Advances in the study of Siouan languages and linguistics

Edited by

Catherine Rudin

Bryan J. Gordon

Studies in Diversity Linguistics 10

language science press

Studies in Diversity Linguistics

In this series:

1. Handschuh, Corinna. A typology of marked-S languages.
2. Rießler, Michael. Adjective attribution.
3. Klamer, Marian (ed.). The Alor-Pantar languages: History and typology.
4. Berghäll, Liisa. A grammar of Mauwake (Papua New Guinea).
5. Wilbur, Joshua. A grammar of Pite Saami.
6. Dahl, Östen. Grammaticalization in the North: Noun phrase morphosyntax in Scandinavian vernaculars.
7. Schackow, Diana. A grammar of Yakkha.
8. Liljegren, Henrik. A grammar of Palula.
9. Shimelman, Aviva. A grammar of Yauyos Quechua.
10. Rudin, Catherine & Bryan James Gordon (eds.). Advances in the study of Siouan languages and linguistics.

ISSN: 2363-5568

Advances in the study of Siouan languages and linguistics

Edited by

Catherine Rudin

Bryan J. Gordon

language science press

Catherine Rudin & Bryan J. Gordon (eds.). 2016. *Advances in the study of Siouan languages and linguistics* (Studies in Diversity Linguistics 10). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/94

ISBN: 978-3-946234-37-1 (Digital)
978-3-946234-38-8 (Hardcover)
ISSN: 2363-5568
DOI:10.5281/zenodo.3741670

Cover and concept of design: Ulrike Harbort
Typesetting: Bryan James Gordon, Sebastian Nordhoff, Catherine Rudin
Illustration: Sebastian Nordhoff
Proofreading: Aaron Sonnenschein, Alec Shaw, Alessia Battisti, Andreas Hölzl, Aneilia Stefanova, Carolyn O'Meara, Christian Döhler, Dominik Lukeš, Eitan Grossman, Elizabeth Zeitoun, Matthew Czuba, Martin Haspelmath, Neal Whitman, Roelant Ossewaarde, Rong Chen, Stathis Selimis, Steve Pepper, Teresa Proto, Varun de Castro Arrazola
Fonts: Linux Libertine, Arimo, DejaVu Sans Mono
Typesetting software: XeLaTeX

Language Science Press
Habelschwerdter Allee 45
14195 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

To Bob, whose knowledge was matched only by his generosity.

Contents

Preface

This volume presents a group of papers representing a range of current work on Siouan[1] languages, in memory of our colleague Robert L. Rankin, a towering figure in Siouan linguistics throughout his long career, who passed away in February of 2014.

Beyond honoring a beloved colleague, our aim in this volume is to bring a variety of issues in Siouan linguistics to the attention of the linguistic community. The Siouan language family is a large and important one, with branches geographically distributed over a broad swath of the North American plains and parts of the southeastern United States. This puts it in contact historically with several other families of languages: Algonquian, Iroquoian, Caddoan, Uto-Aztecan, and Muskogean. Siouan languages are, or were historically, spoken by the members of at least 25 ethnic/political groups. One Siouan language, Lakota, is among the handful of indigenous North American languages with younger speakers today. Siouan languages have occasionally risen to prominence in general linguistics, for instance in the study of reduplication (Shaw 1980); and Omaha and Crow (Apsaalooke) have lent their names to two of the basic categories of kinship systems in anthropology. Nonetheless, the Siouan family has been underrepresented in the descriptive and typological literature, and most of the languages in the family are severely understudied. The majority of work on Siouan languages is unpublished, existing only in the form of conference papers or manuscripts.[2] This volume is a step toward making information on Siouan languages more broadly available and encouraging deeper investigation of the myriad issues they raise.

From the perspective of linguistic typology, Siouan languages have many notable features. Many of these features stand to challenge typological generalizations. Here we briefly sketch a few of the most characteristic features of the Siouan family.

[1] "Siouan" is not to be confused with "Sioux", a controversial term referring to Lakota and Dakota people, rarely to Nakota/Nakoda people too, but never correctly to people of other traditionally Siouan-language-speaking communities.

[2] Many of these unpublished works are collected in the electronic Siouan Archive, maintained by John Boyle at the University of California at Riverside.

All Siouan languages possess a rich variety of applicative affixes, confirming Polinsky's (2013) observation that applicatives are common in North America and adding another language family to her list of applicative-rich families in the area. Helmbrecht (2006) divides the applicatives into three templatic slots: locative applicatives, benefactive applicatives, and applicative markers; all of the Siouan languages sampled by Helmbrecht possess at least two applicative morphemes.

All Siouan languages are strongly head-final, and the consensus among syntacticians working with Siouan languages is that all but the supraclausal projections (and even some of these) are underlyingly head-final in Siouan languages, contra Kayne's (1994) Antisymmetry theory.

All Siouan languages have head-internal relative clauses. A series of strong claims regarding the typological implications of head-internal relative clauses (cf. Cole 1987; Murasugi 2000), including purported distinctions between "Japanese-type" and "Lakota-type" constructions (cf. Watanabe 2004; Williamson 1987; Bonneau 1992), propelled Lakota into the debates of theoretical syntax. It has been pointed out that head-internal relative clauses of the kind found in Lakota and other Siouan languages lack the island restrictions found in other languages. On the other hand, Murasugi (2000) argues that languages with head-internal relative clauses must also have head-external relative clauses, which is not true in Siouan languages.

All Siouan languages have verbal affixes which index subject possession of or relationship with the object. They vary with respect to contexts of obligatoriness of these affixes.

Many Siouan languages have grammaticalized systems of speaker-gender marking, with gender-specific morphology for speech-act markers, address terms, and kinship lexemes.[3] Such usage varies depending on stuational factors, however, especially in the case of speech-act markers; see for instance Trechter (1995).

Many Siouan languages have a modal CCV morpheme shape. This does not necessarily imply a preference for CCV phonetic realizations, but may indicate such a preference in the distant past. Another unusual prosodic feature is the preference for second-syllable stress in most Siouan languages. Ho-Chunk may be the only attested language with default third-syllable stress in the world.

