, this would necessarily apply only to verb stems; rather, it is the PREFIXES that are missing. It should also be mentioned that the Quechua examples are not isolated instances. Of the total of 81 non-duplicated verb forms produced by Max, 46, or 57%, are bare roots or stems. At ages 2;5 to 2;6, Ana produces such forms in 27% of her verb forms. She stops producing such forms at age 2;7, and the older children do not produce any bare roots at all.

4.2 Children build their verb templates outward from the root or stem.

In the hierarchy of accessibility (grammaticizability) briefly discussed by Pinker, the most accessible notions are those grammatical encodings which are the most prevalent in the languages of the world. In a footnote, Pinker further proposes that the most accessible notions, following Bybee (1985), are affixed closer to the root or stem. Children would hypothesize these notions first, moving from the root or stem outward, so that they would never have to reanalyze the developing template in order to insert an affix symbol. The prediction challenges the role of perceptual salience; that is, if perceptual salience is a key determinant in the emergence of verbal affixes, we would expect children first to produce those morphemes occurring at the margins of the verb complex. Our data challenge the proposal that children build their verb templates outward from the root or stem.

Navajo thematic and modal prefixes appear in Positions I and VI in the verb complex, for instance, and the meaning of these prefixes is often opaque. Following Pinker, we would expect children to start off hypothesizing the notion encoded in Position VI, which is closer to the stem. The most proficient pair of children, Albert and Lucy, produced verb forms with these prefixes in both positions. Example utterances produced by these two children are shown in (11) and (12), with the appropriate adult forms indicated underneath the glosses.

(11) Albert (2;11)

Nda.	'(Pretend) he sits down (in the back of the car).'
	(ni- [Pos VI] = terminative: ńdaah.)
Da **dil**tal.	'He's dashing off.'
	(adi- [Pos VI] = relates to arms and legs: dah diiltaaɬ.)

(12) Lucy (3;5)

Biłch'e' **ya**diilte.	'Let's talk to this.'
	(ya- [Pos I] = relates to speech: bich'į'yádiiltih.)
Hanéék.	'Is (the blood about to) come out?'
	(ha- [Pos I] = flows out: haanéék.)

By contrast, Alice's and Rose's productions of the Position I thematic prefix were far more reliable than those of Position VI. Alice produced 80% of the prefixes appropriately for Position I and only 25% for Position VI (a medial conjunct position), WHICH IS CLOSER TO THE STEM.

Rose's errors and omissions were confined to Position VI. The order of accuracy in production clearly relates to prefix position (and thus to perceptual salience)--and NOT to proximity to the stem. The data thus refute Pinker's proposal.

There is further reason for dismissing this proposal: the most accessible notions cannot be generalized as occurring closer to the stem because there is no universal ordering of affixes. In Navajo, the subject prefix is closer to the stem than the object morpheme, while in Quechua, the opposite is true: it is the object-marking suffix that is closer to the stem. Thus, on strictly logical grounds, the proposal fails, since the sequencing of the argument-marking affixes is different in the two languages. Also, although the Navajo child language data do not reveal an acquisitional sequence for these affixes, we find that children learning Quechua produce the outer, person-of-subject suffixes before the inner, object-marking morphology emerges.

4.3 At first, children are unlikely to hypothesize dimensions which are low on the accessibility hierarchy.

The previously discussed accessibility hierarchy derives from Greenberg's (1963) linguistic universals, such that the most accessible notions are those which are the most universal. Pinker draws from Slobin in making this proposal, and he specifically mentions Navajo as having inflectional elements that are relatively inaccessible. More precisely, Pinker proposes that the probability that a child would hypothesize a notion encoding number is higher than the probability that she or he would hypothesize a notion encoding shape.

The Navajo data support this proposal. For example, the most proficient child, Lucy, was the only one who used some roots that are highly specialized for shape or type of movement, e.g., 'move hands abruptly', 'scratch', 'handle object in an open vessel', and 'flat flexible object (such as blood or water) moves freely'. Rose (3;8), on the other hand, relied on the generic *'á* meaning 'handle single round object' instead of *tí* meaning 'move animate object'. This is shown in (13).

(13) Ch'iidoo**'aa**. '(The doctor) will take out (the baby).' (ch'íidool**tééł**.)

Unlike Navajo, Quechua does not exhibit inflectional elements which might be considered highly inaccessible in comparison to others. For this reason, corresponding Quechua data are not available.

4.4 The Unique Entry Principle disallows multiple entries for a single paradigm cell. Paradigm-splitting, with the formation of new cells, is motivated by the Unique Entry Principle.

So far, we have dealt with Pinker's predictions regarding the order children follow in isolating or hypothesizing different parts of the verb complex. We now turn to the mechanism Pinker proposes for the development of inflectional paradigms, that is, the UNIQUE ENTRY PRINCIPLE. In general terms, this principle implies that children seek unique form/function correspondences, and we find evidence for this in the Navajo child production of the direct and oblique object prefixes.

Table 4 shows a list of the direct object Position IV prefixes in Navajo, as well as the object of postposition Position 0 prefixes. The selection of *bi-* versus *yi-* for third person object involves rules of agreement and topicality. Essentially, *yi-* occurs in object position when the

subject is also third person and the subject is topical; *bi-* occurs when the subject is not third person or the object is in focus, as in "passive" constructions. (See Thompson (1996) for a discussion of this alternation.)

Table 4

Direct object Position IV and object of postposition Position 0 prefixes

Direct Object Position IV		Object of Postposition Position 0	
sh-	1st	sh-	1st
nihi-	1st plural	ni-	2nd
bi-/yi-	3rd	bi-/yi-	3rd
'a-	indefinite	ał-	reciprocal
'ahi-	reciprocal 'each other'		

In Table 5, we present a summary of the accuracy attained by each of the four oldest children in the production of these prefixes. The children are presented in ascending order of proficiency. The plus signs (+) indicate consistently accurate production, while the X's represent only partially accurate production. A blank space means that the child did not produce the prefix.

Table 5

Accuracy of four oldest children in production of Navajo IV and 0 prefixes

	sh-	nihi-	yi-	'a-	'ahi	sh-	bi/yi	ni-	ał-
Alice			+	+		+	X		
Rose			+	+		+	X		
Albert			+	+		+	+	+	X
Lucy	+	+	+	+	X	+	+	+	+

Of particular interest here is the production of the third-person *bi-* and *yi-* prefixes in Positions IV and 0. Note that all the children accurately produced the *yi-* prefix in Position IV (no obligatory *bi-* context occurs in these data); 87% of those produced in Position 0 were *bi-*. The less proficient pair of children, Alice and Rose, frequently substituted *bi-* in Position 0 contexts requiring *yi-*. For these children, the two prefixes pattern in complementary distribution, with *yi-* for Position IV and *bi-* for Position 0. It is as if the children were imposing one-to-one form/function correspondences on these affixes. Examples of these *bi-* and *yi-* errors and the target forms are presented in (14) and (15). The unique form/function patterns the children have developed may also relate to input frequency of occurrence.

(14) Alice (3;6)

Ale'e ***bi** naane.	'He is playing with something.'
	(t'áadoo le'é **yee** naané.)
Wagon **yii** da sida.	'He is sitting in the wagon.'
	(VERB: **yii**' dah sidá.)

(15) Rose (4;0)

Alice ***bee** naane.	'He is making Alice cry.'
	(VERB: **yee** naané.)
Ye neeneegii.	'The one he is playing with.'
	(**yee** naanéhígíí.)

A related phenomenon, to be discussed later, comes up in the Quechua data: the children appear to impose mutual exclusivity on suffixes which are not mutually exclusive in adult production.

The Unique Entry mechanism for paradigm-building is especially interesting in light of some of the earlier Quechua-language data. A case in point is the pattern of production in Ana's first-person singular verbs. The pattern changes dramatically from the first age range, 2;5 to 2;6, to the next age range, 2;7 to 2;8. This is shown in Table 6. We see from the first two rows of the table that Ana produced a total of 38 first-person utterances during the first age range. Of these, fully 36 included explicit pronoun subjects. These 38 utterances included 19 forms inflected for first-person future, only 2 with the first-person present suffix, and 17 forms that were either bare roots or inflected in third-person. In the next age range, shown in the third and fourth rows of the table, Ana produces 58 first-person utterances, with the number of explicit pronouns sharply reduced. Now, over half of the verb forms bear the first-person present suffix, and only 4 verbs exhibit missing or incomplete inflections.

Table 6

Ana's first-person singular verb utterances

First-person subject inflection		PRESENT -ni	FUTURE -saq	INCOMPLETE -Ø or -n
2;5 - 2;6	[+Pronoun Subj]	1	18	17
	[-Pronoun Subj]	1	1	--
2;7 - 2;8	[+Pronoun Subj]	6	7	2
	[-Pronoun Subj]	24	17	2

We feel that Pinker's Unique Entry Principle explains these data quite well. In Figure 1, we present a sequence of gradually developing paradigms for the root *puri-* 'to go or walk' with a view to accounting for the observed patterns in Ana's early first-person verb forms. Ana may start off with the undifferentiated cell shown under (a), hypothesizing, at first, that the third-person inflection simply encodes the notion of SUBJECT for all persons. Thereafter, even though future forms such as *purisaq* 'I will go' may briefly occur in free variation with

incompletely inflected forms, the Unique Entry Principle forces a split in the cell. Under (b), the paradigm now represents the newly hypothesized dimension of PERSON-OF-SUBJECT. In the next step, the child notices that *puri-ni* 'I go' also occurs in first-person subject environments. The Unique Entry Principle then forces another split in the paradigm, such that, in (c), the PERSON-OF-SUBJECT dimension now has two levels, FUTURE TENSE and PRESENT TENSE. Eventually, in this way, as shown in (d), the full paradigm for Person-of-Subject is developed.

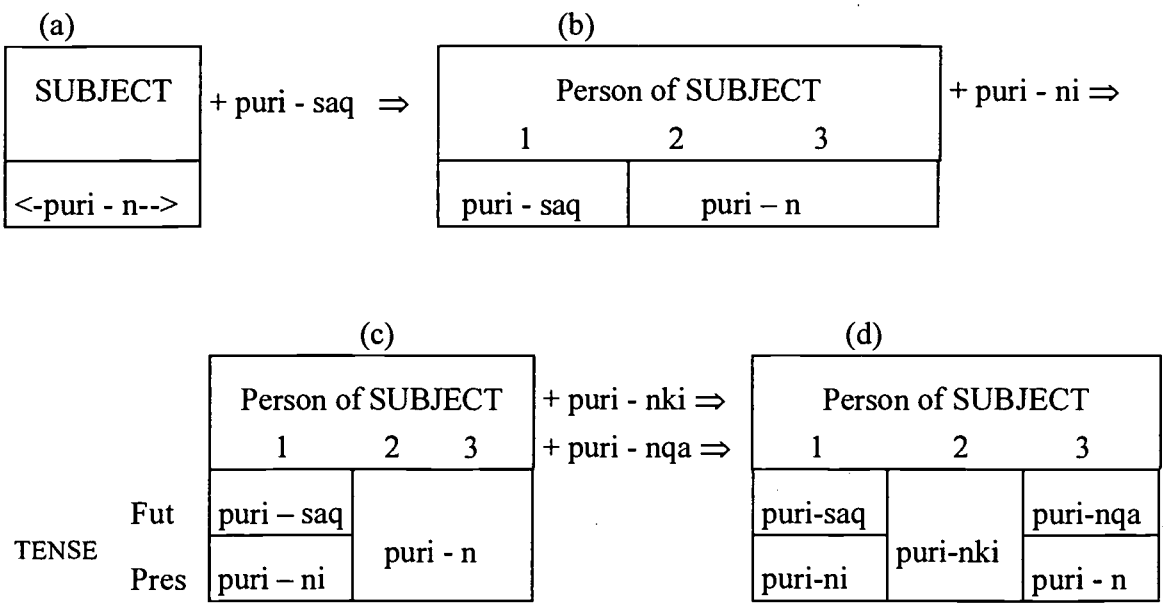

Figure 1

Development of person-of-subject paradigm in present and future singular

4.5 The acquisition of an affix may be delayed if it (1) is homophonous with other affixes or (2) occurs infrequently in the input.

The Navajo progressive prefix *yi-* in Position VII is homophonous with other prefixes, including the previously presented third-person object prefixes and the "peg elements", which are inserted if there is no other syllable preceding the stem. As shown in (16), Albert produced an appropriate *oo-* variant of the progressive but not *yi-*. In (17), Rose has omitted the progressive prefix in an obligatory context (the only obligatory context for this prefix in her first 50 non-duplicate utterances), and Alice does not produce the prefix at all. Homophony could be an explanatory factor in the late appearance of this morpheme.

(16) Albert (2;11)
Bił oolwoł. 'He is riding along.'
(lit. 'The vehicle is moving with him.')

(17) Rose (3;8)
Na#lt'o'. 'It is sucking it.'
(Omits yi- progressive: yiłt'o'.)

The form *yi-* also functions as a perfective prefix with some stems, as does *si-* with others; the *si-* prefix emerges first, even though *yi-* occurs more frequently in adult Navajo. Clearly, input frequency does not determine the order of emergence of these morphemes in child language.

Moving on to the Quechua data, we find that adults very frequently produce verb forms with object markers: these occur in 26.8% of the adult complex verb forms (i.e., those verbs with at least one non-subject suffix appended to the root). Table 7 presents a summary of the object-marking suffixes produced by three of the children. At the top of each column, the numbers indicate the person-of-subject and person-of-object encoded in the suffixes. For example, in the first of these columns, "3 arrow 1" means third-person singular subject, first-person singular object. Note in the last column that only the oldest of the three children produced any verb forms bearing the first-person subject/second-person object suffix *-yki.* This suffix is frequent in the child-directed adult speech, but it is homophonous with the second-person possessor suffix. It may be that homophony delays the PRODUCTION, if not the UNDERSTANDING of the object-marking suffixes.

Table 7

Quechua object markers produced by three children

	3>1 -wa-n	3pl>1 -wa-n-ku	2>1 -wa-nki	2pl>1 -wa-nki-chis	3>2 -su-nki	1>2 -yki
Ana: 2;5-2;6	--	--	--	--	--	--
Ana: 2;7-2;8	x	--	x	--	x	--
Ana: 2;9-2;10	x	--	x	--	x	--
Hilda: 2;10-3;1	x	x	x	--	x	--
Ines: 3;2-3;5	x	x	x	x	x	x

In his complex verb forms, the youngest child, Max, most frequently produces combinations with reflexive *-ku-*, progressive *-sha-*, and augmentative *-yku-*. These suffixes, as well as first-person future *-saq,* are the most frequent suffixes for Ana during the earliest age range (2;5 to 2;6). The very early emergence of both reflexive *-ku-* and augmentative *-yku-* is puzzling, considering that the reflexive is homophonous with an allomorph of the augmentative. It is possible that young Quechua speakers may meld together the semantic notions of intensity and personal involvement, expressed by means of the augmentative and the reflexive, respectively.[4]

[4] It is intriguing, in this regard, that both Juan and Tomas sometimes produce the augmentative suffix where the reflexive is required. For example, instead of constructing the appropriate reflexive intransitives, *phiña-ku-* 'get angry' and *silla-ku-* 'mount', Tomas produced the forms, *phiña-yku-shan* and *silla-yku-shan.* In both forms, the augmentative suffix is substituted for the required reflexive. Juan produces three variants of a verb form intended to mean 'he is mounting': *silla-ku-sha-n, silla-yu-ku-shan,* and *silla-yku-shan.* The apparent interchangeability of the two suffixes, or fusion of meanings, may partially account for their early appearance in child verb production, even

Overall, of the 249 adult complex verbs, 21% are progressive forms and 27% are future forms. With respect to the modifier suffixes, exhortative *-rqo-* occurs more frequently than augmentative *-yku-* (22% and 10%, respectively). Table 8 lists the most frequently occurring contiguous combinations in the corpus of adult complex verb forms, in descending order of frequency. Conspicuously missing from this list are the suffix pairings, *-yku-sha-* and *-ku-sha-*, that is, the augmentative and the reflexive in combination with the progressive. Although these combinations are relatively infrequent in the adult corpus (3.2% and 1.6%, respectively), they are among the most frequent contiguous combinations produced by all the children. The table reveals that exhortative *-rqo-* figures prominently among the most frequent suffix pairings, and yet, only the oldest children, Juan (3;0-3;2), Ines (3;2-3;5) and Tomas (3;7-3;9), produce exhortative combinations with any frequency. These production facts are curious, since the augmentative and the exhortative suffixes are probably equally accessible modifiers on the accessibility hierarchy.

One might speculate that homophony among the allomorphs of the exhortative and the past tense is a factor here; however, we also find homophony in the reflexive and augmentative suffixes that appear early and frequently in the children's verbs. The fact that the *-rqo-* variant of the exhortative straddles two syllables may be another factor. Of the 15 exhortative verbs produced by Ana, the first 10 all bear the one-syllable allomorph *-ru-*. In fact, Ana does not produce a verb form with the *-rqo-* allomorph until age 2;9.

Table 8

Most frequent contiguous suffix combinations in adult complex verbs

-rqo-mu-				10.0%
-rqo-FUTURE				9.0%
-mu-FUTURE				7.0%
-pu-FUTURE				4.8%
-rqo-ku-	-mu-OBJECT			4.4%
-ku-mu-	-pu-OBJECT	-ri-ku-	-chi-ku-	4.0%

4.6 Children's production of affixes is meaningful, even though children often overgeneralize the use of individual affixes.

Whether or not production is meaningful is not easy to determine from an analysis of naturalistic speech corpora, but there is one example in the Quechua data that suggests that children's production is not always semantically driven. In examples (18-23), we find a set of utterances produced by Ana at ages 2;5 to 2;6. Ana's utterances, which are on the right, are all responses to the direct questions provided on the left (IL refers to the interlocutor). In all of these utterances, Ana's verbs end in the combination *-ku-sha-*, which comprises the augmentative allomorph *-ku-* with progressive *-sha-*. (As mentioned previously, the suffix pairings, *-yku-sha-* and *-ku-sha-*, are among the most frequent contiguous combinations produced by all the children.) If her responses seem inappropriate in the English glosses, they seem even more so in the original

though they are often homophonous.