Most Siouan languages have ejective stops. The Dhegiha branch is notable for a four-way glottal-state distinction in its stop series (voiced/lenis, tense/pre-aspirated, ejective and aspirated). Outside of the Dhegiha branch are many Siouan

[3] In the case of kinship terms, lexical choice is driven by the gender of the "ego" deictic center, which coincides with speaker gender when there is 1st-person inflection.

languages which have the unusual feature of a phonemic voicing distinction in fricatives but not in stops.

Verbs play some typologically unusual, prominent roles in Siouan languages. Diachronically, many grammatical items which rarely grammaticalize from verbs in other languages tend to derive from verbs in Siouan languages. For instance, Rankin (1977) documents the derivation of classifiers and articles from verbs. In some Siouan languages, the source verbs and target grammatical items continue to exist in parallel with substantial semantic overlap. The Omaha positional article *t^{h}o^{n}* 'OBVIATIVE ANIMATE SPECIFIC STANDING', for instance, is homophonous with the root of *át^{h}o^{n}* 'stand on'.

This diachronic tendency is mirrored by synchronic flexibility. Siouan languages tend to verb freely — to use nearly any open-class stem as a verb. Thus Lakota *wimačhaša* 'I am a man' is derived from the nominal stem *wičhaša* 'man/person' with the 1st-person stative pronominal ma-.

Dhegiha articles (which have many features in common with positional classifiers in e.g. Mayan languages; see Gordon 2009) are homophonous with postverbal and postclausal functional items like subordinating conjunctions and aspect and evidentiality markers. They have considerable semantic overlap with them too, a fact which comprises another area of blurriness between nominal and verbal syntax: In Ponca, *niáshi^{n}ga-ama* may mean 'the [PROXIMATE ANIMATE PLURAL SPECIFIC] people', but also may mean either 'they are people' or 'I am told s/he was a person'. Plurality is a part of the semantics of -ama in both the nominal and the first clausal interpretation. To make matters more interesting, these kinds of ambiguity are not always easily resolved by context alone, and may suggest a "simultaneity" (cf. Woolard 1998) at work as part of speakers' competence.

This flexibility, that is, the ability of one and the same root to function in both nominal and verbal contexts, has led to some discussion on the status and quality of the noun/verb distinction in Siouan languages (see e.g. Helmbrecht 2002 and Ingham (2001)).

Nominal arguments in general are not required in Siouan languages, thematic relations being signaled by pronominal or agreement markers within the verb — including zero markers. This makes Siouan languages relevant to debates about the existence of "pronominal argument" languagesPronominal Argument Hypothesis (Jelinek 1984) and to the related issues of whether there are languages with truly nonconfigurational or flat structure. The preponderance of evidence in Siouan is for the existence of hierarchical structure, specifically including a VP (for instance, West 2003; Johnson 2016; Johnson, Rosen & Schuck 2016; Rosen 2016).

Although Siouan languages have many remarkable features in common, they vary on many others. Some Siouan languages have noun incorporation, while others do not. Some Siouan languages have stress-accent systems, and others have pitch-accent systems. Dhegiha languages are notable in having as many as eleven definite/specific articles indexing features such as animacy, proximacy/obviation (or case), posture/position, number, visibility, motion and dispersion; meanwhile other Siouan languages have no fully grammaticalized articles at all.

Some Siouan languages reflect longtime cultural presence on the Plains, while others are located as far east as the Atlantic Coast, and many more show cultural aspects of both regions. Dhegiha-speaking peoples (Quapaw, Osage, Kanza, Omaha and Ponca, and likely Michigamea as well (Kasak 2016; Koontz 1995) likely lived at the metropolis at Cahokia, perhaps at a time before any of the descendant groups had separated, and have many Eastern Woodlands-style features of traditional governance and religion, in sharp contrast with the more Plains-typical cultural features of close Lakota and Dakota neighbors and relatives.

One seemingly minor but in fact quite significant issue in Siouan linguistics is the matter of language names and their spelling. Often this involves a self-designation in competition with a name imposed by outsiders. Even when an autonym gains currency among linguists there is sometimes no agreed spelling; so for instance the Otoe self-designation is written <Jiwere> or <Chiwere>. For the most part in this volume the choice of language designations has been left to the individual chapter authors. However, after a volume reviewer pointed out that the language of the Ho-Chunk or "Winnebago" people was spelled no less than ten different ways in various chapters, we encouraged authors to choose one of the two spellings used on the tribe's web site: <Ho-Chunk> or <Hoocąk>. Most have voluntarily complied. In a related move, we decided to transliterate all Lakota data throughout the volume using the orthography of the *New Lakota Dictionary* (Ullrich 2012).

The volume is divided into four broad areas (Historical, Applied, Formal/Analytical, and Comparative/Cross-Siouan) described in more detail in separate introductions to each part of the volume. Part I consists of five chapters on historical themes: Ryan Kasak evaluates the evidence for a relationship between Yuchi and Siouan; David Kaufman discusses the participation of some Siouan languages in a Southeastern Sprachbund; Rory Larson summarizes current knowledge of Siouan sound changes; and Kathleen Danker and Anthony Grant investigate early attempts to write Ho-Chunk, Kanza, and Osage. Part II opens with Linda Cumberland's interview with Robert Rankin about his work with Kanza language programs. Jimm Goodtracks, Saul Schwartz, and Bryan Gordon present three

different perspectives on Baxoje-Jiwere language retention. Justin McBride applies formal syntax to the solution of a pedagogical problem in teaching Kanza. This applied-linguistics section ends with Jill Greer's sketch grammar of Baxoje-Jiwere. Part III contains formal analyses of individual Siouan languages. David Rood proposes an analysis of /b/ and /g/ in Lakota using the tools of autosegmental phonology and feature geometry. John Boyle elucidates the structure of relative clauses in Hidatsa. Meredith Johnson, Bryan Rosen, and Mateja Schuck, in a series of three interrelated chapters, discuss syntactic constructions in Ho-Chunk including resultatives and VP ellipsis, which they argue show the language has VP and an adjective category. Part IV consists of three chapters which take a broader view of grammar, considering data from across the Siouan family. Catherine Rudin compares coordination constructions across Siouan; Bryan Gordon does the same with information structure and intonation, and Johannes Helmbrecht with nominal possession constructions.