Quechua, since Quechua speakers characteristically respond to direct questions by repeating the exact verb stem provided in the question. Example (19) is particularly revealing. Here, the question contains a verb with the resultative suffix *-sqa*. In her response, Ana produces an ill-formed verb containing the root, part of the resultative suffix, and *-ku-sha-n*. Since Ana clearly overuses this suffix pairing, she may be assigning an idiosyncratic non-adult meaning to the combination. She might also be attaching it to verbs because [root + two syllables] forms sound better to her; that is, she may have noticed that the adults around her more frequently produce multisyllabic verb forms. Significantly, in the next age range, Ana stops producing this combination.

(18) IL: Mamayki tusunchu tokakaqtin?
'When it is played, does your mom dance?'
Ana: *Tokan. Tusu-kusha_.
'It plays. She is AUG-dancing.'

(19) IL: Ñachu chayasqa?
'Is it already cooked?'
Ana: *Chayas-kushan.
'It is AUG-cooking.'

(20) IL: Pukllanchu?
'Does she play?'
Ana: *Puklla-kushan.
'She is AUG-playing.'

(21) IL: Imata ruwashani?
'What am I doing?'
Ana: *Kuchu-kushan.
'It is AUG-cutting.'

(22) IL: Noqachu sipirusaq?
'Shall I kill it?'
Ana: **Tu* sipi-kushan.
'You (SP) *is AUG-killing it.'

(23) IL: Haytakusqasunki?
'Had he kicked you?'
Ana: *Hayta-kushan.
'He is AUG-kicking.'

There are intriguing examples of non-meaningful production in child Navajo verb forms, as shown in (24-25). The children sometimes produce verbs with novel prefixes, that is, Navajo-sounding nonsense fillers. Both of the following utterances were produced by Rose at age 4;0:

(24) Money **nanajaa** leh. 'He usually gives me money.'
(Verb should be: shaa néijih ɬeh.)

(25) Líí' **nana**zi. 'It is standing by the horse.'
(ɬíí' yíighahgi sizį́.)

Perhaps, following Peters (1985, 1995), children produce such fillers because they have acquired a phonological template for the verb form without having fully analyzed the individual affixes in the string. However, it is intriguing in this regard that the only Quechua-speaking child who produced such verb forms was Ines (3;2 to 3;5), the most proficient of the six children:

(26) Regala - **a** - **a** - wa - n - mi. 'He has given (it) to me.'
Give ? ? 1 obj 3 subj Aff

(27) Chura - **a** - **a** - wa - n - mi. 'She has put it on me.'

Put ? ? 1 obj 3 subj Aff

Given this child's level of competence, it is clear that she has acquired the individual suffixes as well as suffix combinations. The inserted nonsense morphemes are probably filling in for suffixes that Ines is not able to assemble during production because of processing load.

4.7 Paradigm-building proceeds from word-specific to generalized patterns of inflection.

Pinker claims that the developmental evidence supports a progression from word-specific to general paradigms because children's first inflections are confined to only a very small subset of words. The Navajo and Quechua data suggest that this is the case for some affixes but not for others.

Navajo children, for example, do appear to learn some inflections in relation to specific lexical items. The Navajo utterances provided below suggest that the stem shape change for aspect and mode, the thematic and modal prefixes, and the modal-conjugation marker are all learned in relation to specific roots. In contrast to these inflections, the Navajo argument prefixes appear on all stems at once, as evidence for generalized learning.

4.7.1 Stem shape change for aspect and mode

Albert (2;10-2;11) voiced the final consonant for perfective forms of 'áázh 'two subjects go', but not for gizh 'cut':

(28)	Diit'ash.	'(Now) let's go (outside).'	(Future: 'ash)
(29)	Ch'iniit'a**z**.	'(Pretend that) we went out.'	(Perfective: 'áá**zh**)
(30)	K'ijiłgeesh.	'When one is cutting (the sheep).'	(Momentaneous/ Imperfective: géésh)
(31)	Iniłge**sh**.	'I cut it off.'	(k'ínígi**zh**.)

4.7.2 Thematic and modal prefixes

With one exception, Alice (3;6) categorically either produced or omitted the thematic prefix associated with a particular stem multiple times without variable occurrence. In other words, a thematic prefix either always occurred with a stem (e.g., *daa-* with 'die') or was always omitted (e.g., *ni-* with 'look'), strongly suggesting that these are lexically-linked elements.

(32)	Bee **daa**tsa.	'It died with it (my father's gun).'	(bee **daa**ztsą́.)
(33)	**Daa**tsa.	'It died.'	(**daa**ztsą́.)
(34)	#łí.	'Look at (something).'	(Omits prefix: **nił**'į́.)
(35)	#shí.	'I'm looking for it.'	(Omits prefix **ní**sh'į́.)

4.7.3 Modal-conjugation marker

Alice's overall 50% accuracy rate in this position was also far from random: she always used *si-* appropriately with 'sit', for instance, but never used it with 'die'.

We find a similar pattern in Quechua, where suffixes such as regressive *-pu-* first appear on specific lexical items, while the exhortative modifier appears on a variety of stems right from the outset. This is shown in Table 9, which provides a list of the verb stems produced by the three youngest children with four selected suffixes: the exhortative, causative *-chi-*, directional *-mu-*, and regressive *-pu-*. As seen in the last column, the children's earliest production of regressive *-pu-* is confined to just a single stem, *qo-pu-*, meaning 'give back', while the exhortative appears on a variety of roots. Also note in the second half of the last column, that the suffix *-pu-* eventually starts occurring very frequently in combination with the object marker *-wa-*. Here, *-pu-* is a benefactive suffix, so that the combination *-pu-wa-* means 'for me' or 'for my sake.'

Table 9

First verbs produced by three youngest children bearing four selected suffixes

	Exhortative -rqo-	Causative -chi-	Directional -mu-	Regressive -pu-
Max (2;0-2;2)			apa-mu- 'bring' ayku-mu-'enter'	qo-pu- 'give back'
Ana (2;5-2;6)	taka-ru- urma-ru-	lava-chi-	apa-mu- ayku-mu-	
Ana (2;7-2;8)	pika-ru- (2) wayk'u-ru- (2) chinka-ru- puri-ru- t'oqo-ru- urma-ra-	toka-chi- (2) 'play (recorder)'	apa-mu- aysa-mu-	qo-pu- (5) INCLUDING: qo-[pu-wa] (2)
Ana (2;9-2;10)	sipi-rqu (2) hisp'a-ra muyu-rqa ayku-ru	14 verbs/10 roots INCLUDING: toka-chi (3) qhawa-chi (3)	apa-mu- (6) qechu-mu- (2) saqe-mu- chaya-mu- qo-mu- muyu-[rqa-mu]	qo-pu- (2) apa-ka-pu- (2) pasa-pu- 'go away' ranti-[pu-wa]
Hilda (2;10-3;1)	wañu-ru (2) tuku-ru tuku-rqu ayku-ra rikch'a-ra wikch'a-ra ?qo-wa-rqu	24 verbs/14 roots	apa-mu (2) qhawa-mu ayku-[ra-mu]	pasa-pu (2) mikhu-naya-pu chura-[pu-wa] saka-[pu-wa] ruwa-ri-[pu-wa] toka-chi-[pu-wa]

Children learning Quechua may at first confine their production of suffixes to a limited set of roots. However, they also appear to append a limited set of suffixes and suffix combinations to a wide variety of roots--even novel roots, e.g., *kometa-shan* (from Spanish *cometa* 'kite'). This is illustrated in Table 10, which shows the most common suffixes and suffix combinations produced by Ana at ages 2;5 to 2;6. Ana produced a total of 30 [ROOT + 1] verb forms, 25 of which involved only the four suffixes shown, but nearly every verb had a different root. The same four suffixes, in different pairings, account for 19 out of the 22 two-suffix combinations in Ana's corpus. Of the 4 three-suffix combinations produced by Ana, 3 utilize these same suffixes. Thus, 84% (47) of the 56 complex verbs used by Ana in this age range are formed with only four suffixes.

Table 10

Frequent suffixes and their combinations in Ana's complex verbs (2;5-2;6)

Stem Type	-yku- (Aug)	-ku- (Refl)	-sha- (Prog)	FUTURE (1 singular)	Number of Verbs	Number of Roots	% all Verbs
root +1			x		16	14	
	x				4	3	
				-saq	3	3	
		x			2	2	25/30 83%
root +2	x		x		7	6	
		x	x		5	5	
	x			-saq	4	4	
		x		-saq	3	3	19/22 86%
root +3	x	x		-saq	2	2	
	x	x	x		1	1	3/4 75%
TOTAL: 47 verbs (84%)							

4.8 The template developed by the child is rigidly ordered.

Quite remarkably, none of the Navajo children ever made any errors in the sequencing of prefixes within the verb complex. There was not a single instance of inverted order among prefixes in the production of any of the children; the ordering of constituent positions within the inflected verb was inviolate. Even for Quechua, which allows some flexibility in suffix ordering as well as repetition of suffixes, the children in Chalhuanca showed rigid ordering in the initial template. In fact, for the younger children, some suffixes were mutually exclusive which are not so in adult forms. Figure 2 below represents the early child template. For the youngest children, the suffixes presented in boldface within each column appear to be mutually exclusive. Reflexive *-ku-* and the object marker *-wa-* are in bold letters because they are at first mutually exclusive with respect to each other. At first, children do not repeat elements in the verb complex, nor do they vary the sequence of morphemes.

STEM	Modifier	Causative	Reflexive	Direction	Progress	Object	Tense/ Mood	SUBJECT
ROOT	**-ri-**	-chi-	**-ku-**	**-mu-**	-sha-	**-wa-**	**FUT**	-y
+	**-rqo-**			**-pu-**			**COND**	-n
Deriv.	**-yku-**							-ni
Suffix	**-rpari-**							-nki

Figure 2

Early Quechua template with mutually exclusive suffixes

The adults, by contrast, produce combinations of suffixes which appear to be mutually exclusive for the younger children, repeat elements including double causatives, and vary suffix order. The three oldest children produce similar combinations of suffixes and variable order, as

shown in Table 11. The adults and the two oldest children, Tomas and Ines, also produce forms with the suffixes *-qa-* and *-ta-*, both filler suffixes that lack any semantic substance.

Table 11

Combinations, duplications, variations produced only by oldest children

	COMBINATIONS	DUPLICATIONS	ORDERING VARIATIONS
JUAN 3;0-3;2	-ri- + -rqo- Inch Exhort		-ri- + -rqo- / -rqo- + -ri- Inch Exhort Exhort Inch ?-wa- + -yku- + -mu- 1 obj Aug Direct ?-ku- + -rqo- + -ku- + -mu- ?Refl Exhort ?Refl Direct
INES 3;2-3;5		-ri- + -ri- Inch Inch -rqo- + -rqo- Exhort Exhort	?-pu- + -rqo- Regres Exhort
TOMAS 3;7-3;9	-mu- + -pu- Direct Regres -ri- + -rqo- Inch Exhort ?FUTURE + CONDITIONAL	-rqo- + -rqo- Exhort Exhort	?-sqa- + -ri- Result Inch

In terms of the number of suffixes appended to the root, verb complexity develops gradually, with the oldest children producing adult-like Quechua verb forms bearing four suffixes in addition to the person-of-subject morphology and independent suffixes. (See Courtney 1998).

5 Summary and Discussion

Now that we have reviewed our data from Navajo and Quechua in relation to the proposal in Pinker's (1984/1996) model of paradigm-learning, how does the model fare? In brief, the scorecard shows mixed results. That is, the data provide clear support for some of Pinker's proposals while contradicting other aspects of the model.

5.1 Aspects Supported by the Data

(a) In both Navajo and Quechua verb development, children produce bare roots or stems, in support of Pinker's prediction that children first isolate the verb root or stem. The phenomenon undermines Hyams' stem parameter, which predicts that children acquiring morphologically rich languages will not produce uninflected verb forms. (b) The Navajo data also fortify Pinker's claim that children are unlikely to hypothesize those dimensions which are low on the accessibility hierarchy. Nevertheless, we would have to attribute to other factors children's differential production of the Quechua modifying suffixes, augmentative /-yku-/ and exhortative /-rqo-/, which are presumably equally accessible grammatical notions. (c) We may also infer

from Pinker's discussion that the developing template is rigidly ordered, and this is the case for both languages. Navajo children never make errors in the sequencing of verb morphemes, and very young Quechua speakers are clearly conservative in their early production of complex verbs. They do NOT start with the assumption, as Muysken (1988) has suggested, that both orders are possible, in principle, for any set of two affixes.

Further, (d) we find evidence that the Unique Entry Principle is an important constraint on children's hypotheses in the acquisition of inflectional systems. That is, children seek one-to-one form/function correspondences. This is evident in the apparent effect of homophony on the emergence of different affixes: both Navajo and Quechua children take longer to acquire productive competence for affixes with the same form but different meanings and functions. Further evidence is children's early tendency to impose on pairs of affixes a relation of complementary distribution or mutual exclusivity not found in adult production. With further respect to homophony, however, it is puzzling that young Quechua learners should so rapidly master the homophonous reflexive and augmentative suffixes. One may speculate that very young children fulfill the unique form/function expectation embodied in the Unique Entry Principle by merging the modifying function of the augmentative with one meaning of the reflexive, i.e., the notion of general personal involvement. As to the differential productivity of the augmentative and exhortative suffixes by children acquiring Quechua, it may be that the Unique Entry Principle constrains very young children to attend to only one of these two suffixes, to which they initially assign the general, undifferentiated notion of "intensifier."

5.2 Aspects Contradicted by the Data

While small children acquiring Navajo and Quechua do often produce bare roots and stems, (e) paradigm-building does NOT proceed from the stem outward as Pinker has suggested. Perceptual salience is an important factor here; following Slobin and Peters, among others, children attend to the word edges first, gradually working inward. (f) For both languages, the impact of frequency in the input is also somewhat questionable. For example, some suffixes occurring frequently in child-directed, Quechua adult speech appear early in the child corpus, and others do not, and the most frequent Navajo aspectual prefix in adult input is not the first to be produced.

5.3 Aspects only Partially Supported by the Data

If affix production is meaningful, (g) children do not necessarily assign to the affixes the same semantic features that prevail in the adult grammar. In fact, children appear to acquire the features encoded in affixes gradually, as they refine their inflectional paradigms through the creation of new cells. In fact, there is some evidence that children produce unanalyzed strings of affixes (cf., Peters 1985, Franco and Landa 1998, Rubino and Pine 1998). As to child insertion of nonsense filler morphemes into the verb complex, it is not at all clear whether this reveals a competence deficit, performance difficulty, or a kind of prosodic competence beyond grammatical performance limits.

Finally, (h) Pinker's proposal that paradigm-building proceeds from word-specific to generalized patterns of inflection is only partly right, since, in both Navajo and Quechua, not all affixes first appear on a very small subset of verb roots or stems. In fact, children may develop production routines for the formation of complex verbs by attaching, to a wide variety of roots, only a small set of affixes, both singly and in combination with one another. In this way,

children may develop sets of two- and three-affix mini-templates. We must also consider the types of Navajo and Quechua affixes that tend to emerge all or none, that is, those which are NOT initially limited to a small number of lexical roots. It is intriguing that, in both languages, the morphological expression of arguments appears to emerge categorically. We have also seen that learners of Quechua, at least, exhibit an early preference for the analytic expression of arguments. In the case of Ana's first-person utterances, the explicit pronoun arguments largely vanish as soon as the subject-marking morphology falls into place.

6 Concluding Remarks

This study yields compelling questions for further study. For example, how do children learning Navajo and Quechua isolate the verb roots and stems? Since Quechua roots are invariable, but Navajo stems vary according to aspect and mode, how does this affect the abstraction of bare stems from the input? We have also seen that Quechua children, at least, consistently include explicit pronoun subjects before the person-of-subject paradigm fully emerges. Another issue is the significance of an intriguing coincidence: the uniform appearance of the argument-marking affixes on all roots occurs simultaneously in both languages. Also, to what extent is child production of verb morphology meaningful? What meanings do children initially assign to individual affixes and their combinations?

Although we believe the naturalistic production data reported in this study offer significant insight, the exploration of this and other issues requires additional research and would profit from complementary formal measures of elicited production and comprehension.

References

Aksu-Koç, Ayhan and Dan I. Slobin. 1985. The acquisition of Turkish. In *The crosslinguistic study of language acquisition, Volume 1*, ed. Dan I. Slobin. Hillsdale, NJ: Lawrence Ehrlbaum Associates.

Bybee, Joan. 1985. Morphology: A study of the relation between meaning and form. The Netherlands: John Benjamins.

Cerrón-Palomino, Rodolfo. 1987. *Lingüística Quechua*. Cuzco, Peru: Centro de Estudios Rurales Andinos Bartolomé de Las Casas.

Courtney, Ellen. 1998. Child acquisition of Quechua morphosyntax. Doctoral dissertation. University of Arizona.

Creamer, Mary Helen. 1974. Ranking in Navajo nouns. *Navajo Language Review* 1:29-38.

Franco, Jon and Alazne Landa. 1998. Symmetries and asymmetries in the acquisition of Basque inflectional morphology by Spanish-speaking children and adults. *Southwest Journal of Linguistics* 17(1):31-48.

Greenberg, Joseph H. 1963. *Universals of language*. Cambridge, MA: The MIT Press.

Hoff-Ginsberg, Erika. 1997. *Language development*. Pacific Grove: Brooks/Cole Publishing Company.

Hyams, Nina. 1994. Nondiscreteness and variation in child language: Implications for principle and parameter models of language development. In *Other children, other languages: Issues in the theory of language acquisition*, ed. Yonata Levy. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hyams, Nina. 1986a. Core and peripheral grammar and the acquisition of inflection. Paper presented at the 11th Annual Boston University Conference on Language Development, Boston.

Mithun, Marianne. 1989. The acquisition of polysynthesis. *Journal of Child Language* 16:285-312.

Muysken, Pieter. 1986. Approaches to affix order. *Linguistics* 24:629-643.

Muysken, Pieter. 1988. Affix order and interpretation: Quechua. In *Morphology and Modularity*, ed. M. Everaert, A. Evers, R. Huybregts and M. Trommelen. Holland and USA: Foris Publications.

Olguin, Raquel and Michael Tomasello. 1993. Twenty-five-month-old children do not have a grammatical category of verb. *Cognitive Development* 8:245-272.