All four of the areas represented by this volume are ones to which Bob Rankin contributed. His scholarly publications centered primarily around Siouan historical phonology, but included works ranging from dictionaries to toponym studies, from philological investigation of early Siouanists to description of grammaticalization pathways. He was deeply involved in language retention efforts with the Kanza Language Project. Other interests included archeology, linguistic typology, Iroquoian and Muskogean languages, and the history of linguistics.

Bob was a major figure in Siouan linguistics, a mentor to nearly all living Siouanists, and a mainstay of the annual Siouan and Caddoan Linguistics Conference meetings for decades. Trained in Romance and Indo-European linguistics, with a specialty in Romanian (Ph.D. University of Chicago 1972), he shifted gears soon after leaving graduate school, and became an expert in Siouan languages, especially the Dhegiha branch, with special focus on Kanza. From the mid 1970s through the end of his life, he devoted himself to Siouan studies, both practical and scholarly. His long association with the Kaw Nation led to a grammar and dictionary of that language (see Cumberland 2016), and he also produced a grammar of Quapaw, and briefly conducted field work on Omaha-Ponca and Osage. At the University of Kansas he directed dissertations on Lakota (Trechter 1995) and Tutelo (Oliverio 1996) as well as several M.A. theses on Siouan languages, and taught a wide variety of courses including field methods and structure of Lakota and Kanza as well as more theoretical courses in phonology, syntax, and historical linguistics. Perhaps Bob's greatest gift to the field was his encouragement of others. At conferences and on the Siouan List email forum, he was unfailingly patient and encouraging, answering all questions seriously, explaining linguistic

terms to non-linguist participants and basic facts of Siouan languages to general linguists with equal enthusiasm and lack of condescension.

Following his untimely passing, a special session was held at the 2014 Siouan and Caddoan Linguistics Conference to organize several projects in Bob's honor: The first of these was publication of the Comparative Siouan Dictionary, an immense project comparing cognates across all the Siouan languages, undertaken by Rankin and a group of colleagues in the 1980s. It had been circulated in various manuscript forms but never published. Thanks to David Rood (another founding member of the CSD project), with help from Iren Hartmann, the CSD is now available online (Rankin et al. 2015). The second project was a volume of Bob's conference papers and other previously unpublished or less accessible work, to be collected and edited by a group headed by John P. Boyle and David Rood; that volume, tentatively titled *Siouan Studies: Selected Papers by Robert L. Rankin*, is currently in progress. The third project was a volume of Siouan linguistic work in Bob's memory, which has taken the shape of the present volume.

References

Bonneau, José. 1992. *The structure of internally headed relative clauses*. Montréal: McGill University (Doctoral dissertation).

Cole, Peter. 1987. The structure of internally headed relative clauses. *Natural Language and Linguistic Theory* 5(2). 277–302.

Cumberland, Linda A. 2016. In his own words: Robert Rankin recalls his work with the Kaw people and their language. In Catherine Rudin & Bryan James Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 119–132. Berlin: Language Science Press.

Gordon, Bryan James. 2009. *"Artifiers" in Mississippi Valley Siouan: A novel determiner class*. Paper presented at the annual meeting of the Linguistic Society of America, San Francisco, 8–11 January 2009.

Helmbrecht, Johannes. 2002. Nouns and verbs in Hocank (winnebago). *International Journal of American Linguistics* 68(1). 1–27.

Helmbrecht, Johannes. 2006. *Applicatives in Siouan languages: A study in comparative Siouan grammar*. Paper presented at the Siouan and Caddoan Languages Conference, Billings, MT.

Ingham, Bruce. 2001. Nominal and verbal status in Lakhota: A lexicographical study. *International Journal of American Linguistics* 67(2). 167–192.

Jelinek, Eloise. 1984. Empty categories, case and configurationality. *Natural Language and Linguistic Theory* 2. 39–76.

Johnson, Meredith. 2016. A description of verb-phrase ellipsis in Hocąk. In Catherine Rudin & Bryan James Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 287–313. Berlin: Language Science Press.

Johnson, Meredith, Bryan Rosen & Mateja Schuck. 2016. Evidence for a VP constituent in Hocąk. In Catherine Rudin & Bryan James Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 339–363. Berlin: Language Science Press.

Kasak, Ryan M. 2016. A distant genetic relationship between Siouan-Catawban and Yuchi. In Catherine Rudin & Bryan James Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 5–37. Berlin: Language Science Press.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Koontz, John E. 1995. *Michigamea as a Siouan language*. Paper presented at the 15th annual Siouan and Caddoan Languages Conference, University of New Mexico – Albuquerque.

Murasugi, Keiko. 2000. An antisymmetry analysis of Japanese relative clauses. In Artemis Alexiadou, Paul Law, André Meinunger & Chris Wilder (eds.), *The syntax of relative clauses*, 231–264. Amsterdam: John Benjamins.

Oliverio, Giulia. 1996. *A grammar and dictionary of Tutelo*. Lawrence: University of Kansas (Doctoral dissertation).

Polinsky, Maria. 2013. Applicative constructions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/109 (23 August, 2015).

Rankin, Robert L. 1977. From verb, to auxiliary, to noun classifier and definite article: Grammaticalization of the Siouan verbs 'sit', 'stand', 'lie'. In Robert L. Brown, Kathleen Houlihan, Larry G. Hutchinson & Andrew MacLeish (eds.), *Proceedings of the 1976 Mid-America Linguistics Conference*, 273–83. Minneapolis: University of Minnesota Linguistics Department.

Rankin, Robert L., Richard T. Carter, A. Wesley Jones, John E. Koontz, David S. Rood & Iren Hartmann (eds.). 2015. *Comparative Siouan dictionary*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://csd.clld.org (25 September, 2015).

Rosen, Bryan. 2016. On the structure and constituency of Hocąk resultatives. In Catherine Rudin & Bryan James Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 313–338. Berlin: Language Science Press.

Shaw, Patricia A. 1980. *Theoretical issues in Dakota phonology and morphology*. New York: Garland Publishing.

Trechter, Sara. 1995. *The pragmatic functions of gendered clitics in Lakhota*. Lawrence: University of Kansas (Doctoral dissertation).
Ullrich, Jan F. (ed.). 2012. *New Lakota dictionary*. 2nd edn. Bloomington, IN: Lakota Language Consortium. http://www.lakotadictionary.org/nldo.php.
Watanabe, Akira. 2004. Parametrization of quantificational determiners and head-internal relatives. *Language and Linguistics* 5. 59–97.
West, Shannon Louise. 2003. *Subjects and objects in Assiniboine Nakoda*. Victoria: University of Victoria (Master's thesis).
Williamson, Janis Shirley. 1987. An indefiniteness restriction for relative clauses in Lakhota. In Eric Reuland & Alice ter Meulen (eds.), *The representation of (in)definiteness*, 168–190. Cambridge, MA: MIT Press.
Woolard, Kathryn. 1998. Simultaneity and bivalency as strategies in bilingualism. *Journal of Linguistic Anthropology* 8(1). 3–29.

Acknowledgments

We are grateful to everyone who helped with this project, especially the many people who served as chapter reviewers and volume reviewers. Special thanks to Deniz Rudin for his invaluable typesetting/LaTeX assistance, to Jan Ullrich for orthography checking, and to Sebastian Nordhoff at Language Science Press for his patient help with all the details of producing the final manuscript.

Part I

Historical linguistics and philology

Introduction to Part I

The relative degree of "genetic" relatedness of the major branches of the Siouan language family is quite well-established: the Catawban languages split off first, then the Missouri Valley Siouan languages, followed by the Southeastern Siouan and Mississippi Valley Siouan languages. Among the latter branch, the Dakotan languages split off first, followed by the Dhegiha and Jiwere-Ho-Chunk sub-branches. Rankin (1988; 1998; Rankin, Carter & Jones 1998; etc.) contributed much to developing and supporting this understanding, alongside advances in rigorous application of the comparative method. Open questions include the possibility of relationships with Yuchi, Iroquoian languages and Caddoan languages, and areal connections. Chapters in Part 1 of this volume address some of these issues, as well as considering what we can learn from early attempts to write Siouan languages.

Ryan Kasak ("A distant genetic relationship between Siouan-Catawban and Yuchi") argues that evidence exists to link Yuchi to the Siouan family. Though scarce and not fully conclusive, the evidence includes phonological and morphological correspondences strong enough to make a reasonable case for a genetic relationship between Yuchi and Siouan-Catawban.

David Kaufman ("Two Siouan languages walk into a Sprachbund") details the effects on Ofo and Biloxi of their participation in the Lower Mississippi Valley (LMV) language area; these languages share many lexical, phonetic and grammatical traits with genetically unrelated languages across the southeastern present-day United States.

Rory Larson ("Regular sound shifts in the history of Siouan") summarizes the current state of knowledge of the sound changes and correspondences distinguishing each branch and sub-branch of the Siouan family. These phonetic correspondences were worked out as part of the Comparative Siouan Dictionary project (recently made available online as Rankin et al. (2015)), of which Bob Rankin was a central member. This concise catalog of all the known sound changes will be invaluable to anyone working with Siouan etymologies or cognates in the future.

Kathleen Danker ("Ba-be-bi-bo-ra: Refinement of the Ho-Chunk syllabary in the 19th and 20th centuries") presents a glimpse into the process of formation of a Native writing system. This syllabary was inspired by one used by neighboring Algonquian peoples, then progressively changed to better represent the phonologically quite different Ho-Chunk language before being supplanted by writing in English, and by an alphabetic Ho-Chunk orthography.

Anthony Grant ("A forgotten figure in Siouan and Caddoan linguistics: Samuel Stehman Haldeman (1812–1880)") is another study of writing systems, in this case an early attempt to write Kanza and Osage. Haldeman developed a universal phonetic orthography, one of several precursors to the modern IPA, which he tried out on a variety of languages including these two Siouan ones. While not entirely successful at representing all the sounds of Kanza and Osage, Haldeman's word lists do provide some insights into the pronunciation of these languages at a time earlier than other available information on them.

References

Rankin, Robert L. 1988. Quapaw: Genetic and areal affiliations. In William Shipley (ed.), *In honor of Mary Haas: From the Haas Festival conference on Native American linguistics*, 629–650. Berlin: Mouton de Gruyter.

Rankin, Robert L. 1998. Grammatical evidence for genetic relationship and the Macro-Siouan hypothesis. In Marie-Lucie Tarpant (ed.), *Actes du 21ième colloque annuel de l'Association de Linguistique des Provinces Atlantiques*, 20–44. Halifax: Mount Saint Vincent University.

Rankin, Robert L., Richard T. Carter & A. Wesley Jones. 1998. Proto-Siouan phonology and grammar. In Xingzhong Li, Luis López & Tom Stroik (eds.), *Papers from the 1997 Mid-America Linguistics Conference*, 366–375. Columbia: University of Missouri Press.

Rankin, Robert L., Richard T. Carter, A. Wesley Jones, John E. Koontz, David S. Rood & Iren Hartmann (eds.). 2015. *Comparative Siouan dictionary*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://csd.clld.org (25 September, 2015).

Chapter 1

A distant genetic relationship between Siouan-Catawban and Yuchi

Ryan M. Kasak

A lack of ancient written records is no impediment to establishing genetic relationships between languages at great time depths. While scholars like Sapir (1929) have proposed genetic groupings based on particular lexical similarities, other scholars have utilized a multifaceted approach to arguing for relatedness by comparing both lexical items and morphological material, given the fact that the latter is less prone to change over time than the former (Goddard 1975; Vajda 2010). This paper assesses Rankin's (1996; 1998) earlier analysis of the plausibility of a relation from common descent between Siouan-Catawban and Yuchi, which is currently considered by most an isolate. By comparing cognates and establishing possible sound correspondences, and by examining the peculiarity of the verbal template with respect to the placement of the first person plural marker vis-à-vis the preverb and verb, and the use of nasal ablaut in Yuchi to mark future tense that is similar to *iŋ*-ablaut in Dakotan languages, this paper builds upon Rankin's original case for a genetic link between Yuchi and Siouan-Catawban. While more constrained in scope than Chafe's (1976) Macro-Siouan proposal, this paper adds to the body of support for Siouan-Catawban and Yuchi sharing a common ancestor.