Peters, Ann M. 1995. Strategies in the acquisition of syntax. In *The handbook of Child Language*, ed. Paul Fletcher and Brian MacWhinney. Cambridge, MA: Basil Blackwell, Inc.

Peters, Ann M. 1985. Language segmentation: Operating principles for perception and analysis of language. In *The crosslinguistic study of language acquisition, Volume 2*, ed. Dan I. Slobin. NJ: Lawrence Erlbaum Associates.

Pinker, Steven. 1984, 1996. *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Pizzuto, Elena and Maria Cristina Caselli. 1994. The acquisition of Italian verb morphology in a cross-linguistic perspective. In *Other children, other languages: Issues in the theory of language acquisition*, ed. Yonata Levy. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pye, Clifton. 1983. Mayan telegraphese: Intonational determinants of inflectional development in Quiché Mayan. *Language* 593:583-604.

Rubino, Rejane B. and Julian M. Pine. 1998. Subject-verb agreement in Brazilian Portuguese: What low error rates hide. *Journal of Child Language* 25:35-59.

Saville-Troike, Muriel. 1996. Development of the inflected verb in Navajo child language. In *Athabaskan language studies: Essays in honor of Robert W. Young*, ed. Eloise Jelinek, Sally Midgette, Keren Rice, and Leslie Saxon. Albuquerque: University of New Mexico Press.

Slobin, Dan I. 1985. Crosslinguistic evidence for the construction of language. In *The crosslinguistic study of language acquisition*, Volume 2, ed. Dan I. Slobin. NJ: Lawrence Erlbaum Associates.

Thompson, Chad. 1996. The history and function of the yi-/bi- alternation in Athabaskan. In *Athabaskan Language Studies:Essays in Honor of Robert W. Young*, ed. Eloise Jelinek, Sally Midgette, Keren Rice, and Leslie Saxon. Albuquerque: University of New Mexico Press.

Young, Robert W. 1987. *The Navajo language: A grammar and colloquial dictionary*, Revised Edition. Albuquerque: University of New Mexico Press.

(Courtney)
Department of Foreign Languages and Literatures
East Carolina University
Greenville, North Carolina 27858

courtneye@mail.ecu.edu

(Saville-Troike)
Department of English
University of Arizona
Tucson, Arizona 85721

msaville@u.arizona.edu

Appendix

Terms for abbreviations used in tables and interlinear glosses for Quechua forms

Independent suffixes	Aff - Affirmative marker	/-mi/
	Add - Additive suffix	/-pis/
Nominal Suffixes	Poss - Person-of-possessor	1st: /-y/, 2nd: /-yki/, 3rd: /-n/
Verbal Suffixes	Subj - Person-of-subject	lst: /-ni/, 2nd: /-nki/, 3rd: /-n/
	Obj - Person-of-object	1st: /-wa-/, 3rd>2nd: /-sunki/ 1st>2nd: /-yki/
	Fut - Future	1st: /-saq/, 2nd: /-nki/, 3rd: /-nqa/
	Past	/-r(q)a-/
	Reflex - Reflexive	/-ku-/
	Prog - Progressive	/-sha-/
	Result - Resultative	/-sqa-/
	Imper - Imperative	/-y/
	Pot - Potential	/-na-/
	Cond - Conditional	/-man/
	Dir - Directional	/-mu-/
	Regr - Regressive	/-pu-/
	Desid - Desiderative	/-naya-/
	Inch - Inchoative	/-ri-/
	Exhort - Exhortative	/-rqo-/, /-ru-/
	Aug - Augmentative	/-y(k)u-/, /-(y)ku-/
	Asst - Assistive	/-ysi-/
	Caus - Causative	/-chi-/

Toward an OT Account of Yaqui Reduplication*

Jason D. Haugen

1 Presentation of Data

Yaqui[1] is a Uto-Aztecan language spoken in southern Arizona and in Sonora, Mexico. There are two kinds of reduplication in Yaqui: primary and secondary.[2] These can be characterized as morphologically distinct because of their differing semantics: primary reduplication indicates "habitual action", whereas secondary reduplication indicates "distributive" or "iterative" action. Relevant data for these two kinds of reduplication are given in (1) and (2), respectively (the reduplicant is indicated in boldface). I should point out that there are extremely interesting effects of vowel shortening and stress in these data, which are discussed in Demers et al. (1999). These effects are ripe for a treatment in Optimality Theoretic terms, but I will leave this for a more full treatment in the future. For the purposes of this paper I will only be interested in the shape of the reduplicants.

(1) *Primary Reduplication: Red1 = "habitual action"*

a. vaane	**va.**vane	'irrigate'
b. vusa	**vu.**vusa	'awaken'
c. vamse	**vam.**vamse	'hurry'
d. chepta	**chep.**chepta	'jump over'
e. patta	**pat.**patta	'cover'
f. 'eta	**'e.**'eta	'shut'
g. 'amuse	**'a.**'amuse	'go hunting'
h. 'eecha	**'e.**'echa	'plant'
i. 'ivakta	**'i.**'ivakta	'hug someone'
j. hia	**hi.**hia	'sounds'
k. wiuta	**wi.**wiuta	'tear it down'
l. suale	**su.**suale	'believe'

The shifting of stresses and shortening of vowels in the reduplicated forms can be accounted for by following Demers et al.'s (1999) analysis of prominence in Yaqui words. According to their analysis (and based upon argumentation delivered therein), Yaqui coda consonants do not typically count as moraic units. In each case of primary reduplication, the reduplicant copies the first mora of the base, and carries with it any coda consonant which

* I would like to thank the following individuals for helpful discussion of this data and/or earlier versions of this paper: Mike Hammond, Sean Hendricks, Bob Kennedy, Terry Langendoen, Katie Russell, as well as the editors of this volume. Any residual craziness is my own.

[1] All Yaqui data in this paper are taken from Escalante (1985) and/or Shaul et al. (1999). Following the tradition set forth by these sources, I use the official Yaqui orthography throughout (the one exception being my making word-initial glottal stops overt on words which the orthography leaves vowel-initial).

[2] The terms "primary reduplication" and "secondary reduplication" were used in Escalante (1985), and have been subsequently adopted by Demers et al. (1999) and Shaul et al. (1999), and I use these terms here.

appears in the same syllable as that vocalic mora. This trend fulfills a template constraint[3] for primary reduplication in Yaqui (henceforth "Red1"): "Red1=μ". The copying of a coda consonant in a reduplicant only if it occurs in the same syllable as the reduplicated mora is unusual in that it seems to require reference to surface syllable position, in contrast to normal assumptions of syllabification which maintain that syllabification is a surface phenomenon. I will discuss in detail the issues raised by the reduplication of coda consonants (and not onset consonants) in my analysis below, and will provide a way out of this apparent dilemma.

The reduplicant in Yaqui secondary reduplication (henceforth "Red2") always appears as a closed syllable with two moras, where the second mora is occupied through gemination of the onset consonant of the base. I will assume that this is an example of "compensatory lengthening" (Hayes 1989), where the gemination of the onset consonant to coda position of the reduplicated syllable occupies the second mora. Examples can be seen in the data in (2):

(2) *Secondary Reduplication:*[4] *Red2 = "iterative"*

a. bwiika	**bwib.**bwika	'sing'
b. bwana	**bwab.**bwana	'cry'
c. teeka	**tet.**teka	'lay it across'
d. vahume	**vav.**vahume	'swim'
e. tuuke	**tut.**tuke	'flicker out'
f. yena	**yey.**yena	'smoke (tobacco)'

Unfortunately, I have not found any examples in the literature on Yaqui reduplication which discuss secondary reduplication in words which contain word-medial consonant clusters. For now I will assume that consonants in coda position in the input word (i.e. the unreduplicated form) do not count as moraic units, and those which geminate (i.e. in the reduplicated form) do count as moraic units. It is this difference in moraic weight which accounts for gemination (I will discuss this further, and make predictions, in section 2.2.1).

In sum, there are two distinct reduplicative morphemes: Red1 ("habitual") and Red2 ("distributive/iterative"). The grammatical distinctness of these two abstract morphemes can be demonstrated by examining the different reduplication patterns with a single verb. Here, I will examine the verb *bwiika*, 'to sing':

(3) *Yaqui Reduplication*

/bwiika/		'to sing'
	aapo bwiika	'he is singing'
Red1	aapo **bwi**bwika	'he sings'
Red2	aapo **bwib**bwika	'from time to time he sings'

I will now turn to an Optimality Theoretic analysis of this data.

---

[3] A template constraint, based upon McCarthy and Prince (1993), is defined as "Morphological Category = Prosodic Category".

[4] Some verbs have semantic restrictions on one or both of these kinds of reduplication, which gives the reduplication of those verbs specialized meanings:

e.g. koce 'sleeping' **kok**koce 'falling asleep (on and off)'
nooka 'speaking' **non**nooka 'gossiping'
(Shaul et al. 1999)

2 Analysis

I will offer an analysis of Yaqui reduplication in terms of Optimality Theory (OT) (Prince and Smolensky 1993; McCarthy and Prince 1993), and specifically Correspondence Theory (McCarthy and Prince 1995). Crucial to my analysis and to the theory in general is the strict ranking of constraints to rule out certain candidates while preserving others. The consideration of all constraints and all candidates in parallel will yield the optimal candidate for each underlying input, given the particular set of ranked constraints.

2.1 Yaqui Primary Reduplication

The first constraints that are required will define the reduplicant and where it is aligned in the word. These constraints include AlignRedLeftWordLeft (AlignRedLWdL) and L-AnchorBR. AlignRedLWdL aligns the reduplicant to the left edge of the word, yielding a prefix, rather than an infix or a suffix. The anchor constraint makes sure the left edge of the reduplicant corresponds to the left edge of the base. The result of this anchoring is that the reduplicant and the base both begin with an identical segment.

As mentioned above, we will assume a template constraint for Red1: Red1=μ. This template constraint requires the reduplicant to be composed of a syllable containing only one mora. Because coda consonants generally do not count as moraic units in Yaqui, coda consonants which get reduplicated should not count as violations of this template. Thus, for a word such as *vusa* 'awaken', and given just this one template constraint, the template can be met by either of two output forms, as shown in (4), wherein for the time being the other constraints (AlignRedLWdL and L-AnchorBR) are assumed): [5]

(4)

Red1 + /vusa/	Red1=μ
a. ☺ **vu** vusa	
b. ☹ **vus** vusa	
c. **vusa** vusa	*!

Candidate c is immediately ruled out because it, being a full copy of a disyllabic base, violates the template constraint allowing a syllable with only one mora. Since coda consonants in Yaqui do not always constitute moraic units (and I will not formalize here when they do), candidate b does not violate the template. Instead, we can rule out candidate b through some other constraint, such as *Coda, which insists that syllables do not have codas. This is shown in (5):

(5)

Red1 + /vusa/	Red1=μ	*CODA
a. ☺ **vu** vusa		
b. **vus** vusa		*!
c. **vusa** vusa	*!	

[5] The symbols that I will use are as follows: '☺' denotes the correct output form that is correctly chosen by a particular constraint ranking; '☹' denotes the correct output which fails to be chosen by a particular constraint ranking; and '☹' denotes an incorrectly selected candidate.

With just these two constraints (in addition to the assumed AlignRedLWdL and L-AnchorBR) we predict the correct winning candidate and rule out other candidates for this input. However, these two constraints alone will not be sufficient for the cases where there is a word-medial consonant cluster, such as the form *vamse* 'hurry'. As we see in (6), the same constraint ranking that worked before now predicts the wrong output candidate:

(6)

Red1 + /vamse/	Red1=μ	*Coda
a. ☹ **va** vamse		
b. ☹ **vam** vamse		*
c. **vamse** vamse	*!	

This time we actually want candidate b to win, yet it is ruled out by a well-formedness constraint, and another candidate is incorrectly deemed optimal.

At first glance what seems to be required here is some kind of constraint which not only allows what surface as coda consonants in the input word to be reduplicated with (or perhaps even as a part of) the mora associated with the vowel which gets copied, but also requires this. What I will temporarily propose is a correspondence constraint which requires that segments in the reduplicant be faithful to the syllabic position of those segments in the base. This constraint is formalized in (7):

> (7) MAX-σ-POSITION: Every coda consonant in the input must be in coda position in the output, and every onset consonant in the input must be in onset position in the output.[6]

If we rank this faithfulness constraint between the template and *Coda, then we produce the desired results, as demonstrated in (8):

(8)

Red1 + /vusa/	Red1=μ	MAX-σ-position	*Coda
1a. ☺ **vu** vusa			
b. **vus** vusa		*!	*
c. **vusa** vusa	*!		

Red1 + /vamse/	Red1=μ	MAX-σ-position	*Coda
2a. **va** vamse		*!	
b.☺ **vam** vamse			*
c. **vamse** vamse	*!		*

Thus, if we were to assume such a constraint, we see that the crucial violation of candidate b in (5) above (1b in (8)) was not actually *Coda, but MAX-σ-position. That is, *Coda can be violated without punity if the constraint that requires syllable faithfulness (MAX-σ-position) is ranked above *Coda.

[6] Presumably, such a constraint, if we were to keep it, would also require that vocalic peaks remain vocalic peaks.

Crucial to such an analysis is the fact that all information in a copied syllable must appear in the reduplicated syllable. I would like to pose the constraint MAX-σ-position as a variation on the more commonly utilized MAX-BR, which is defined as the requirement that every segment which appears in the base should also appear in the reduplicant. This has the effect of requiring that reduplicants be as large as is possible given the other constraints at work in any given tableau. The base is usually taken for granted as consisting of what I am here calling the "input word". In the case of Yaqui primary reduplication, it appears to be the case that this maximization is relevant only for the syllable that gets copied: the first syllable of the input word. Given that the theory already has an independently motivated constraint for base-reduplicant identity (MAX-BR), and since, under OT assumptions, syllabification does not occur until the output level, the positing of a new constraint appealing to syllable-placement is to be dispreferred over the use of constraints that are already established. If we just assume that in Yaqui primary reduplication the base *is* the first syllable of the input word, then what I have proposed here as "MAX-σ-position" is actually MAX-BR: the base is the first syllable of the input word, and the reduplicant in Yaqui primary reduplication copies the entire base.

In (8) above, violations of "MAX-σ-position" accrued both for codas in the base that were incorrectly not copied (e.g. ***vavamse***), as well as for extra-base onsets which were incorrectly copied as codas on the reduplicant (e.g. ***vusvusa***). With the re-evaluation of this constraint as MAX-BR, only codas in the base (the first syllable of the word) which fail to be copied in the reduplicant can count as violations of this correspondence constraint. Onsets which get copied are not a part of the base, and so they can be ruled out by making another constraint undominated. DEP is a constraint which will do the trick: the DEP family of constraints forbids anything not in an input from appearing in an output. Since the onset to a second syllable in a word is not in the first syllable of the input word (the domain for defining the base), if a word-medial onset is copied then the fact that a segment does not appear in the base but does appear in the reduplicant violates DEP. With MAX-σ-position redefined as MAX-BR and with the addition of DEP, (8) above becomes (9) below. Two candidates have been added (candidate d for each input) which fail to reduplicate the entire base. The first syllable of the input word, which is the base, is underlined.

(9)

Red1 + /vusa/	DEP	Red1=μ	MAX-BR	*Coda
1a. ☺ **vu** vusa				
b. **vus** vusa	*! (s)			*
c. **vusa** vusa	*!* (sa)	*		
d. **v** vusa			u!	

Red1 + /vamse/	DEP	Red1=μ	MAX-BR	*Coda
2a. **va** vamse			m!	
b.☺ **vam** vamse				*
c. **vamse** vamse	*!* (se)	*		
d. **v** vamse			a!m	

While the template constraint itself appears to have no functional role in the tableau as given, I assume that it is the work of this template which defines the base as the first syllable of

the input word, and it is this definition which allows the DEP constraint to reject candidates which supersede the base domain, as defined by the template.

In (10) below we see how all of the constraints mentioned so far act in concert to yield an optimal output form. The reduplicant-defining constraints (L-Anchor-BR, AlignRedLWdL, and Red1=μ) are undominated. The only crucial rankings are those mentioned so far:

(10) Primary Reduplication: DEP, Red1=μ >> MAX-BR >> *Coda

/Red1+ bwiika/	L-Anchor BR	AlignRed LWdL	Red1=μ	DEP	Max-BR	*Coda
a. ☺ **bwi** bwika						
b. **bw** bwika					i!	
c. bwika **bwi**		*!				
d. **ka** bwika	*!					
e. **bwik** bwika				*! (k)		*
f. **bwika** bwika			*	*!* (ka)		

This table illustrates the effects of the constraint rankings for the verb *bwiika* in primary reduplication. Candidate b violates MAX-BR, which requires that all segments in the base appear in the reduplicant, and is ruled out. Candidate c incorrectly aligns the reduplicant to the right edge of the base, and this suffixation is a fatal violation of AlignRedLWdL. Candidate d anchors the right edge of the reduplicant with the right edge of the base, and this is a fatal violation of L-AnchorBR. Candidate e has the correct anchoring and alignment, but its inclusion of a coda segment in the reduplicant which is not present in the base violates the DEP constraint, and this rules it out. Since candidate f can be ruled out through two violations of the DEP constraint, it again appears to be the case that the template plays no role in this tableau. But, once again, I assume that the template itself is what defines the base domain, so the inclusion of the template is crucial. The result of this position is that, formally, the DEP constraint can be considered a part of the template constraint (perhaps through conjunction). The winning candidate is a, which survives the constraint evaluation entirely unscathed, and thus is the optimal candidate.

2.2 *Yaqui Secondary Reduplication*

Let us now turn our attention to Red2, which in each case yields a reduplicant with a closed syllable where the coda of the reduplicant syllable corresponds with the onset segment of the input word (i.e. through gemination). Following Demers et al. (1999), I will assume that for each of these cases the necessary template constraint is "Red2=μμ". This constraint should be construed as one which is defined as a single syllable consisting of two moras, and any candidate which fails to have both of its moras in the same syllable (e.g. a candidate with two syllables consisting of 1 mora each) would violate it.[7]

Unlike in primary reduplication, it may not be essential here to assume that the base for reduplication is the first syllable of the input word. As mentioned above, the literature does not

[7] In other words, the template constraints for both Red1 and Red2 are orthographically represented as "Red1=μ" and "Red2=μμ", respectively, for convenience, although in fact they could more accurately be written as "Red1=σ/μ" and "Red2=σ/μμ". I follow Demers et al. (1999) in counting the mora as the unit for reduplication, although syllable onsets and non-moraic codas are also copied.

report on what happens to words with medial consonant clusters (e.g. *vamse* 'hurry'). Thus, for now I will focus on analyzing those words listed above in (2): those with open syllables as first syllables in the input word (e.g. *bwiika* 'sing'), leaving the analysis of other kinds of words open to future research.