1 Introduction

The absence of pre-contact written records and the increasing loss of native speakers are major problems in researching linguistic change in North American indigenous languages. Unlike Hittite and Tocharian, whose written records provided major breakthroughs in our current understanding of the spread and evolution of Indo-European languages, no such breakthrough is likely to be found in the archaeological record. Rather, deeper genetic relations among North American languages must be found by sifting through old hymnals, Jesuit memoirs, and page after page of field notes left behind by past researchers.

The purpose of this paper is to re-visit the idea of a plausible common ancestry for the Siouan-Catawban language family and the language isolate Yuchi (a.k.a.

Ryan M. Kasak. 2016. A distant genetic relationship between Siouan-Catawban and Yuchi. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 5–39. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.120
DOI:10.17169/langsci.b94.120

"Euchee"). This line of inquiry grows out of earlier work done by the late Bob Rankin (1996; 1998). Though Yuchi has been grouped with Siouan-Catawban in the past, there is no consensus on its status as a distant relative or simply a language that may have the occasional similarity here and there.

The overarching goal is twofold. Firstly, I wish to summarize the state of linguistic scholarship up to this point for both Siouan-Catawban and Yuchi. In addition to looking at purely linguistic data, I piece together what is known about these language groups to demonstrate that there are significant non-linguistic factors to support the idea that Siouan-Catawban and Yuchi can be related, notably drawing from historical accounts of the proximity of the Yuchi to Siouan and Catawban peoples from the sixteenth century onwards. Secondly, I wish to support the notion that Yuchi is a distant relative of the Siouan-Catawban languages by providing lexical and morphological evidence.

The task of establishing genetic relationships among languages of the New World is not altogether dissimilar from using methods traditionally reserved for languages with robust written histories. Bloomfield (1925) manages to apply the same methods used by Indo-Europeanists to create proto-forms for Central Algonquian languages. By using the comparative method, he is able to convincingly reconstruct numerous proto-forms for seven different groups of Algonquian languages, further cementing the idea that measuring linguistic change in North America is not a lost cause. His work, however, focuses on languages whose relation is readily apparent and well-accepted. A far trickier task is to connect two languages or language families whose relationship is not well-accepted, particularly when dealing with isolate languages.

Goddard (1975) demonstrates a similarity in morphology between Proto-Algonquian and the Californian languages Yurok and Wiyot. In addition to morphology, Goddard finds a small group of lexical items and posits a set of plausible reconstructions to explain how such linguistically and geographically divergent languages can ultimately be shown to be quite similar. Today, Yurok and Wiyot are grouped with Algonquian as part of the Algic family, and their status as related languages is readily accepted. Deeper genetic relationships are more easily accountable through the comparison of morphological features than lexical items due to the fact that lexical change is possible at a faster rate than morphological change. The fact that most Balto-Slavic languages have a robust case system despite thousands of years of separation from Proto-Indo-European, or the widespread use of the inherited Proto-Polynesian passive marker *ia in its daughter languages is testament to the more languid pace at which morphology changes. A more recent example of the power of comparative morphology is that

of the Dené languages of North America and the Yeniseian languages of central Siberia. Despite a potential time depth of 13,000 to 15,000 years, it is still possible to show cognacy between a variety of inflectional and derivational morphology, as well as individual lexical items (Vajda 2010).

Dunn (2009) has even taken this concept to its next logical step by utilizing Bayesian methods of phylogenetic inference to provide evidence of similarity in the morphology of Austronesian and Papuan languages of Island Melanesia. While Dunn's methods are not employed in this work, the idea behind them is the same: can a convincing argument be made for a relationship between two different language groups using morphological evidence despite a lack of lexical similarity? Ultimately, I argue in favor of the morphological similarities found in both Siouan-Catawban and Yuchi to be more than coincidental. Such similarities are most likely due to genetic inheritance.

This paper is organized into six sections. In §2, I summarize the historical scholarship leading up to the classification of the Catawban family and the Siouan family, along with their eventual classification as related language groups. In addition to pointing out the development of the notion of a Siouan-Catawban language family, I explore the work previously conducted on Yuchi. §3 highlights past research into the relationship Siouan-Catawban may have with Yuchi, and the critiques of those groupings made by other linguists.

In §4, I give a list of lexical items that appear to be cognates and posit a probable set of sound correspondences between Siouan-Catawban and Yuchi. §5 is dedicated to building on the lack of cognates with morphological cognacy, while §6 offers a conclusive summary of the data found within this paper in addition to adding further commentary on some other possible long-distance genetic relationships between Siouan-Catawban and other indigenous North American languages.

2 Previous scholarship on Siouan-Catawban and Yuchi

Siouan-Catawban, often called simply Siouan, is a language family whose speakers are predominantly found on the Great Plains, though some languages were once spoken around the Ohio River Valley and along the Atlantic and Gulf coasts (Mithun 1999). Catawban split from Siouan proper some time in the distant past, which Rankin (1996) posits is at least 4,000 years before the present.[1]

[1] For a more detailed phylogenetic look at the relationship between the Siouan languages, see Rankin (2010).

2.1 Catawban language family

Catawban only has two attested members, Woccon and the eponymous Catawba. Woccon went extinct in the early 1700s, and is only known by a list of 143 words printed in 1709. Carter (1980) identifies the first known attempts to link Woccon and Catawba by Western philologists in the early nineteenth century to Adelung & Vater (1816), who created a side-by-side wordlist to compare the two languages. Gallatin (1836) builds upon the notion that Woccon and Catawba share a common descent and expands the comparison between the two languages.

2.2 Siouan language family

In addition to examining the relationship between Woccon and Catawba, Gallatin (1836) also is credited for being the first scholar to posit a Siouan language family, named after their most well-known members, the Sioux.

2.2.1 Early classification of languages within Siouan

His first classification breaks the Siouan language family into four clades: 1) Ho-Chunk (also called "Winnebago"); 2) Sioux, Assiniboine, and Cheyenne; 3) Hidatsa (sometimes called "Minetare" in older literature), Mandan, and Crow; and 4) Osage, Kaw, Quapaw, Iowa, Otoe-Missouria, and Omaha-Ponca (Parks & Rankin 2001). His classification of Cheyenne as a Siouan language is now outmoded, as Cheyenne is now uncontroversially classified as an Algonquian language.

Gallatin's (1836) identification and classification of Siouan languages serves as the jumping point for all future research into the internal relationships between the Siouan languages. Mandan started out as something of an issue for early scholars, as it bears a strong lexical affinity to the Missouri Valley languages Hidatsa and Crow. However, it clearly was more distinct from these two languages, both lexically and grammatically. Thus, several groupings done by scholars after Gallatin struggled to place Mandan within the language family, alternatively lumping it with Crow and Hidatsa or giving it a phylum of its own. Regrettably, the status of Mandan is still somewhat suspect even today, as Rankin (2010) advocates recognizing Mandan as its own branch of the Siouan family tree due to its distinct morphology.

2.2.2 Discovery of an additional branch of the Siouan family tree

Hale (1883) breaks new ground by documenting the Tutelo language of Virginia, which bears a strong lexical and morphological affinity to the Siouan languages

of the Great Plains. Thus, for the first time, there is conclusive evidence of a more widespread relationship between the Siouan languages and those farther east of the Mississippi.

Hale further recounts the recent discovery of documents that place peoples bearing Siouan names around the Appalachian mountains from modern-day Virginia to North Carolina. Swiss-American linguist Albert Gatschet identifies the language Biloxi, spoken on the Gulf of Mexico, as a Siouan language (Dorsey 1893). Earlier, Gatschet had visited the Catawbas of South Carolina in 1881 and remarked on how similar this language is to the Siouan languages (Gatschet 1900).

2.2.3 A possible eastern origin

The realization that Tutelo and Biloxi had a clear affinity with the Dakotan languages of the Great Plains was a great discovery. However, the question arose of whether there was a migration of some ancestral group from the east to the west or vice versa. Swanton's (1909) subsequent analysis of Ofo (a.k.a. "Mosopelea"), a now-extinct language recorded in northern Mississippi, helped to paint a picture of westward migration from around the Ohio River Valley. Records from early explorers to the region strengthened his hypothesis. Before Swanton's field work, Ofo was assumed to be a Muskogean language. The intuition that Ofo belonged with the Muskogean languages is in part due to the fact that /f/ is not found in Siouan languages, but common among languages of the western regions of the Deep South (Rankin 2004; Kaufman 2014).

Hanna (1911) states that the Ofo lived in eight villages between the Muskingum and Scioto rivers, north of the Ohio River. When French explorer Jean-Baptiste-Louis Franquelin explored the vicinity of these eight villages in 1684, he noted that they had all been destroyed. The Ofo had been attacked and driven from their homeland by the Haudenosaunee (a.k.a. "Iroquois"),[2] whose aggression during the Beaver Wars had caused other groups from the eastern Great Lakes region to flee to safer lands in the West (Reed 1952). Connecting the Ofo with Siouan languages, combined with the anthropological and historical data on the Mosopelea, created a more complete picture of the time frame of when many Siouan languages shifted westward towards the Great Plains or southward towards the Gulf Coast.

[2] Thank you to the annonymous reviewer who suggested that I use the endonym "Haudenosaunee," rather than the exonym "Iroquois", whose etymology is typically considered to be derogatory (cf. Day 1968).

2.2.4 Support from missionary texts

Koontz (1995a,b) adds support to an eastern origin for an ancestral homeland for Siouan-Catawban speakers being somewhere in the East by classifying the understudied Michigamea language as Siouan. The Michigamea were a member of the Illinois Confederacy and were thought to have spoken a dialect of Miami-Illinois, whose range stretched from northern Arkansas to near Lake Michigan, though the northernmost extent of their habitation is somewhat in question.

A Jesuit who visited the Michigamea in 1673 found himself unable to communicate with them in any of the six other languages he spoke, including Illinois. Curiously, the Michigamea were also regarded as go-tos for dealing with the nearby Quapaw tribe, whose language was clearly Siouan. The recorded evidence of Michigamea is scant, but two complete sentences were enough to clearly show that Michigamea is not an Algonquian language, but a Siouan one. Its status within the Siouan language family is not completely understood, but Michigamea shows very strong affinities to the Dhegihan branch (Koontz 1995a,b).

2.2.5 Support from historical toponymy

In addition to searching for information in missionary texts, more modern scholarship by Booker, Hudson & Rankin (1992) examines toponyms and ethnonyms documented by Spanish explorers during three expeditions in the 1500s to corroborate the idea that there were once Siouan peoples along the Eastern seaboard. Hernando de Soto, Tristán de Luna y Arellano, and Juan Pardo all led multi-year exploratory missions into the American mainland from Spanish-held territory near modern-day Tampa, Florida.

The expeditions took place approximately a decade apart from each other, and together covered territory spanning modern-day Florida, Alabama, Tennessee, Georgia, and the Carolinas. Booker, Hudson & Rankin (1992) outline each of the place names and give the most likely modern analysis for what language group under which to classify them. The names, written in Spanish orthography, give strong clues that the Spaniards had visited a large number of Catawban villages, and possibly one non-Catawban village that of possible Siouan stock: the Chequini. If these people were Siouan, they were likely speakers of a Virginia Siouan language.

When English-speaking settlers began to settle the Atlantic coast of North America in large numbers, there was often little to no trace of the inhabitants described by the Spanish. No doubt, the spread of disease and conflicts among the indigenous groups played an enormous role in the large-scale demographic shifts of the Southeast (Mann 2006).

2.2.6 Siouan tribes in the midwest

The presence of Siouan tribes in the Great Lakes region during the 1600s makes sense when couched in a historical context. That is, around the time of more aggressive European colonization and expansion, the Haudenosaunee and Algonquian peoples of the Eastern seaboard became entangled in the Beaver Wars, in which the aforementioned groups vied for dominance of the fur market and trade with Europeans, pushing refugees west over the Mississippi or into the Southeast, displacing other autochthonous peoples.