How do we get this closed syllable? I will argue that what is needed to account for the reduplicant in secondary reduplication is the combination of the template constraint and two anchoring constraints, wherein each edge of the reduplicant is anchored to the left edge of the base. This yields gemination from the left edge of the input word to the right edge of the reduplicant. The possibility for a double-anchoring analysis first came to my attention through Hendricks' (1999) a-templatic analyses of bare-consonant reduplication, in which reduplicants of the form CC get defined and expressed minimally by anchoring both edges of the reduplicant to opposite edges of the base. Hendricks' compression theory was initially proposed as an alternative to templatic accounts of reduplication. This move was required in order to adequately derive reduplicants in bare-consonant reduplication. In this kind of reduplication, the first and last consonantal segments of a word get reduplicated: the "compression" occurs when Anchoring constraints are ranked higher than constraints such as AlignRootLeft, which demands that the root appear at the far left of the word. For the anchoring constraints to be satisfied, this alignment constraint must be violated. If both anchoring constraints are ranked higher than this alignment constraint, then the minimal unit which could satisfy both constraints anchoring the left and right edge of the base would be a reduplicant with two segments: one corresponding to the left-most segment of the base, satisfying L-AnchorBR, and one corresponding to the right edge of the base, satisfying R-AnchorBR.

As Hendricks shows, this kind of reduplication cannot be accounted for by the use of a template constraint, since the bare-consonant reduplicant does not constitute a prosodic unit. Hendricks' theoretical stance, "compression", accounts for bare-consonant reduplication by using Anchor constraints *without* template constraints (or at least with the template constraints ranked so low as to be irrelevant); the factorial typology of Optimality Theoretic constraint ranking allows for this move, as well as for an extension of this move: the use of Anchor constraints *with* template constraints.[8] In the former case Anchoring occurs on opposite edges of the base; in the latter Anchoring can occur on the same edge.

Secondary reduplication in Yaqui can be accounted for by anchoring both edges of the reduplicant with the left edge of the base. The constraints that will be required are as follows. A template constraint (Red2=μμ) will be required to force the reduplicant to surface as a single syllable with two moras, as motivated by Demers et al. (1999). Still considering the verb *bwiika*, so far this will yield an onset and vowel with some consonant as a coda:[9] giving us *bwiC*. Since secondary reduplication is realized as a prefix, it will have the same alignment as primary reduplication: AlignRedLWdL. This will align the reduplicant to the left edge of the word, yielding a prefix: ***bwiC***.*bwika*. It is the right-edge anchoring of the secondary reduplicant which differentiates it from the primary reduplicant. Because Red2 has an onset and coda which are copies of the same segment of the base, Red2 requires two anchoring constraints:

[8] Hendricks (among others, see discussion in Hendricks 1999) in fact rejects the use of any (prosodic) templates whatsoever. While the rejection of templatic constraints in some cases seems to be the correct move (as Hendricks has ably shown for bare-C reduplication), I do not share Hendricks' conviction that this entails that the notion of "template" should be abandoned altogether. I would like to keep the notion of "templates", although I agree with Hendricks that morphological templates need not always be prosodic units.

[9] See Demers et al. for reasons for the shortening of the vowel and the shift of the stress. Here I assume that the vowel occupies only one mora.

AnchorRedLBaseL, which requires correspondence between the left edge of the base and the reduplicant (here, generating a *bw* on the left edge of the reduplicant), and AnchorRedRBaseL, which requires correspondence between the right edge of the reduplicant and the left edge of the base (here, generating a *b* on the right edge of the reduplicant[10]).

The OT account of secondary reduplication using these constraints can be seen in detail in (11). Here I omit the constraint MAX-BR, for reasons which I will make clear in section 2.2.1:

(11) Secondary Reduplication with a Doubly-Anchored Reduplicant

Red2+/bwiika/	Red2=μμ	AlignRedL Word L	Anchor RedLBaseL	Anchor RedRBaseL
a. **bwi** bwika	*!			*
b. ☺ **bwib** bwika				
c. bwika **bwib**		*!		
d. **kab** bwika			*!	
e. **bwik** bwika				*!
f. **bwk** bwika	*!			*

Because candidate a is not composed of a syllable with two moras, it is ruled out because it fails to meet the template. Candidate c is ruled out because the suffixation of the reduplicant violates the AlignRedLWdL constraint. Candidate d violates the constraint which forces the left edge of the reduplicant to correspond to the left edge of the base. E is ruled out because the right edge of the reduplicant is not anchored to the left edge of the base. In other words, the effect of the two anchoring constraints acting in conjunction is to yield templatic syllable edges of a certain form: a closed syllable with identical onset and coda. Candidate f, in which the reduplicant is not a well-formed syllable, is ruled out because it violates the template.[11] The only candidate which matches the template and survives the other constraints is candidate b, which is the actual output form. Thus, these constraints correctly predict the correct output form; since the winning candidate violates none of these constraints, their ranking cannot be determined, and is not crucial.

2.2.1 The Importance of Edge-Anchoring for Red2

In (11) above I omitted the constraint MAX-BR, primarily because it is uncertain what the domain for defining the base is in secondary reduplication. This domain would define the scope of the constraint, and is required in order to assess violations of this constraint. What is clearly necessary is the anchoring of the right edge of Red2 to the left edge of the input word, as is seen in the gemination which occurs in this reduplicant.

[10] For the purposes of this analysis I will assume that the *w* in *bw* is strictly an onglide and as such appears only prevocalically; thus, for the purposes of assessing input-output faithfulness, *b* (which only appears syllable-initially on borrowings from English or Spanish) should be considered to be the correspondent of *bw* preconsonantally, and especially in situations of gemination.

[11] Actually, this candidate is a bare-consonant reduplicant, and the violation here is really a Peak violation, which I assume is contained covertly in the template constraint: RED must be a *well-formed* syllable with two moras. Presumably, at least the first consonant in this case would be an onset, which does not carry moraic weight.

The importance of this right edge anchoring can be seen by contrasting Red2 in Yaqui with its counterpart in a hypothetical (but possible) related language, which could have re-ranked two constraints to yield a differently shaped reduplicant for Red2. I will call this hypothetical language *Yaqui. This hypothetical equivalent of Yaqui secondary reduplication is not amenable to my anchoring analysis, since its secondary reduplication does not anchor both edges of the reduplicant with the left edge of the base. In other words, the second mora required by the template in secondary reduplication is not filled through gemination, but through the copying of further consonantal material from the input word, as seen in (12) (c.f. (3)):

(12) **Yaqui Reduplication*

/bwiika/		'to sing'
	aapo bwiika	'he is singing'
Red1	aapo **bwi**bwika	'he sings'
Red2	aapo **bwik**bwika	'from time to time he sings'

So the form *bwika* 'to sing' yields the reduplicated forms ***bwi****bwika* and ***bwik****bwika*. In *Yaqui, this can be explained by considering the entire input word as the base, and by using the constraint MAX-BR the correct output is predicted. This is shown in (13), where the left Anchor constraint is assumed:

(13) *Yaqui Secondary Reduplication: MAX-BR >> AnchorRedRBaseL

Red2+/bwika/	Red2=μμ	AlignRedLWdL	MAX-BR	AnchorRed RBaseL
a. **bwi** bwika	*!		ka	*
b.☺ **bwik** bwika			a	*
c. **bwika** bwika	*!			*
d. bwika **bwik**		*!	a	*
e. **bwib** bwika			k!a	
f. **bwb** bwika	*!		ika	

In this tableau, candidates a, c and f are ruled out for violating the template constraint: they are not comprised of a single syllable comprised of two moras. Candidate d is ruled out because the reduplicant is incorrectly suffixed, violating the left-alignment constraint for the reduplicant. The two interesting candidates are b and e, each of which fulfills the template. The difference between these two lies in how they fulfill the template: candidate b fulfills the template by copying as much of the base (here, the entire input word) as will fit, giving it one violation of MAX-BR, and one violation of the lowly-ranked AnchorRedRBaseL. Candidate e does the opposite: it sacrifices an extra violation of MAX-BR in favor of AnchorRedRBaseL. Since MAX-BR is a gradient constraint, and because it is crucially ranked higher than the R-Anchor constraint, this second violation of MAX-BR rules candidate e out in favor of candidate b.

Turning back to actual Yaqui, we see how the re-ranking of two constraints yields the different outputs generated by two closely related grammars: by re-ranking the AnchorRedRBaseL constraint above MAX-BR, the opposite result holds (and here the L-Anchor constraint is again assumed):

(14) Yaqui Secondary Reduplication: AnchorRedRBaseL >> MAX-BR

Red2+/bwika/	Red2=μμ	AlignRedLWdL	AnchorRed RBaseL	MAX-BR
a. **bwi** bwika	*!		*	
b. **bwik** bwika			*!	a
c. **bwika** bwika	*!*		*	
d. bwika **bwik**		*!	*	a
e.☺ **bwib** bwika				ka
f. **bwb** bwika	*!			ika

Once again candidates a, c and f get ruled out for violating the template, and d for suffixing instead of prefixing. In the Yaqui grammar the constraint requiring the anchoring of the right edge of the reduplicant to the left edge of the base is crucially ranked higher than MAX-BR, so the candidate which violates the former instead of the latter (here, b) does so at its peril: the candidate (here, e) which does not violate the higher-ranked constraint is deemed more optimal than the one that does violate the higher-ranked constraint. Because of this crucial ranking, the domain for defining the base is irrelevant: since e does not violate anything before MAX-BR comes into play, e is the winning candidate.

3 Conclusions and Areas for Further Research

As mentioned in the introduction, one area for future analysis is the integration of the data that I have covered here with the facts of vowel shortening and stress (pitch accent) shift. Considering only the shape of the two Yaqui reduplicants, I see several interesting implications for OT. First, Yaqui Red1 indicates that explicit notions of consonantal moraicity are needed for constructs, such as reduplicative templates, which require reference to moraicity. Second, the explication of what constitutes the "base" in reduplication needs to be clear for each analysis of each set of data: as demonstrated by the comparison of the data from Yaqui Red1 and Red2, the base is not always necessarily defined in the same way, even within the grammar of a given language. Third, the effect of the current analysis is to demonstrate that the tools used to force compression, Edge Anchoring, rather than circumventing the need for templates, actually can define the realization of templates. In the present analysis, the two anchoring constraints, AnchorRedLBaseL and AnchorRedRBaseL, acting in conjunction, define the edges of a reduplicative syllable template: Yaqui Red2. How this analysis can handle the secondary reduplication of verbs with word-medial consonantal clusters is not presently known: I have not yet found data of this kind. Thus, for example, whether the data will actually yield ***vav.vamse*** or ***vamv.vamse*** (or some other possibility) from *vamse* is an empirical question which can only be answered by gathering more data. I hope to gather such data to supplement the analysis found in this paper in the near future, and to expand this analysis to include the effects of vowel shortening and the interaction of pitch accents.

References

Demers, Richard, Fernando Escalante, and Eloise Jelinek. 1999. Prominence in Yaqui words. *International Journal of American Linguistics* 65:1.

Escalante, Fernando. 1985. *A preliminary view of the structure of Yaqui.* Master's thesis: University of Arizona, Tucson.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20:253-306.

Hendricks, Sean. 1999. *Reduplication without templatic constraints: A study in bare-consonant reduplication.* Doctoral dissertation: University of Arizona, Tucson.

McCarthy, John, and Alan Prince. 1993. *Prosodic morphology: Constraint interaction and satisfaction.* Ms., University of Massachussetts, Amherst and Rutgers University, New Brunswick, NJ.

McCarthy, John, and Alan Prince. 1995. *Faithfulness and reduplicative identity.* Ms., University of Massachussetts, Amherst and Rutgers University, New Brunswick, NJ.

Prince, Alan, and Paul Smolensky. 1993. *Optimality theory: Constraint interaction in generative grammar.* Ms., Rutgers University, New Brunswick, NJ and the University of Colorado, Boulder.

Shaul, David, F. Molina, and H. Valenzuela. 1999. *Yoeme-English English-Yoeme standard dictionary: a language of the Yaqui tribe in the American southwest and northern Mexico, with a comprehensive grammar of Yoeme language.* New York: Hippocrene Books.

Department of Linguistics
University of Arizona
200 E. Douglass Bldg.
Tucson, AZ 85721

jhaugen@u.arizona.edu

Clitics, Scrambling and Parsing

William Lewis

1 Introduction

The purpose of this paper is to devise optimal algorithms for parsing linguistic structures that contain P2 (Wackernagel) clitics. Since many languages that have P2 clitics also allow scrambling, any algorithms for parsing P2 clitics must also contain algorithms for parsing scrambled structures. Most of the energy of this paper, however, will be focused on P2 parsing.

Although many languages have P2 clitics, I have focused most of my attention on Native American languages (with some exceptions). There is one major reason for this: languages of the Americas are almost entirely ignored by the computational and parsing literature, which focuses on languages of the Indo-European language family (and almost always on English, at that). By doing so, researchers deprive themselves of data and linguistic structural diversity that can help in devising more widely applicable parsing algorithms.

This is a computational paper, the intention of which is to develop parsing procedures. Little attention will be paid to a specific syntactic/morphological theory, nor will much attention be paid to the form of the output. These are concerns that can be addressed in a later stage of parser design.

What is an "optimal" parsing algorithm? I shall define the optimality of a given solution by the criteria in (1) below:

(1)

1) The optimal solution is one which uses devices and formalisms whose generative capacity is as low as possible on the Chomsky hierarchy.
2) The optimal solution uses as few "rules" or "devices" as possible.

Obviously, it will be necessary to strike a balance between these two criteria. For this reason, the issue of optimality may be somewhat flexible, depending on how much weight is given to each criterion. The most optimal solutions might require the power of context-sensitive rules, but these may be used in concert with context-free or even finite-state rules.

2 Definitions

Because this paper may appeal to a diverse audience, I will first define some of the terms used herein. These definitions will be the focus of this section.

2.1 Parsing

The term *parsing* evolved from the Latin *pars orationis*, which means "part of speech" (Russell and Norvig 1995). The task of parsing involves analyzing input, identifying its components, and assembling these components into some valid, "parsed" structure. For natural language input, a parser will generally identify the words within a sentence, assign them parts of speech, and

assemble them into phrases. The output might be a simple listing of the word and phrase categories in the input, or might contain a more involved tree or hierarchical structure. A parser might also indicate whether the input is ungrammatical, and might indicate where it fails to be grammatical. (Please note: The sense of *parser* is described here in the traditional computational sense. There are certainly many other types of parsing that will not be addressed in this paper.)

2.2 The Chomsky Hierarchy

The generative capacity of a grammatical formalism describes the set of languages or structures that it can represent. The four classes of grammatical formalisms of the Chomsky hierarchy, in order of generative power, are shown in (2):

(2)	a) Recursively Enumerable (a.k.a. Type 0, Unrestricted)	$AB \rightarrow C$
	b) Context-sensitive	$AB \rightarrow BA$, $ASB \rightarrow AXB$
	c) Context-free	$S \rightarrow a\ S\ b$
	d) Regular (finite-state)	$S \rightarrow a\ S$

(Please note that upper-case letters denote non-terminal symbols, and lower-case denote terminals. For a thorough discussion of the Chomsky hierarchy, see Russell and Norvig 1995, Allen 1995, or Hopcroft and Ullman 1979.)

Appealing to the first criterion of the definition of optimality described in (1), those languages appearing lower in the hierarchy are the more optimal (*regular* being the most optimal). Thus, algorithms that express regular grammars are more optimal than those that express context-free grammars, context-free more optimal than context-sensitive, and context-sensitive more optimal than recursively enumerable. It should also be noted that those grammars shown higher in the hierarchy subsume the lower grammars, but the reverse is not true. For example, all context-free grammars are context-sensitive grammars, but not all context-sensitive grammars are context-free. To say that a grammar is *strictly* of its type is to say that it cannot be described by grammars lower in the hierarchy. A strictly context-sensitive grammar, for instance, cannot be described in context-free terms. All future references to grammars in this paper will be biased towards the strict interpretation.[1]

Numerous devices manifest the above described grammars. *Finite State Machines* (FSMs), also called *Finite State Automata* (FSA) or *Simple Transition Networks* (STNs), describe the set of devices that are equivalent in expressive power to regular grammars. The following FSM describes (partially) the nominal inflections of Brazilian Quechua, a member of the Quechuan family spoken throughout South America:

[1] Optimality, as defined previously, must be appealed to here. It can be argued that many grammars (perhaps *all*) described as context-free can be expressed in finite-state terms. However, certain constructions may be more optimally described by a context-free grammar than a finite-state one, because of a sizeable reduction in the number and repetition of steps. The same may hold between context-sensitive and context-free grammars.

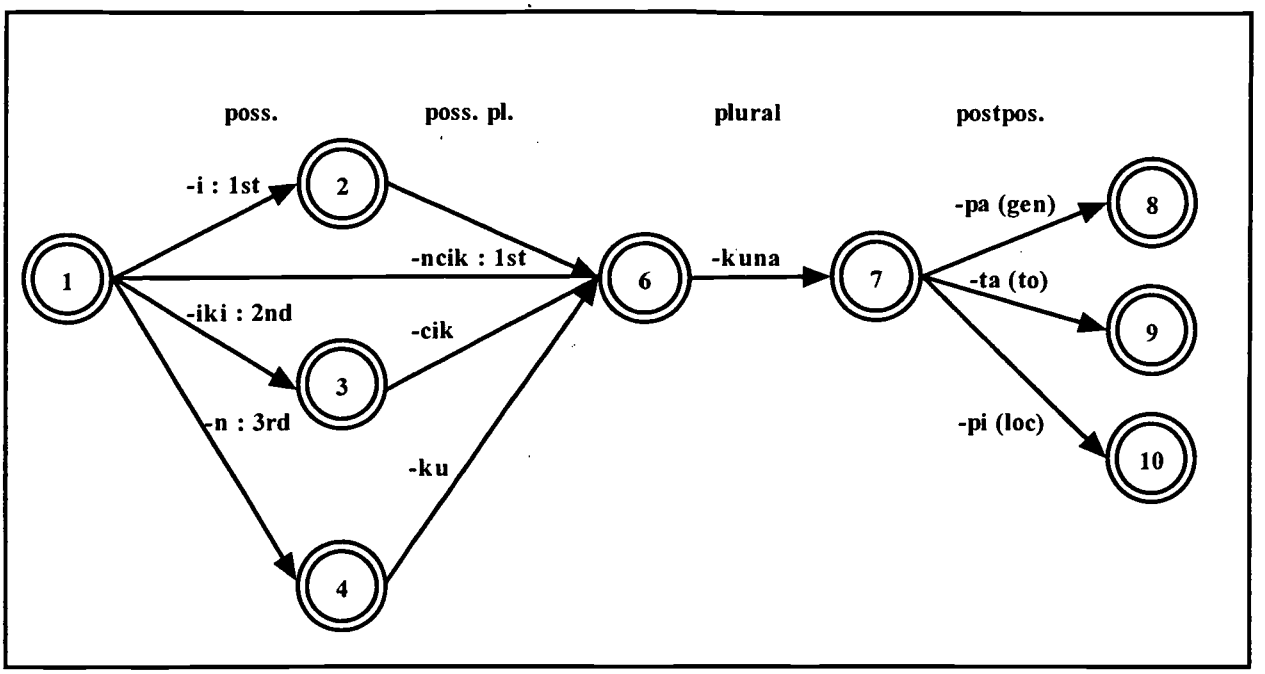

Figure 1 - STN or FSM of Brazilian Quechua Nominal Inflections

In this diagram, numbered nodes represent specific states within the FSM. Node 1 represents the start state.[2] Labeled arcs represent possible paths between nodes. Nodes with double circles represent possible end-states. If this FSM were used to check the validity of strings, (3a) would succeed, but (3b) would fail:

(3) a) -n-ku-kuna-ta
b) -iki-ku-kuna-ta

To describe context-sensitive grammars, devices that have memories and can recurse are needed.[3] *Recursive Transition Networks* (RTNs) are just such devices. The RTNs in Figure 2 describe the grammar required for parsing simple NPs in Yaqui. There are two RTNs shown: one for the noun phrase (NP) and one for the noun (N). Each node indicates the current position in the parse. Each arc represents the legal steps that can be taken from a given node to the next node. For example, from NP1, two steps are legal: N or adj.

[2] The nominal stem was not included in the diagram of this FSM. For completeness, a set of arcs representing the possible nominal stems should precede node 1. To simplify the diagram, these arcs were not included.
[3] This is to allow the expression of recursive context-free rules, such as:
S → VP (COMP) S

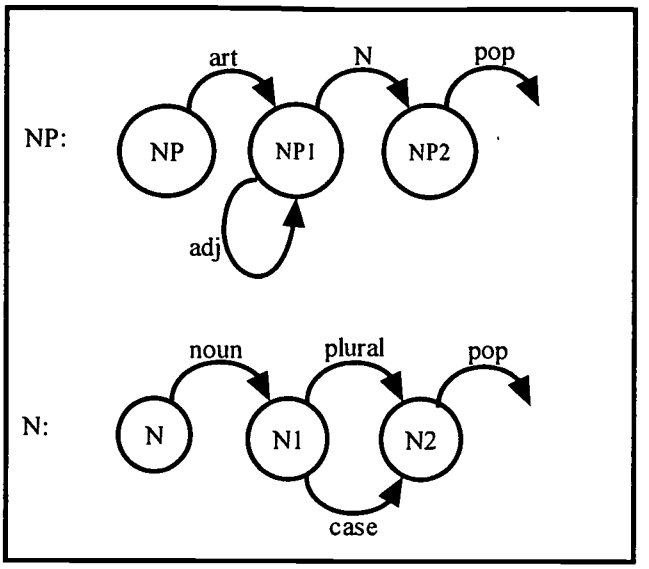

Figure 2 - RTNs are context free

Lower case letters on arcs indicate lexical items or morphemes. Upper case letters on arcs indicate calls to other networks (notably the N arc connecting nodes NP1 and NP2). The *pop* arc on each RTN indicates the end of the current RTN, whereupon control is returned to the calling RTN (a stack, a particular type of memory, is required for such an operation). For example, if the NP RTN calls the N RTN from node NP1, after the N RTN is finished processing, control is returned to the next node in the NP network, namely NP2.

For those unaccustomed to reading diagrams of this kind, two interesting points should be made about the arcs:

1. The *adj* arc in the NP RTN can be repeated multiple times. This is possible because the starting node and ending node for the arc are the same. This is representative of the grammar of Yaqui since multiple adjectives can modify a noun.
2. When traversing the N RTN, there are two possible paths between N1 and N2: plural or case. Only one of these can be chosen, not both. This is not a mistake. In Yaqui, the plural morpheme is mutually exclusive of the case morphemes.

Linear Bounded Automata (LBA) describe a set of devices that have the expressive power of context-sensitive languages. These will be discussed in further detail later.

Turing machines encompass the entire set of aforementioned devices, and are the only devices that can be used to describe recursively enumerable languages. Since no algorithms described in this paper will be recursively enumerable, further discussion of Turing machines will be avoided.

2.3 Scrambling

Scrambling, or *free word order*, is a property exhibited by numerous languages. Although languages that exhibit scrambling do have basic, canonical word orders, there is significant freedom in the surface expression of phrasal constituents. For example, in Yaqui the basic word order is SOV (Escalante 1990)[4], yet often sentences violate this order. For example, all the sentences in (4) are valid, and express basically the same concept, 'Peter kicked the cat'.

(4)	a.	Peo Peter	miisita cat:ACC	tetemuk kick:PERFECTIVE
	b.	Miisita cat:ACC	Peo Peter	tetemuk. kick:PERFECTIVE
	c.	Tetemuk kick:PERFECTIVE	Peo Peter	miisita. cat:ACC

'Peter kicked the cat.'

Various semantic, pragmatic, syntactic, and prosodic principles of a language motivate scrambling. These principles can be grouped into three major classes: contrastive focus or topicalization, anti-focus, or definiteness/specificity (the fronting of definite objects over indefinites) (Choi 1996). Whatever the motivation for scrambling, languages that exhibit this property tend to have a fairly rich inflectional morphology. In addition, the domain within which scrambling can occur is often constrained. Rich inflectional morphology and a constrained domain for scrambling will be critical to the design of solutions for parsing scrambled structures.

2.4 P2 Clitics

Second position clitics, also called *P2* or *Wackernagel* clitics, are a special class of clitics which must occupy second position within a phrase or sentence. Interpretation of "second" is subject to significant cross-linguistic variation. In some languages, second position means second phrasal constituent, as in Walpiri, an aboriginal language of Australia:

(5)	a.	wawiri kangaroo	kapi-na FUTURE (clitic)		purami cook	(Zwicky 1977)
	b.	wawiri kangaroo	njampu this	kapi-na FUTURE	purami cook	

[4] Describing the grammar using a basic word order, such as SOV, oversimplifies the notion of constituent order within a language, since there are many constituents in any language which do not fall neatly within the classes of Subjects, Objects and Verbs (consider nominal and verbal complements and adjuncts, all forms of oblique objects, etc.). This is a convenience necessary to make the discussion of parsing algorithms possible. It is also necessary to permit cross-linguistic analyses. Additional work would be necessary to integrate all such other objects into similar word order templates.

In many languages, however, second position is much more loosely defined. In Yaqui, for example, the first position constituent is defined as either the first phrase, or the first word, depending on what is receiving focus. In (5a), the phrasal constituent *uka ili'uusita* is the focused constituent, and therefore the host of the 1st person singular clitic *ne*. In (6b), just the determiner *uka* is the object of focus, and therefore the landing site for the clitic.[5]

(6) a. uka ili'uusita=**ne** aniak
DET:ACC child-ACC=1sg help:PERF
'I helped the child' (Escalante 1990)

b. uka=**ne** ili'uusita aniak
DET:ACC=1sg child-ACC help:PERF

There is even evidence that some languages allow enclisis, or "infix cliticization." One notable example is from an Iranian language Pakhto, shown in (7).

(7) a. axistə ba ye (Zwicky 1977)
buy it will he
'He would be buying it.'

b. á=ba=ye-xistə

In (7), [ba] and [ye] are members of the second position clitic cluster. In (7a) these clitics appear following the first *syntactic* element of the sentence. In (7b) they follow the first *phonological* element of the sentence. The gloss for both sentences is the same (although the placement of the clitic cluster probably has some undocumented impact on the interpretation).

Based on the examples given, it should be obvious that in order to successfully parse P2 clitics, one must consider syntactic, morphological, and possibly phonological properties of the language.

3 Parsing and Clitics

It has been argued that most, if not all, morphology can be expressed in finite-state terms (Anderson 1992, Hammond p.c.). (A notable exception might be reduplication.) For instance, the nominal affix morphology of Brazilian Quechua is easily described by the FSM shown in Figure 1.

Because clitics appear to operate at the phrasal level, yet in many ways behave as affixes, they present interesting challenges to the design of a parser. For example, the English possessive [s] is representative of most clitics, in that it is not limited as to the types of elements to which it can attach. This can be seen in the following examples:

[5] A note about notation: I will consistently separate clitics from their hosts and from other clitics by using the equal sign "=". Affixes, when appropriate to the discussion, will be separated by the dash "-". (This notational convention was borrowed from Jelinek 1996.) These are merely conveniences to make the identification of clitics and affixes easier for the reader. It is assumed that a parser would not have access to such conveniences.

(8) The student's essay.
(9) My mother and father's house.
(10) The guy we went to school with's brother.
(11) The truck driver who drove so slowly's friend.

In (8) and (9), the possessive clitic attaches to a noun, in (10) to a preposition, and in (11) to an adverb. Adopting a "lexical" strategy similar to that taken in Figure 1, one might propose the RTN fragments in Figure 3.[6]

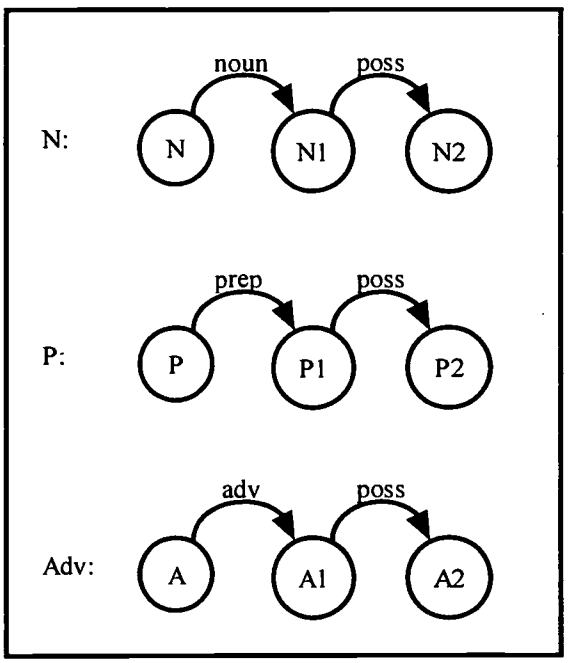

Figure 3 - RTN fragments for English possessive

The main failing of these RTNs, however, are their inability to capture an obvious generalization: the English possessive does not attach to the specific lexical category noun, preposition, and adverb, but rather, it binds to the specific phrasal category *noun phrase*, as shown in the NP RTN in Figure 4 (next page).

[6] The devices shown in figure 4 are arguably finite-state. If these fragments were integrated into a larger device, this certainly would be true. However, because the devices are given distinct names, it is assumed that other devices would reference, and *call*, these devices by the given names. *Calling* some other device uses requires memory, specifically a type of memory called a *stack*, which is a context-free mechanism.

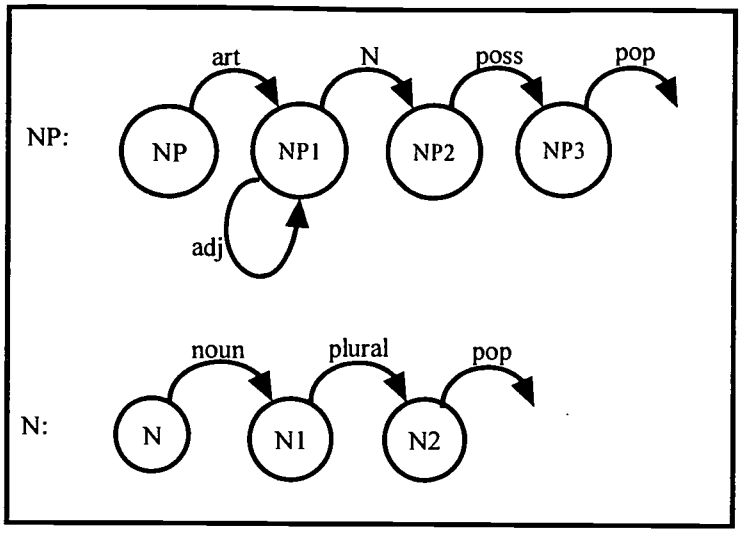

Figure 4 - English possessive as a phrasal affix

Is the device shown in Figure 4 finite-state? Arguably not. If one were to expand the shorthand notation for NPs shown in this figure to include all the necessary arcs to express all possible NPs--the forms shown in (8)-(11), as well as all others--one *would* have a finite-state device. However, the fact that the shorthand is used, indicating that the arc is referencing some other device called "NP,"one has elevated the descriptive power of the device in Figure 4 to that of a context-free grammar. This shorthand would optimally be needed if multiple references to NPs were likely. Such a situation *is* likely in a large network describing the entire grammar of a natural language. The device in Figure 4 is best described as an RTN, not an FSM.

4 Parsing and P2 Clitics

The advantage of clitics of the type of the English possessive--what Zwicky 1977 would refer to as simple clitics--is that they attach to specific phrasal types. One can express this easily in a context-free grammar, by specifying the phrasal category as the host for the clitic. P2 clitics are a little more problematic, since they do not bind to constituents of a particular phrasal type. Rather they bind to whatever object is occupying first position. If one examines the Yaqui data in (12)-(14), one will note that the subject clitic *ne* attaches to the verb in (12), to a temporal adverb in (13), and to the object noun phrase in (14).

(12) koche=**ne**
sleep:IMPERF=1sg
'I am sleeping.' (Jelinek, p.c.)

(13) tuka=**ne** tekipanoak
yesterday=1sg work:PERF
'I worked yesterday' (Escalante 1990)

(14) uka ili'uusita=**ne** aniak
DET:ACC child-ACC=1sg help:PERF
'I helped the child' (Escalante 1990)

How should one devise a parsing strategy for P2 clitics? The strategy that immediately presents itself is to create separate rules for each possible phrasal host. For the above examples, we would need rules or devices for verbs, adverbs, and noun phrases, as shown in Figure 5.

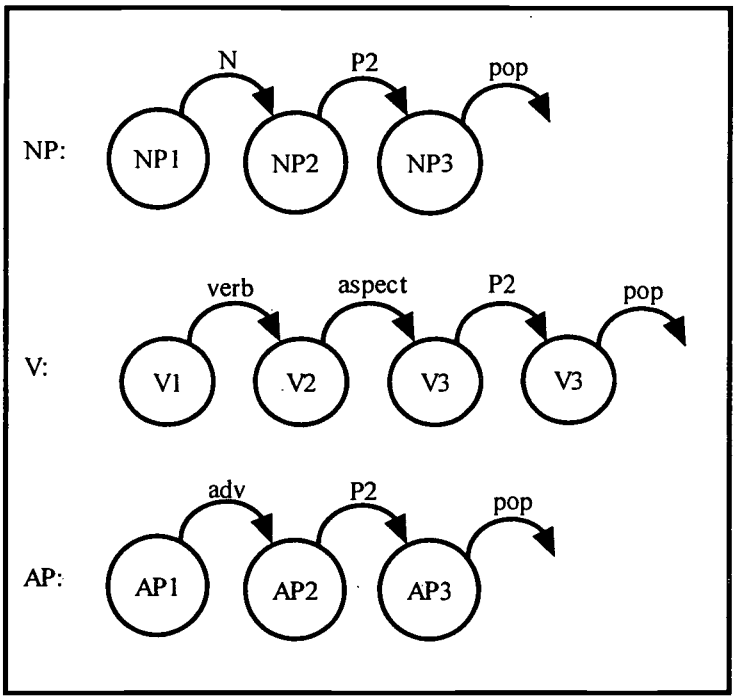

Figure 5 - P2 Parsing Devices

There is a problem with this approach, however: Since P2 clitics can attach to any host occupying first position, a large number of rules would be necessary to describe all potential first position hosts, one for each potential host. In addition, more rules would be necessary to accommodate all instances where the potential hosts do not occupy first position, effectively doubling the number of rules required. Although the number of rules are finite, a more optimal solution (by the second criterion of optimality described in (1)) would be one which could collapse these rules, the purpose of which is to parse the P2 clitic, into a smaller set.

What is really needed is a device similar to that shown in Figure 6, where X_1 refers to any phrasal constituent occupying 1st position. The problem with this approach is finding a method to integrate it into the set of devices describing the entire grammar, such as the devices in Figure 7 (next page). Figure 7 shows the set of devices that are designed to parse canonically ordered Yaqui sentences. How can both the device in Figure 6 and the devices in Figure 7 be active at the same time?