Jennings (1990) also places the Lakota near modern-day Chicago near the southern tip of Lake Michigan in 1648, meaning that they had not yet crossed the Mississippi River until some time in the late seventeenth or early eighteenth century. The Manahoac tribe, whose autonym was identical to that of the Tutelo, were likewise driven from the Piedmont Plateau of Virginia by the Haudenosaunee, who claimed their territory as hunting grounds by right of conquest.

2.2.7 Summary of the eastern origin question

The overlap of linguistic, anthropological, and historical data together support the idea that the majority of all Siouan-Catawban peoples resided in or around the Ohio River Valley by the seventeenth century, only to join numerous other tribes in flight before the aggression of the Haudenosaunee.

2.3 Yuchi language

Currently, Yuchi is fluently spoken in Oklahoma by a small group of elders, with Linn (2000) stating that their number was around a dozen, though it is likely less now. There are some middle-aged heritage speakers who passively understand the language, but are mostly unable to engage in conversation in Yuchi. The language is considered an isolate, though that idea is called into question in subsequent sections of this paper.

2.3.1 Early records

The first records of the Yuchi place them in the Southeast near the Upper Tennessee Valley during the middle of the sixteenth century (Gatschet 1885). The Chisca are associated in literature with the Yuchi. Hernando de Soto encountered these people on his expedition, and he sent two men to find their villages, as they were reported to have gold.

Later, in 1567, the Spanish burned down two of their villages after the rumors of the Chisca having gold turned out to be false. There is no convincing argument as to why the Chisca are to be associated with the Yuchi other than the fact that the Cherokee, Yuchi, and Koasati of the area all seem to share several loanwords to support the idea of cultural contact (Booker, Hudson & Rankin (1992)).

2.3.2 Removal from the southeast

The Yuchi moved south into what is now Alabama, Georgia, and South Carolina during the seventeenth century due to pressure from the migration of the much stronger Cherokee into their ancestral lands (Jackson 2012). The main bulk of the Yuchi people were known to have resided in northern Georgia during the early nineteenth century. The Yuchi have had a long relationship with the neighboring Creek people, having been members with the Creek Confederacy. The greater part of the Yuchi nation was forced into moving to the newly designated Indian Territories out West between 1836 and 1840. The Yuchi settled in what is now Oklahoma alongside the Creek, though some Yuchi left with other Creek allies to go south into Florida, where they were absorbed into the Seminole nation (Mithun 1999).

2.3.3 Known linguistic work on Yuchi

Gatschet's (1885) work among the Yuchi in Oklahoma was the first major effort in the study of the Yuchi language that goes beyond the creation of simple word lists. This effort was followed up some time later by the German-American anthropologist Günter Wagner (1934). Wagner's grammar of Yuchi was the most comprehensive analysis of the language until Mary Linn's (2000) dissertation.

3 Previous attempts at a Siouan-Yuchi connection

Given what is now known about the location of ancestral Siouan-Catawban and Yuchi peoples in the early days of European expansion, there are a few major points worth mentioning explicitly: 1) the Yuchi lived in close proximity to Catawban and some Siouan people during the sixteenth century; 2) the Yuchi continued to live in close proximity to the Catawba and to several other Siouan languages well into the eighteenth century; and 3) it is quite possible that the Yuchi and the Siouan-Catawban peoples had lived in close proximity for longer periods of time before the the early exploration of the Southeast and Northeast by

Europeans. It is with these facts in mind that I propose the connection between Yuchi and Siouan-Catawban is more than just geographic.

3.1 Initial suspicions of common ancestry

Following Gallatin's (1836) grouping of Siouan with Catawban, other scholars began to posit relationships of other languages to Siouan. Latham (1856) first attempted to link Siouan with Iroquoian, saying that they appeared to belong "to some higher class" that may even include other languages, such as Catawba, Cherokee, Choctaw, and Caddo. Morgan (1871), after becoming interested in the Haudenosaunee following his law school research into treaties with the Cayuga, believed that the Iroquoian languages were offshoots of the Dakotan languages. Both Latham and Morgan, however, based their assumptions on very small word lists, appealing to the idea that both languages were related based on spurious data and broad claims.

3.2 Sapir's "Hokan-Siouan"

The previous attempts to link Siouan and Catawban with other languages, including Yuchi, never produced a satisfactory connection. The question of how indigenous languages were related to each other greatly interested Edward Sapir, who famously lumped Siouan with many other languages into a family called Hokan-Siouan (Sapir 1929).

Some of the groupings he made were rather spurious, based on very small or suspect sets of data. He mentioned that "a certain amount of groping in the dark cannot well be avoided in the pioneer stage of such an attempt at this," acknowledging the fact that he still had much to flesh out in his explanation for proving genetic relationships between Hokan and other indigenous languages (Sapir 1920: 289).

This pioneer stage developed into a massive putative phylogeny of North American languages, where the languages of North America are divided into six "superstocks:" 1) Eskimo-Aleut, 2) Algonquian-Wakashan, 3) Na-Dené, 4) Penutian, 5) Hokan-Siouan, and 6) Uto-Aztecan. Of these six classifications, Hokan-Siouan, later to be called Macro-Siouan, was the amalgamation of several major language families, including Siouan-Catawban, Iroquoian, Caddoan, Muskogean, Natchez, Tunica, Yuchi, and several languages of the Southwest (Sapir 1929).

This concept of a "Greater" Siouan language family has waxed and waned, but of all the possible relationships put forth, it seems that Siouan-Yuchi was one of the more accepted relationships (Campbell & Mithun 1979a).

3.3 A possible link to Proto-Gulf

In her attempts to make a case for a Gulf language family, Haas (1951; 1952) references the Proto-Siouan reconstruction by (Wolff 1950a,b,c,d) and compares two words in Proto-Siouan and Yuchi with her reconstructions for their analogues in Proto-Gulf, noting an interesting correspondence between them. On the basis of two possible proto-forms, Haas compares Wolff's Proto-Siouan construction *amą ‘land, earth’ and *miní ‘water’ with her Proto-Gulf *(ih)aγ^{w}añi(γa) ‘land’ and *ak^{w}ini ‘water.’