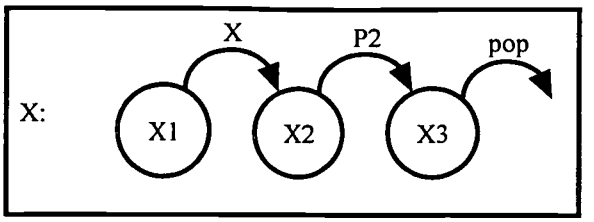

Figure 6 - Second Position RTN

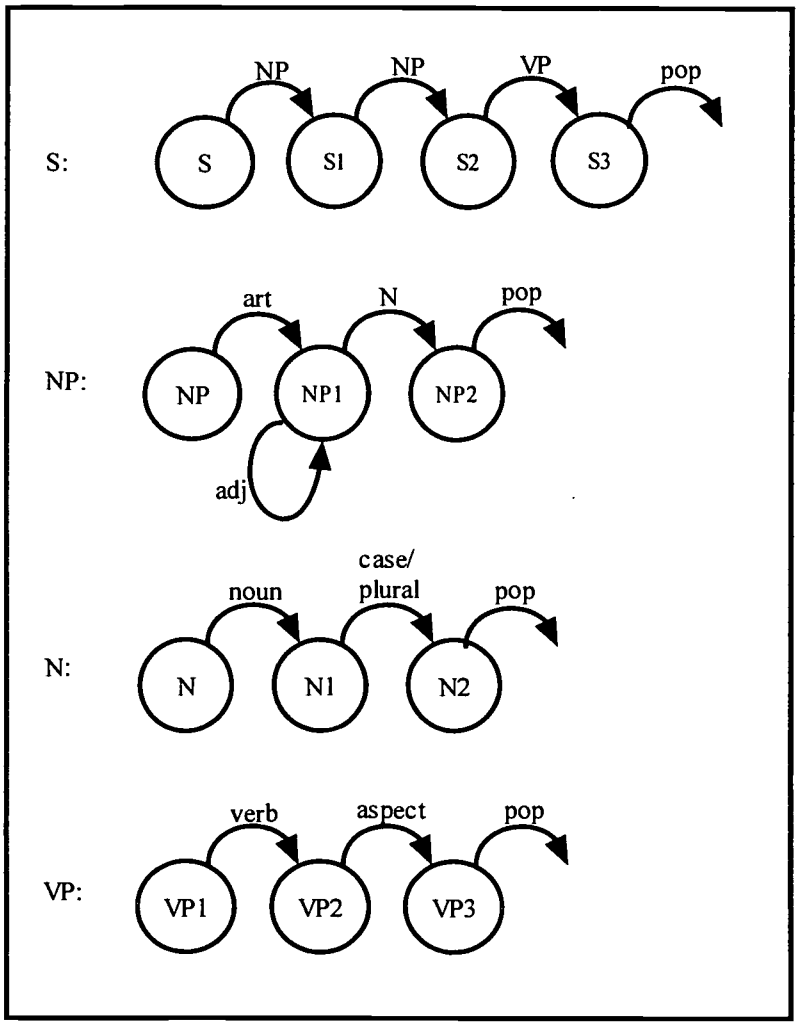

Figure 7 - A Simple Yaqui Parser (Canonical)

Since the P2 clitic is position sensitive, another approach would be to devise context-sensitive devices. Such devices could be employed at an initial phase of the parsing process, their sole purpose being to "unpack" the P2 clitic. The output of such a device would be the input to another device (as those in Figure 7) which performs the actual parse.

What type of device would have the descriptive power of a context-sensitive grammar? Mentioned briefly in the Definitions section of the paper is the device called the *Linear Bounded Automaton (LBA)*. From Hopcroft and Ullman 1979, an LBA is defined as a nondeterministic

Turing machine with the expressive power of context-sensitive grammars. The LBA satisfies the following two conditions:

1. Its input alphabet includes two special symbols ¢ and $, the *left* and *right endmarkers*, respectively.
2. The LBA has no moves left from ¢ or right from $, nor may it print another symbol over ¢ or $.

The ¢ and $ symbols essentially define the domain over which the device can operate. An LBA that transduces the string ABC into the string BAC would have the form shown in Figure 8 (next page).

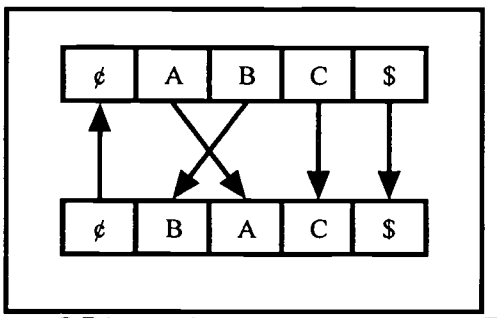

Figure 8 Linear Bounded Automaton (LBA), shown as a transducer

In the case of parsing P2 clitics, the LBA's domain would be minimally confined to the first and second positions of the input string. As discussed previously, we would want such an LBA to generate an output which would serve as input to the parsing device. This output should have the form which most closely resembles the canonical ordering of the language being processed (SOV in Yaqui). In Yaqui, since P2 clitics are always subject clitics (the full list of Yaqui subject clitics is shown in (15)), these clitics should always be transduced into first position. Figure 9 (next page) shows an LBA designed to remove the clitic from its host, a process which I refer to as *unpacking*. I have called this LBA a *Clitic Unpacker*. In this figure, X refers to any constituent, S-c refers to the P2 subject clitic, and Y and Z refer to all succeeding constituents (if any) within the domain of the LBA.

(15)

Singular	Plural	(Escalante 1990)
1. =ne	=te	
2. ='e	='em	
3. =∅	=m	

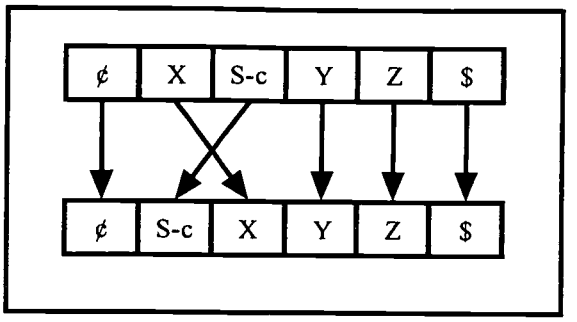

Figure 9 - P2 LBA, Clitic *Unpacker*

Applying this LBA to the sentence in (16) would generate the "sentence" in (17) (X = *koche*, and P2 = *ne*). Of course, the device accepting this string as input would need to recognize the clitic *ne* as a special object: clitics cannot exist in Yaqui detached from their hosts. Prefixing the clitic with "=" in the output will be used to indicate its special status (as shown).

(16) kochene
koche=**ne**
'I am sleeping'

(17) =**ne** koche

5 Constraining X and P2 - Parametric Variations

There are two problems in designing the P2 LBA: characterizing the form of X (the object in first position), and characterizing the form of the P2 clitic. First, let's consider the various forms that X can take in Yaqui. (12)-(14) are repeated here for convenience. Example (18) is a variant of (14), where the clitic is imbedded within the object NP.

(12) koche=**ne**
sleep:IMPERF=1sg
'I am sleeping.'

(13) tuka=**ne** tekipanoak
yesterday=1sg work:PERF
'I worked yesterday'

(14) uka ili'uusita=**ne** aniak
DET:ACC child-ACC=1sg help:PERF
'I helped the child'

(18) uka=**ne** ili'uusita aniak
DET:ACC=1sg child-ACC help:PERF

Without (18), one would be tempted to assume that X would refer to the first phrasal constituent, and the physical host of the P2 clitic would be the final lexical constituent of that phrase. Including (18), however, one must also allow for the potential for X to refer to the initial lexical constituent of that first phrase as well. Since some languages also allow enclisis (infixation of a clitic within its host), the form of X can be seen to vary parametrically across a continuum with three distinct realizations, as shown in Figure 10. For any given language, the potential value of X includes the entire set of values to its right, inclusive of the starting point. For instance, since Yaqui accepts W_1+P2 (the first word as the attachment site for the clitic), it also accepts XP_1+P2 (the first phrase as the attachment site for the clitic). The Clitic Unpacker LBA for languages far to the left must be progressively more flexible than those to their right.

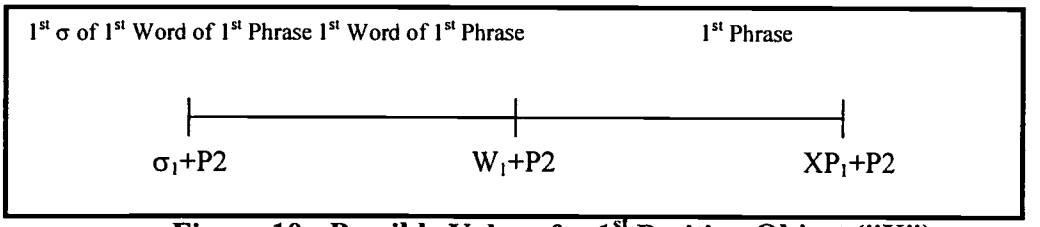

Figure 10 - Possible Values for 1st Position Object ("X")

The form of the clitic varies cross-linguistically, as well. In the case of Yaqui, the type of elements that can be cliticized varies. Many languages, however, allow multiple constituents to be cliticized simultaneously. In such cases, interestingly, the placement of these clitics with respect to one another is strongly regimented. This is especially intriguing if one considers languages in which the word order is normally free.

Lummi, a Straits Salish language of the Northwest Coast of North America, is a language that has so-called *clitic clusters*. The full inventory of clitics is shown in (19)[7]:

(19)

a. Tense clitics
   =sə' Future
   =lə' Past

b. Modal clitics
   =yəq
   =yex^{w}
   =č'ə'
   =q

c. Mood clitic
   =ə

d. Subject clitics (NOM, local)

=sən	'I'
=sx^{w}	'you'
=ɬ	'we'
=sx^{w}helə	'you pl.'

[7] Jelinek 1996 is the source for all Lummi data shown in this paper.

Clitic clusters in Lummi occupy second position. Again, as in Yaqui, the host of the clitics is whatever object is in first position. However, unlike Yaqui, multiple clitics can occur in the second position simultaneously. Within this clitic cluster, clitics are ordered strictly according to the paradigm in (20):

(20) =Q/MODAL=TENSE=SUBJECT[8]

Some examples from Lummi, showing the P2 clitic cluster in several environments, are shown in (21)-(25). Example (21) shows a single clitic attached to an adjective. Example (22) shows a two member clitic cluster, consisting of the past tense clitic followed by the 2nd person singular nominative clitic, attached to the predicate. Example (23) shows the verb followed by a three-member clitic cluster, consisting of the question clitic, the past clitic, and the clitic for 2nd person singular nominative. Example (24) is similar to (23). Example (25) shows a noun followed by the future and 2nd person singular nominative clitics.

(21) ɬčikʷəs=ɬ
tired=1pl:NOM
'We are tired.'

(22) čey=lə=sxʷ
work=PAST=2sg:NOM
'You worked.'

(23) leŋt-oŋəɬ =ə=lə'=sxʷ
see-1pl:ACC =Q=PAST=2sg:NOM
'Did you see us?'

(24) kʷəniŋt-óŋəɬ =ə=sə'=sxʷhelə
help-1pl:ACC =Q=FUTURE=2pl:NOM
'Will you (pl.) help us?'

(25) si'em=sə'=sxʷ
chief=FUT=2sg:NOM
'You will be a chief.'

Designing a device to unpack and parse the Lummi clitic string is a little more complicated than the equivalent task in Yaqui. First, it must be determined to what extent unpacking is necessary. Second, if unpacking is necessary, a description of the device to unpack the string is required. Third, since the clitic string in Lummi can contain multiple clitics, this device, or some other, needs to parse and check the constituency of the clitic cluster (to test for grammaticality, if nothing else).

[8] =Q/MODAL indicates that a question or modal clitic can occupy this position in the cluster. Because they occupy the same position in the clitic cluster, a Lummi clitic cluster cannot contain both a question and a modal clitic simultaneously.

Does the clitic cluster in Lummi need to be unpacked? To answer this question, one must determine what the canonical constituent ordering is in Lummi. If the order is rigidly enforced and consistent in the local domain of the cluster, than no such unpacking algorithm is necessary. Although constituent order is much more rigid in Lummi than in Yaqui, there is some flexibility in the category of the object that can occupy first position.[9] In (22)-(24), a verb is the host; in (21), an adjective is the host; and in (25), a noun is the host. Assuming constituent order is roughly VOS[10] (note that I did not indicate "word" order, since the constituents referenced here are not words), the flexibility in the choice of 1st position constituent argues for some form of a clitic unpacker.

But what form should this clitic unpacker take? In Yaqui, the design of such a device was much simpler, since the clitic string consists of one clitic, which represents the subject. Since Yaqui is subject initial, the LBA was designed to ensure that the *subject* clitic was moved from second position to first. In Lummi, the clitic cluster is far more complex, and in many ways, the cluster itself dictates the constituent ordering. For (22)-(24), the surface order of the constituents, including the host verb, appears to match the canonical VOS order. (In these cases, transitivity of the verb dictates the presence or absence of the object. We ignore the tense and question clitics, for the moment.) What proves to be problematic are examples (21) and (25), where there is no overt verb (and, depending on one's definition of the copula and choice of syntactic theory, possibly no covert verb either).

The choice may require defining two devices to unpack the clitic cluster, one to unpack sentences of the type in (22)-(24), and one to unpack sentences of the type in (21) and (25). The templatic structure of the clitic cluster can be used to these devices' advantage, by making the assumption that all slots have values. This seems warranted and justified if one considers (21), which has no overt tense. In this case, the null value of this slot can be assumed to have a default tense value: present. The LBA in Figure 11 acts as an identity device, mapping the constituents to the same relative places in the output, confirming the canonical order.[11] (Tense was moved next to the verb, since this seemed an appropriate place for a verb "related" object.)

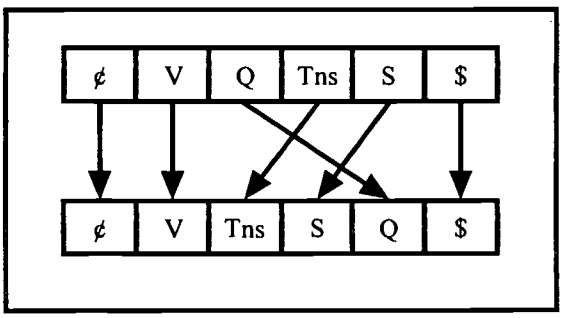

Figure 11 - Lummi Clitic Unpacker #1 (Identity)

[9] Although, arguably, whatever occupies the first position is the predicate (Jelinek, p.c.).

[10] This constituent order was derived from the examples containing overtly spelled out verbs, namely (21)-(24), most specifically, (23) and (24). Since the object is a suffix of the verb, it occupies a position closer to the verb than the subject, which surfaces as a clitic.

[11] It is unclear what should be done with the question particle, so it was placed at the end. If the question particle were ignored by this device--mapped to the same position in the output--then this device could easily be expressed as an FSM.

The LBA in Figure 12 acts as a clitic unpacker, placing the tense and subject clitics to the front of the output string. In this device, X represents any object but a verb. The succeeding device, whatever form it takes, would necessarily need to interpret the nascent tense clitic as an indication of the presence of the copula.[12]

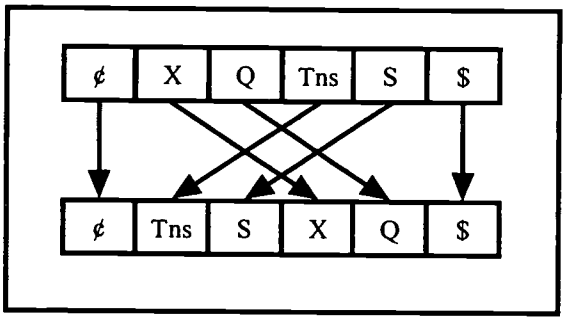

Figure 12 - Lummi Clitic Unpacker #2

6 Scrambling and Parsing

The purpose of the Yaqui Clitic Unpacker, an LBA shown in Figure 10, is to unpack the clitic from its host and make it and the other material in the sentence available to whatever devices follow. This Clitic Unpacker, whatever its expressive capacity, would need to parse successfully a large number of structures, and to generate structures similar to those shown in (26)-(28) (these examples are *not* Yaqui, but Yaqui after being processed by the Clitic Unpacker; the host "unnaturally" follows the clitic in these examples):

(26)	='e	uka	kava'i-ta	nee	miika-k
	=2sg	DET:ACC	horse-ACC	1sg:DAT	gave-PERF
	'You gave me the horse.'				
(27)	='e	miika-k	uka	kava'i-ta	nee
	=2sg	gave-PERF	DET:ACC	horse-ACC	1sg:DAT
(28)	='e	nee	miika-k	uka	kava'i-ta
	=2sg	1sg:DAT	gave-PERF	DET:ACC	horse-ACC

[12] This analysis hinges on the assumption that there exist distinct categories of nouns and verbs in Lummi. Jelinek 1996 argues against this: "A persistent problem in the analysis of the Salish languages [of which Lummi is a member] is the question of whether they show a contrast between *noun* and *verb* as lexical categories.... While it is possible to separate Salish roots into various classes, including nominal roots, on the basis of their lexical features, roots and the predicates derived from them are not associated with distinct maximal projects such as NP or VP. There is no copula (in any paradigm) to derive sentences with a 'predicate noun' or adjective; thus we can say that every sentence contains a verb only if we call all words (except the closed list of particles and adverbs) 'verbs'." If we make this assumption about Lummi and apply it to the design of a parser, then the canonical and surface orders shown in (21)-(24) are identical. In this case, no unpacker would be necessary. In fact, a parser for such a language could be described with an FSM. I will not adopt this argument in this paper.

However, the output of the Clitic Unpacker is only an intermediary step in the parsing process; the Unpacker does not parse its input, it only unpacks the clitic. What of the device or devices that parse its output? What is the their nature, and what is their expressive capacity? Initially, it would appear that the best approach would be to design a device that could descramble these structures into their canonical forms. The device would then be able to take the structures represented in (26)-(28) and output the "underlying" form of SOIV (where *I* represents the indirect object). Figure 14 (next page) shows a device which takes output similar to (26)-(28) and generates an SOIV output. This device consists of a set of three integrated LBAs, which, combined, I call a *Descrambler*:

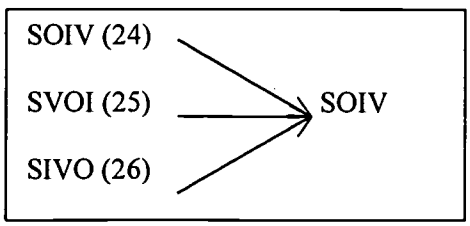

Figure 13 - Descrambler

But, is such a descrambler necessary? Could the structures shown in (26)-(28), and all other potential structures, be successfully parsed by a device whose expressive power is context-free, rather than context-sensitive? For example, the RTNs shown in Figure 15 would successfully be able to parse the structures in (26)-(28). Granted, a large number of such devices would be necessary. However, the same would be true if one were to choose context-sensitive devices instead (there are three integrated LBA in Figure 14). In fact, the number of devices would be the same in *either* paradigm. That being the case, the context-free paradigm--being more optimal--would win.[13]

[13] Although the form of the output has been mostly ignored in this paper (as it was indicated it would be), it should be noted that its form might have some bearing on the type of grammar chosen. If one wished to have an output whose structure was hierarchical and consistent from parse-to-parse, a context-sensitive grammar would seem appropriate: it could freely match input words and phrases to nodes in the output, whatever form these might have. For example, if the output were to have the structure $[_S\ [_{NP}\ S]\ [_{VP}\ O\ V]]$, a context-sensitive device would more readily be able to parse the various scrambled structures--SOV, SVO, OSV--into the output structure, whereas a context-free device would not.

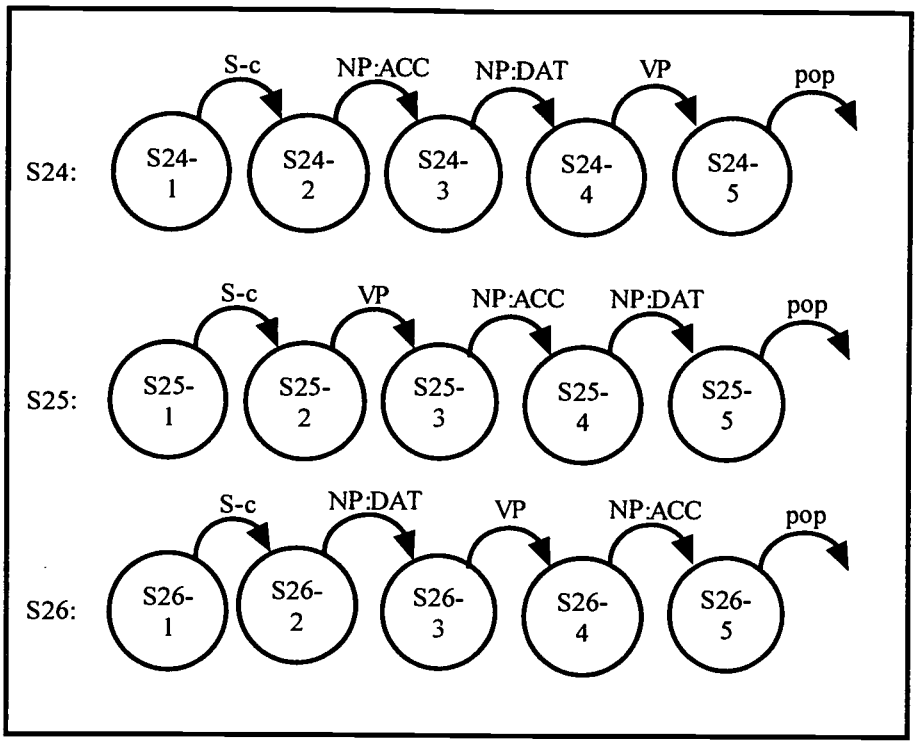

Figure 14 - RTNs to Parse the Scrambled Structures (24)-(26), *Scrambled Parser*

7 P2 Clitics, Scrambling and Parsing

In section 6, we devised an algorithm to parse scrambled sentences. The context-free *Scrambled Parser*, being more optimal, was the winner over the context-sensitive *Descrambler*. In section 5, we discussed devices to unpack clitics, so-called *clitic unpackers*. Are both types of devices necessary? For Lummi, arguably no, since the clitic unpacker for Lummi *is* a descrambler.[14] For Yaqui, the question is a little more complicated.

To unpack the clitic in Yaqui, a rule of the type XC → CX was used, where C corresponds to the clitic, and X any first position constituent. Since such a rule is context-sensitive, the power of an LBA is required. Could such an LBA be integrated into a parsing device of the type discussed in section 6? At first glance, the answer would be "no", primarily because of the general nature of the rule XC → CX, *versus* the more specific rules in Figures 13 and 14. Before accepting this conclusion, however, the question should be a little more carefully analyzed.

Consider for a moment the "unpacked" sentences shown in (26)-(28), shown here again in (29)-(31) in their unaffected forms:

(29)	uka	kava'i-ta='e	nee	miika-k
	DET:ACC	horse-ACC=2sg	1sg:DAT	gave-PERF
	'You gave me the horse.'			

[14] Based on the discussion in section 6 and here, one might question whether the power of context-sensitivity shown for the Lummi clitic unpacker in section 5 is necessary. It might be possible to formulate devices with the more limited power of context-free grammars.

(30)	miika-k='e	uka	kava'i-ta	nee
	gave-PERF=2sg	DET:ACC	horse-ACC	1sg:DAT
(31)	nee='e	miika-k	uka	kava'i-ta
	1sg:DAT =2sg	gave-PERF	DET:ACC	horse-ACC

The clitic unpacker corresponding to the rule XC → CX was devised as a solution to convert any sentence with a second position clitic into one with the subject clitic fronted. By so doing, we eliminated specific phrasal or lexical rules needed to unpack the clitic for each category, rules of the form NC → CN, VC → CV, etc. However, considering the rather large number of rules required to parse scrambled structures, it should be questioned whether we really have gained anything at all. Figure 14 shows a descrambler designed to parse the already unpacked structures in (24)-(26). Now consider the descrambler in Figure 15 (where lower case *s* denotes the subject clitic). This device (or set of devices) could descramble the sentences in (29)-(31), the un-unpacked counterparts of (26)-(28). (For the moment, ignore the optimality of this device. A more optimal context-free device could be devised.) If we posit such a device, then there is no need for the Clitic Unpacker described earlier.

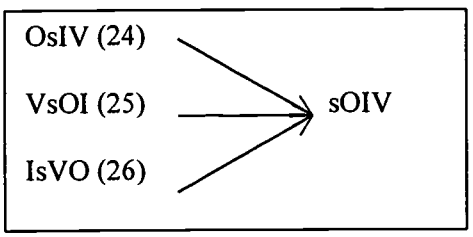

Figure 15 - Descrambler II

This conclusion, however, may be hasty. The rules shown in Figure 15 treat the clitic as a phrasal affix. Although generally considered phrasal affixes (see Anderson 1992 for discussion), clitics do have a high degree of flexibility cross-linguistically as to how they attach to their hosts. The rule XC → CX has the advantage because X can represent *any* first position object: the first phrase, the first word, or the first stressed phonological unit (syllable). If the interpretation is so commonly varied, then the rules described in Figure 15 would need to be expanded to accommodate all the possible clitic hosts, an increase in the number of rules that the rule XC → CX prevents. The clitic unpacker as a *separate* layer does appear fully warranted, after all.

Let us return to Lummi. As described, the clitic unpacker for Lummi *is* a descrambler/parser. A separate layer, as described for Yaqui, appears unwarranted However, if one extends the principle of optimality to include solutions that are more general--a definition not proposed in this paper, but one compatible with my definition--then one would want a solution whose structure is similar cross-linguistically. Mirroring the solution for Yaqui, one could propose for Lummi a parser that includes a layer devoted to clitic unpacking, and another layer devoted to parsing. Although such a solution would not be locally optimal--by proposing an additional set of rules that would be unnecessary for Lummi, violating the second criterion of

optimality--one would adopt a solution that is generally optimal, since it will work cross-linguistically. By *not* adopting such a definition of optimality, one would be forced to design a parser for Yaqui that would look and behave very differently than its counterpart for Lummi.

8 Concluding Remarks

The purpose of this paper was to devise optimal algorithms for parsing P2 clitics. Optimality was defined by the two criteria shown in (1), repeated here for convenience:

(1)

1) The optimal solution is one which uses devices and formalisms whose generative capacity is as low as possible on the Chomsky hierarchy.
2) The optimal solution uses as few "rules" or "devices" as possible.

Although the solution to unpack P2 clitics, generically described by the context-sensitive rule XC → CX, is quite appealing and optimal, no solution would be possible without parsing the scrambled structures in which P2 clitics are likely to occur. Appealing to context-sensitive devices such as LBAs to descramble such structures offers some potential solutions which may decrease the total number of devices or rules required for such a parse. The structure of a parser might require "layers" of devices, each layer feeding devices in the next, until finally parsed output is generated. An example design is shown in Figure 16.

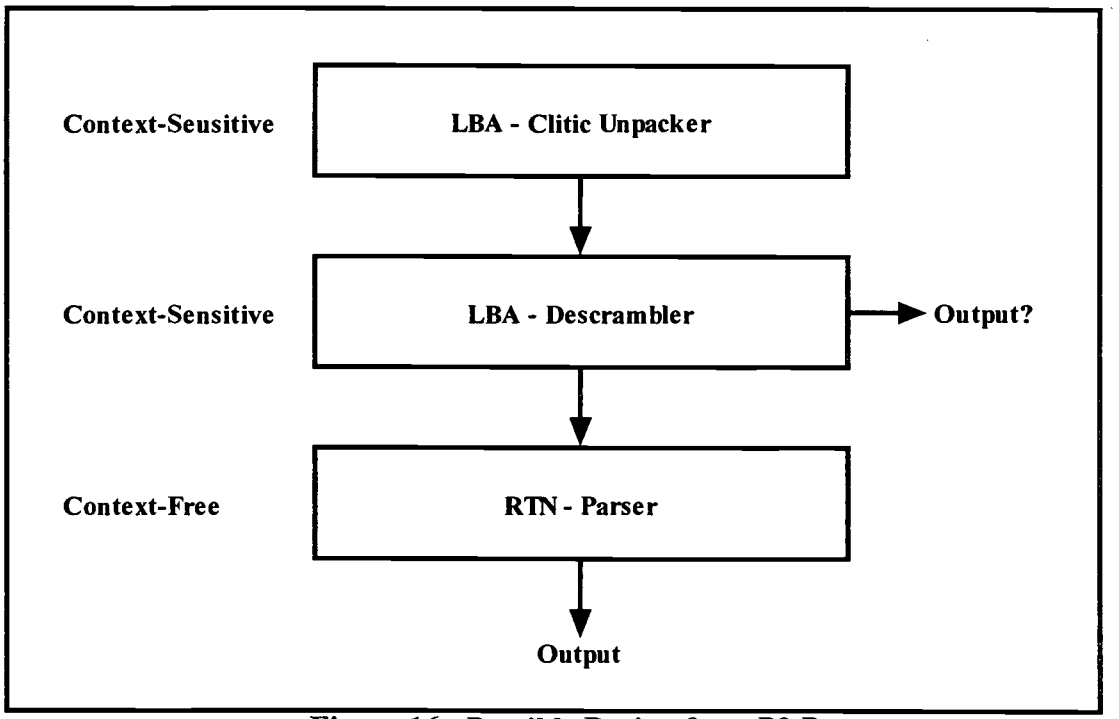

Figure 16 - Possible Design for a P2 Parser

However, appealing to a definition of optimality that is so rigidly constrained by the Chomsky Hierarchy may not be sufficient enough to devise the most efficient and universal algorithms. In the previous section an amendment to the definition of optimality was proposed. This definition proposes that the most optimal solutions are those that operate unchanged cross-linguistically. Of course, this new criterion would compete with the other two criteria proposed, in some ranking yet to be defined.

But perhaps there is more required for a definition of optimality. Could the rigid placement of the P2 clitic, as well as the rigidly enforced ordering of constituents within the P2 clitic cluster itself, offer some insight into a solution? Perhaps the introduction of interrupt-driven devices--devices designed to interrupt the normal flow of a parse--for the purpose of handling unexpected input, might provide for more efficient solutions. Using such devices could allow the parser to parse scrambled structures, while simultaneously allowing it to parse P2 clitic clusters. The added advantage is that the parser would only interrupt the flow of the parse when and if P2 clitics are present within a given structure. Interrupt-driven devices are not easily described by the levels of the Chomsky hierarchy. Other solutions, however, might involve defining meta-grammars or rule schemas, which themselves define or interact with the rules of a grammar. These *do* fit within the confines of the Chomsky hierarchy. The optimality of such meta-grammars would be determined either by the grammar generated or by the meta-grammar, depending on which is lower in the Chomsky hierarchy (Langendoen 1976).

It is obvious that parsing P2 clitics and scrambled structures is a difficult problem. This paper introduces the reader to the complexity of the problem, and offers some solutions. Only a very small set of solutions are discussed. Additional research is required to truly resolve the problems presented.

References

Allen, James. 1995. *Natural language understanding*. Redwood City, CA: Benjamin/Cummings Publishing Company.

Anderson, Stephen R. 1992. *A-Morphous morphology*. Cambridge, UK: Cambridge University Press.

Choi, Hye-Win. 1996. *Optimizing structure in context: The case of scrambling*. Doctoral dissertation proposal, Stanford University.

Escalante, Fernando. 1990. *Voice and argument structure in Yaqui*. Doctoral dissertation, University of Arizona, Tucson.

Hammond, Michael. 1999. Personal communication.

Hopcroft, John, and Jeffrey Ullman 1979. *Introduction to automata theory, languages, and computation*. Reading, MA: Addison-Wesley Publishing Company.

Jelinek, Eloise. 1996. Definiteness and second position clitics in Straits Salish. In *Approaching second*, ed. A. Halpern and A. Zwicky. Stanford, CA: CSLI Publications.

Langendoen, D. Terence. 1976. On the weak generative capacity of infinite grammars. Paper presented at CUNYForum, CUNY, New York City.

Newman, Paula S. 1996. *Representing 2P clitic placement in LFG*. Stanford, CA: Xerox PARC. Web: http://www-csli.stanford.edu/publications/LFG/newman.ps

Maxwell, John T., and Christopher D. Manning. 1998. *A theory of non-constituent coordination based on finite-state rules*. Stanford, CA: Xerox PARC. Web: http://www.sultry.arts.usyd.edu.au/cmanning/papers/coordination.ps

Russell, Stuart, and Peter Norvig. 1995. *Artificial intelligence: A modern approach.* Upper Saddle River, NJ: Prentice-Hall.

Zwicky, Arnold. 1977. *On clitics*. Bloomington, IN: Indiana University Linguistics Club.

Department of Linguistics
University of Arizona
200 E Douglass Bldg.
Tucson AZ 85721

wlewis@u.arizona.edu

Mohave Language Planning: Where has it been and where should it go from here?*

Jessica P. Weinberg *Susan D. Penfield*

The Colorado River Indian Tribes (CRIT), in Parker, Arizona, include four Native Arizonan tribes, Mohave, Chemehuevi, Navajo, and Hopi. These tribes function politically as a unit, although they are distinct in terms of language, culture, and history. While all Native American languages are endangered today, for two of the Colorado River Indian Tribes, the Mohave and Chemehuevi, the language situation is critical. In this paper, we will be concerned only with language planning as it relates to Mohave. As a background for the current language planning situation for Mohave, we briefly discuss the history and current circumstances of the CRIT reservation. We provide a short history of linguistic work on Mohave, we discuss current language planning efforts focused on Mohave, and finally, we make recommendations for continued language preservation and revitalization of Mohave.[1] We conclude that language planning on the CRIT reservation must involve efforts focused on each of the four tribal languages as well as the blending of language planning efforts for all four CRIT languages to reflect the integrated social reality of the CRIT.

1 History and Current Linguistic Situation of the CRIT

The uniqueness of the CRIT reservation lies in the fact that the four tribes function politically as one unit. The history that contributes to this situation involves the collocation of two tribes indigenous to the area, Mohave and Chemehuevi, and members of two other tribes, Hopi and Navajo, who relocated to this reservation following World War II.[2] The Tribal Council of the CRIT reservation is composed of members of all four tribes. As of 1999, the total membership of the California River tribes, both on and off the reservation, is 3,403, Mohaves being the most numerous members.[3]

Mohave language use has been heavily affected by the Mohaves' history of schooling. In 1879, the first Colorado River boarding school was opened near what is now the town of Parker. Prior to that, Mohave students had been sent to boarding school at Fort Mohave in Needles, California. The boarding school experience had a devastating effect on native language use. Attempts to "civilize" and educate Indians entailed severe English-only policies, realized in the physical and verbal punishment of Indian students who spoke their native language. As a result, native language use is limited today. There are only approximately 30 speakers of Mohave on the CRIT reservation.

It is important to recognize that the four CRIT languages are each at different stages of language shift and language planning. At the moment, language loss is most critical for

* We would like to thank Amelia Flores for her invaluable assistance with this paper.

[1] We recognize that there are language planning efforts underway among the Mohave at Fort Mohave in Needles, California. We are primarily concerned here with language planning as it relates to Mohave members of the CRIT reservation.

[2] Of course, the Hopis and the Navajos each also have their own reservations separate from the CRIT reservation.

[3] At the time of this writing, specific data on individual tribal membership is not available.

Chemehuevi, with an estimated 13 speakers remaining. However, the Chemehuevis have been actively pursuing language planning. Mohave is the next most endangered language on this reservation. While it is known that there are many more speakers of Hopi and Navajo, particularly on their reservations in northeastern Arizona, these languages are still both considered endangered. We have no estimates at this time regarding the number of speakers of Hopi and Navajo at the CRIT reservation.

2 A Linguistic History of Mohave on the CRIT Reservation

Linguistic study of the Mohave language on the CRIT reservation may be divided into two domains: 1) that which was initiated and conducted by outside researchers (non-tribal members) and 2) that which was initiated and conducted by tribal members.

2.1 Outside Research

The earliest work on Mohave by a non-tribal member was that done by A. L. Kroeber in 1911 ("The Phonetic Elements of the Mohave Language"). Kroeber's extensive work on Mohave language and culture still stands as a valuable and accurate record. J. P. Harrington is known to have worked on both Mohave and Chemehuevi, but his work has never been published. Several narrative texts in Mohave, with some accompanying analysis, were collected by Susan Penfield in 1969. None of this work has been published to date, but it remains on file in the CRIT Library/Archive. Subsequent work on the language by outsiders has been largely focused on issues involving syntax. Judith G. Crawford published a discussion of the Mohave verb 'be' in 1976 and also published several narrative texts which she had collected in Mohave around that same period of time. The most extensive linguistic contribution to the study of Mohave has been made by Pamela Munro. Her doctoral dissertation, *Mojave Syntax*, was published in 1976. She has published numerous articles on various aspects of Mohave grammar, and she culminated her work with the publication of a Mohave dictionary (with Nellie Brown and Judith Crawford) in 1992.

2.2 Research by Mohave Tribal Members

Linguistic research and planning efforts by tribal members have taken different forms over the years. After Pamela Munro's work in the 1970s, and with her help, Mohave language classes were formed. These were taught by Luther Swick, with materials prepared by Munro. A very talented native linguist, Edward Swick, also worked on Mohave. He wrote extensively about the language and published some of his work in the tribal newspaper, *Smoke Signals*, in the 1970s and 80s. Recently, there have been two notable linguistic efforts, specifically language planning efforts, by tribal members. These language planning efforts will be discussed in detail in the next section.

3 Current Mohave Linguistic Work

The CRIT Library/Archive has recently sponsored two grant proposals for language-related projects which involve Mohave. One project is a Mohave coloring book for children.[4] The proposal for this grant was written by Amelia Flores, head of the CRIT Library/Archive, for funding under the Arizona Department of Library, Archives and Public Record Library Services and Technology Act. The Mohave received this grant, and are currently finishing the coloring book. The second proposal was written by Amelia Flores, Mohave, and Bonita Fernandez, Chemehuevi, to the Administration for Native Americans (ANA), a federal granting agency responsible for funds for the Native American Languages Act. The ANA proposal was for a project to assess the needs of the four Native languages that make up the CRIT, and to develop a plan to meet those needs.

We can see in these proposals two different levels of language planning that are being undertaken by the CRIT. On the one hand, the CRIT have written small grants for focused projects, such as the one received for Mohave and Chemehuevi coloring books. On the other hand, large-scale projects such as the one proposed to the ANA must involve all four tribes because of the governmental structure of the CRIT. In addition, language preservation and revitalization efforts must at some level involve all four tribes because of the reality of the multi-ethnic and multi-lingual character of the CRIT reservation. This is contrary to Fishman (1991), who argues that language revitalization efforts require ethnic insulation.

3.1 The Mohave Coloring Book Project

The Mohave Coloring Book project has taken shape through working sessions among a small group of Mohave elders, who are native speakers of Mohave, with assistance from Amelia Flores, head of the CRIT Library/Archive. Though Amelia is not a native speaker of Mohave, she is keenly interested in her heritage language and has been a driving force behind much oral history and language-related work by the Mohave for many years. The coloring book group decided on a list of words to include in the coloring book, and with the help of the Brown-Munro-Crawford Mohave Dictionary, generated Mohave equivalents for these words.

Susan Penfield, who had been written into the coloring book grant proposal as a consultant, was asked to help with a few areas of discrepancy between the elders' knowledge and the Brown-Munro-Crawford Dictionary. Susan invited her student Jessica Weinberg to come to Parker with her for a two-day trip that would involve a coloring book working session.

The coloring book working session was attended by Susan, Jessica, Amelia, and three elders whose first language is Mohave (Leona Little, Joe Sharp, and Ione Dock). We taped the three-hour working session for two reasons. First, taping provided us with an extra record of the decisions made in the session, along with our notes. Second, we believe that language-related working sessions involving groups of native speakers of the language(s) of interest yield many useful types of data other than simply the Mohave-to-English word equivalencies needed for the coloring book. As argued by Penfield and Flores (1999), these other types of data include Mohave conversational and nonverbal interactional data, expanded lexical data (for example, other than simply the items in question for the coloring book), code-switching data, evidence of beginning Mohave literacy, and attitudes, or ideologies, of Mohaves about the Mohave language.

[4] The project also includes a Chemehuevi coloring book.

This last type of data is crucial for language planning efforts because most of these efforts depend on either accommodating or working to change existing language attitudes of a community. We believe, following Ruiz (1995), that language planning efforts should ideally originate from within the community. However, language planning efforts, either from within or outside the community, must be supported by positive attitudes from the community toward those efforts and toward the language that is the target of those efforts (Fishman 1991).

While the initial idea for the coloring book was to make it bilingual in Mohave and English, we recommended the coloring book be in Mohave only. We based this recommendation on the insight from bilingual education that when material is presented in two languages (a method known as "concurrent translation"), one which the students are more familiar with than the other, students will focus on the familiar language and tune out the less familiar language (Crawford 1995:125). The current plan is to publish the coloring book as Mohave-only, except for explanatory material in English in the back of the book, as a kind of appendix.

We recommend that the tribe undertake more small projects like the coloring book (such as more coloring books with different themes, readers for young children, books of children's songs with accompanying tapes and/or CDs, interactive CDs for children, etc.) for several reasons. First, grants for such projects are generally easier to apply for and to receive than grants for long-term projects. Second, such projects involve a manageable amount of work in a relatively short period of time, and they gratifyingly result in a tangible product. Third, it is our hope that the increased availability of Mohave language materials (e.g., books, CDs, etc.) will generate more interest in the preservation and revitalization of Mohave among younger generations of Mohaves, in particular children and their parents. Finally, Mohave language materials are necessary to facilitate Mohave language school curricula. This brings us to the issue of the multi-ethnic, multi-lingual nature of the CRIT, which poses challenges for curricular implementation. This issue is addressed in part by the second project we discuss here, the ANA grant being developed by the CRIT.

3.2 Grant Proposal to the Administration for Native Americans

As we noted above, the CRIT applied for an ANA grant to implement a project to assess the needs of the reservation's four Native languages (Mohave, Chemehuevi, Navajo, and Hopi) and to develop a plan to meet those needs. The application was not successful the first time the CRIT submitted it. However, they are revising the application with plans to resubmit it in the spring of 2000. Our primary recommendation for the revision of the proposal involves the Language Committee to be established in the course of the project. The committee proposed by the ANA grant application, as currently written, would consist of at least two members of each tribe. As emphasized by the wording of the grant application, the four tribal languages differ greatly in status and needs. Therefore, we suggest that one committee for all four languages may be insufficient. We recommend the grant proposal include the establishment of four separate language committees, one for each of the four languages, plus a coordinating committee with two representatives from each language committee.

In addition, we recommend that the grant proposal include a specific commitment to recruit younger generations, especially parents of children, to sit on these committees. There is an urgent need to generate interest among these generations in having their children learn their ancestral language, and in learning it themselves. Fishman (1991) argues that the most important

piece of the puzzle for reversing language shift is to ensure "intergenerational mother tongue transmission," that is, children acquiring their ancestral language as their first language from their parents and grandparents. The challenge, as we noted, is in finding a way to facilitate this transmission of languages in a multi-ethnic, multi-lingual situation such as the CRIT. One possibility for intergenerational transmission that is especially well suited to languages such as Mohave, with only a small number of elderly speakers, is the Master-Apprentice Program (MAP). MAP, in which a native speaker elder is paired for intensive language teaching with a younger friend or relative who wants to learn their heritage language, has been primarily used for Native Californian languages (Hinton 1994). In fact, there has already been a MAP team at Fort Mohave. Implementation of the MAP method is currently being considered for Mohave on the CRIT reservation.

The four Colorado River Tribes cannot be completely insular, as Fishman recommends for communities attempting to reverse language shift, because of the organization of the CRIT government, because of intermarriage among the tribes, and because all reservation children, as well as children from the surrounding town of Parker, Arizona, attend school together. The four languages need separate committees to assess and address their distinctive needs, but the fates of these four languages must be connected because of the social, political, and economic realities of the CRIT. We hope that this connection will be addressed by a coordinating committee consisting of representatives from the four language committees.

In addition, we expect that the situation will require a commitment by the four tribes to work together and be open to innovative solutions, including cross-tribe language teaching. Our expectation is that most of those interested in learning Mohave would be of Mohave heritage. However, Mohave materials cannot necessarily be reserved for Mohave children only, both because there are many children of mixed heritage and because the children of all four tribes must be taught in the same classrooms. While Fishman (1991) is skeptical about the usefulness of schools in facilitating the kind of intergenerational language transmission that is necessary for language revitalization, the schools necessarily play an important role in generating interest among the youngest generations in their heritage languages. In addition, McCarty and Nicholas (1999) argue that schools can in fact be co-opted by language activists for language revitalization purposes.

4 Conclusions

We conclude that language planning on the CRIT reservation must reflect the integrated social reality of the Colorado River Indian Tribes, while still attending to the specific needs of each individual language. To support this goal, we recommend that younger adult generations of Mohaves (and the other Colorado River Tribes) participate in programs such as the American Indian Language Development Institute (AILDI) at the University of Arizona. Participation in such programs allows tribal members to see how other tribes are handling their language planning efforts, which can be a source of ideas and inspiration. Such programs also provide training in linguistic analysis and language and literacy curriculum development, which facilitates self-generated theoretical and applied linguistic efforts that we maintain are crucial to language preservation and revitalization among American Indians.

References

Brown, Nellie, and Pamela Munro. 1984. Mojave texts. In *Spirit Mountain: An anthology of Yuman story and song*, ed. Leanne Hinton and Lucille Watahomigie, 285-288. Tucson: Sun Tracks and University of Arizona Press.

Crawford, James. 1995. *Bilingual education: History, politics, theory, and practice, third edition*. Los Angeles, CA: Bilingual Educational Services.

Crawford, Judith G. 1976a. Seven Mohave texts. *Yuman texts, International Journal of American Linguistics Native American Texts Series 1*, ed. Margaret Langdon, 32-42. Chicago, IL: University of Chicago Press.

Crawford, Judith G. 1976b. The reduction of idu: 'be' in Mohave. *Proceedings of the First Yuman Languages Workshop, Southern Illinois University Museum Studies 7*, ed. James E. Redden, 45-54. Carbondale: University Museum, Southern Illinois University.

Crawford, Judith G. 1978. Coyote and crane (Mohave). *Coyote stories, International Journal of American Linguistics Native American Texts Series, Monograph 1*, ed. William Bright, 121-123. Chicago, IL: University of Chicago Press.

Fishman, Joshua. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon: Multilingual Matters.

Hinton, Leanne. 1994. *Flutes of fire: Essays on California Indian languages*. Berkeley, CA: Heyday Books.

Kroeber, Alfred L. 1911. The phonetic elements of the Mohave language. *University of California Publications in American Archaeology and Ethnology* 10:45-96.

McCarty, Teresa, and Sheilah Nicholas. 1999. What if the children forget the language? Language planning cases and issues in American Indian communities. Paper presented at the 1999 Annual Meeting of the Society for Applied Anthropology, Tucson, Arizona.

Munro, Pamela. 1973. Nominalization and plurality in Mojave. In *You take the high node and I'll take the low node: Papers from the comparative syntax festival, the difference between main and subordinate clauses*, ed. Claudia W. Corum et al., 53-64. Chicago, IL: Chicago Linguistic Society.

Munro, Pamela. 1976a. *Mojave syntax*. New York: Garland Publishing.

Munro, Pamela. 1976b. Subject copying, predicate raising, and auxiliarization: The Mojave evidence. *International Journal of American Linguistics* 42:99-113.

Munro, Pamela. 1976c. Two Stories by Nellie Brown. In *Yuman texts, International Journal of American Linguistics Native American Texts Series 1*, ed. Margaret Langdon, 43-50. Chicago, IL: University of Chicago Press.

Munro, Pamela. 1976d. Mojave modals. *Proceedings of the First Yuman Languages Workshop, Southern Illinois University Museum Studies 7*, ed. James E. Redden, 55-62. Carbondale: University Museum, Southern Illinois University.

Munro, Pamela. 1977. From existential to copula: The history of Yuman BE. *Mechanisms of syntactic change*, ed. Charles N. Li, 445-490. Austin: University of Texas Press.

Munro, Pamela. 1979. Reduplication in Mohave – and Yuman. *Proceedings of the 1978 Hokan-Yuman Languages Workshop*, ed. James E. Redden, 42-55. Carbondale: Southern Illinois University Department of Linguistics.

Munro, Pamela. 1980. Types of agreement in Mojave. *Proceedings of the 1979 Hokan-Yuman Languages Workshop, Occasional Papers in Linguistics 7*, ed. James E. Redden, 1-14. Carbondale: Southern Illinois University Department of Linguistics.

Munro, Pamela. 1981. Mojave k and m: It ain't necessarily so. *Proceedings of the 1980 Hokan Languages Workshop, Occasional Papers on Linguistics 9*, ed. James E. Redden, 124-129. Carbondale: Southern Illinois University Department of Linguistics.

Munro, Pamela, and Judith G. Crawford, ed. 1984. Mojave literature. In *Spirit Mountain: An anthology of Yuman story and song*, ed. Leanne Hinton and Lucille Watahomigie, 279-290. Tucson: Sun Tracks and University of Arizona Press.

Munro, Pamela, Nellie Brown, and Judith G. Crawford. 1992. *A Mojave dictionary, UCLA Occasional Papers in Linguistics 10.* Los Angeles: University of California Department of Linguistics.

Penfield, Susan. 1969. *Mohave language and oral history, The Doris Duke project.* Ms., on file at the Arizona State Museum, Tucson, Arizona, and at the Colorado River Indian Tribes Library/Archive.

Penfield, Susan, and Amelia Flores. 1999. A translation paradigm. To appear in the *Proceedings of the 1999 Stabilizing Indigenous Languages Conference.* Flagstaff: Northern Arizona University, Center for Excellence in Education.

Ruiz, Richard. 1995. Language planning considerations in indigenous communities. *Bilingual Research Journal* 19:71-81.

Swick, Edward. Various issues of *Smoke Signals* (Mohave Tribal newspaper).

(Weinberg)
Department of Linguistics
University of Arizona
200 E. Douglass Bldg.
Tucson, AZ 85721

weinberg@u.arizona.edu

(Penfield)
Department of English
University of Arizona
445 Modern Languages Bldg.
Tucson, AZ 85721

sdp@u.arizona.edu

Orders for publications from University of Arizona Linguistics Circle should be sent to the following address:

Linguistics Circle
Department of Linguistics
Douglass 200E
University of Arizona
Tucson, AZ 85721
U.S.A.

Shipping and handling, per individual volume:
U.S. - $2.00; Canada - $5.00; Foreign: Air - $17.00, Surface - $3.00

Arizona residents please add 5% tax. Checks should be made payable to the University of Arizona Linguistics Circle.

COYOTE PAPERS: WORKING PAPERS IN LINGUISTICS

Previously published in this series:

Volume 1:	COMPLEX VERBS AND THE LEXICON by Shigeru Miyagawa	Available by special order
Volume 2:	PROCEEDINGS OF THE ARIZONA CONFERENCE ON JAPANESE LINGUISTICS: THE FORMAL GRAMMAR SESSIONS edited by Ann Farmer and Chisato Kitagawa	Available by special order
Volume 3:	STUDIES ON ARABIC, BASQUE, ENGLISH, JAPANESE, NAVAJO AND PAPAGO edited by Thomas G. Larson	$3.00
Volume 4:	EXPLORING LANGUAGE: LINGUISTIC HERESIES FROM THE DESERT edited by Esmerelda Martin-Callejo-Manandise	$3.00

Volume 5:	STUDIES ON NATIVE AMERICAN LANGUAGES, JAPANESE AND SPANISH edited by Stuart Davis	$3.00
Volume 6:	COYOTE PAPERS: WORKING PAPERS IN LINGUISTICS FROM A → Z edited by Paul Saka	$3.00
Volume 7:	UNIFICATION BASED APPROACHES TO NATURAL LANGUAGES edited by Kenneth F. Drozd and Kaz Fukushima	$7.00
Volume 8:	COYOTE PAPERS: WORKING PAPERS IN LINGUISTICS FROM A → Z edited by Chip Gerfen and Pilar Pinar	$10.00
Volume 9:	COYOTE PAPERS: WORKING PAPERS IN LINGUISTICS FROM A → Z edited by Colleen M. Fitzgerald and Andrea Heiberg	$10.00
Volume 10:	COYOTE PAPERS: WORKING PAPERS IN LINGUISTICS edited by Sean Hendricks, Sachiko Ohno, and Amy Fountain	$10.00

ARIZONA PHONOLOGY CONFERENCE PROCEEDINGS

Previously published in this series:

Volume 1:	ARIZONA PHONOLOGY CONFERENCE 1988 edited by Megan Crowhurst	$7.00
Volume 2:	ARIZONA PHONOLOGY CONFERENCE 1989 edited by S. Lee Fulmer, Masahide Ishihara, and Wendy Wiswall	$8.00
Volume 3:	ARIZONA PHONOLOGY CONFERENCE 1990 edited by James Myers and Patricia E. Perez	$10.00
Volume 4:	ARIZONA PHONOLOGY CONFERENCE 1991 edited by Jean Ann and Kyoko Yoshimura	$10.00
Volume 5:	PROCEEDINGS OF THE 1995 SOUTHWESTERN WORKSHOP ON OPTIMALITY THEORY (SWOT) edited by Keiichiro Suzuki and Dirk Elzinga	$10.00

Volume 6: PROCEEDINGS OF THE 1998 SOUTHWESTERN WORKSHOP ON OPTIMALITY THEORY (SWOT) $10.00
edited by Jessica Maye and Mizuki Miyashita

OTHER SPECIAL VOLUMES

PROCEEDINGS OF THE POSTER SESSION OF THE 18TH ANNUAL WEST COAST CONFERENCE ON FORMAL LINGUISTICS (WCCFL 18) $10.00
edited by Peter Norquest, Jason D. Haugen, and Sonya Bird

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Coyote Papers Volume 11, Special Volume on Native American Languages	
Author(s): Jessica P. Weinberg, Erin L. O'Bryan, Laura A. Moll, Jason D. Haugen (Eds.)	
Corporate Source: University of Arizona Graduate Student Linguistics Circle (not for profit)	Publication Date: 2000

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature:	Printed Name/Position/Title: Erin L. O'Bryan / Publisher	
Organization/Address: Coyote Papers, Linguistics Circle 200E Douglass Bldg, Univ of Arizona Tucson AZ 85721 USA	Telephone: (520) 6266593	FAX: (520) 6269014
	E-Mail Address: Coyote@u.arizona.edu	Date: 6/13/2001

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on
Languages & Linguistics
4646 40TH ST. NW
WASHINGTON, D.C. 20016-1859

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com