Haas (1951; 1952) notes that *γ^{w} and *k^{w} correspond to Proto-Siouan *m in both words, and that in both words, the vowel quality following the labialized velar is the same. That vowel, in turn, is followed by a nasal stop. She supposes that the Yuchi word *tse* ‘water’ is likewise analogous to her Proto-Gulf form, suggesting that Yuchi /ts/ originates from an earlier *k^{w}, which would make Yuchi *tse* a cognate of *ak^{w}ini. While she does not overtly say that there must be a connection between Siouan-Yuchi and her Gulf family, she certainly implies that a link is plausible, though quite difficult to prove.[3]

3.4 Chafe's “Macro-Siouan”

The Hokan-Siouan language family has seen various incarnations in the literature, most notably in Chafe's (1976) scaled-back version of a super-family that includes Siouan, Caddoan, and Iroquoian. Chafe does not claim to have conclusively proven the existence of Macro-Siouan, though he describes his findings as “tantalizing, if inconclusive.” His chief argument comes in the form of lexical resemblances shared between the three language families. Campbell (2000) devotes a sizable amount of space to the idea of Macro-Siouan, stating that it has a twenty percent probability, and a seventy-five percent confidence level, though those percentages are not explicitly given concrete metrics. His appraisal of Chafe's work is largely dominated by personal communication from Robert Rankin, who picks apart several lexical items as being false cognates.

[3] An anonymous reviewer points out that this similarity may not be due to common genetic descent, but borrowing due to long-term contact. Kaufman (2014) argues that the languages of the Lower Mississippi Valley form a Sprachbund, explaining certain similarities of Ofo and Biloxi to surrounding languages. However, this region is not thought of as a possible Siouan Urheimat (Parks & Rankin 2001: 104), which either leaves us with a reduced likelihood of these resemblances being purely due to borrowing through long-term contact or signifying that the language groups found in this region have frequently been in and out of contact with each other multiple times during prehistory.

In addition to issues with the choice of lexical items, Chafe's (1976) argument for a possible Macro-Siouan lacks any kind of systematic sound correspondences. Campbell (2000) reports that Rankin disagrees with Chafe's assessment of Caddoan preverbs being related the Siouan instrumentals, as the two morphological phenomena are believed to derive from different sources; Siouan instrumental prefixes derive from verbal roots, while Caddoan preverbs derive from incorporated nouns. Explaining the presence of non-verbal preverbal morphology as being derived from the same source is problematic, and Campbell remarks that Chafe is simply trying to connect two items that could have easily evolved independently or could be part of some areal feature. The fact that they could have arisen as areal features is interesting in of itself, as it would point to the Urheimat of each language family being close to one another at the time each language developed.

3.5 Siouan-Yuchi

Carter (1980) lists several Yuchi words in his comparison of Woccon and Catawba, showing that the two languages have some small degree of cognacy. Rankin (1996; 1998) remains agnostic on the connection of Siouan to Caddoan and Iroquoian, but also makes the case that Yuchi belongs to the Siouan family. The case for a Siouan-Yuchi connection originates from Sapir (1929), and Haas (1952) notes that Sapir had viewed Siouan and Yuchi as closely related based on a small set of lexemes.

Rankin's 1998 most recent attempt to show a relationship between Yuchi and Siouan largely skips over lexical data and concentrates on establishing a correspondence between the morphology of Siouan and Yuchi. He notes that the Proto-Siouan-Catawban word *ree 'go there' and Yuchi *ɬa* 'go' bear a strong resemblance, which Kasak (2012; 2013) builds upon by matching the Proto-Siouan-Catawban motion verbs to cognates in Yuchi, to be explained below.

4 Phono-lexical evidence

A classic method for arguing for genetic relationships is the establishment of regular sound correspondences between cognates. This section examines the posited reconstructed phonemic inventories of Yuchi and Proto-Siouan. In addition, I posit several regular correspondences between Proto-Siouan and Yuchi, adding Catawba cognates where available. While a complete reconstruction of what a Proto-Siouan-Catawban-Yuchi would look like is not within the scope of

this paper, I make the case that at least some correspondences are possible based on the limited set of cognates discovered so far. The Proto-Siouan forms come from the Comparative Siouan Dictionary (Rankin et al. 2015).

4.1 Proto-Siouan sound inventory

Rankin, Carter & Jones (n.d.) posit the following sound inventory for Proto-Siouan:

Table 1: Consonant inventory for Proto-Siouan

	Labial	Dental	Palatal	Velar	Glottal
Plosives					
preaspirated	^{h}p	^{h}t		^{h}k	
postaspirated	p^{h}	t^{h}		k^{h}	
glottalized	p’	t’		k’	’
plain	p	t		k	
Fricatives					
plain		s	š	x	h
glottalized		s’	š’	x’	
Resonants					
sonorant	w	r	y		
obstruent	W	R			

In addition to the consonants listed above in Table 1, Proto-Siouan is assumed to have had five oral vowels /a e i o u/ with two contrasting lengths, as well as three nasal vowels /ą į ų/, which also had a length distinction. Furthermore, Proto-Siouan likely had a pitch accent, marking high versus non-high pitch, and possibly a falling pitch as well. The obstruent resonants are denoted as *W and *R because it is not entirely certain what sounds they might have been, but they both have distinct reflexes in the modern languages.

4.2 Yuchi sound inventory

The modern Yuchi language, as described by Linn (2000), carries a much larger consonant inventory than that of Proto-Siouan; see Table 2.

Table 2: Yuchi consonant inventory

	Labial	Dental	Palatal	Velar	Glotta
Plosives					
postaspirated	pʰ	tʰ		kʰ	
glottalized	p’	t’		k’	
plain	p	t		k	’
voiced	b	d		g	
Affricates					
postaspirated		tsʰ	čʰ		
glottalized		ts’	č’		
plain		ts	č		
voiced		dz	j		
Fricatives					
glottalized	f’	s’	š’		
plain	f	s	š		h
Lateral Fricatives					
glottalized		’ɬ			
plain		ɬ			
Liquids					
glottalized		’l			
plain		l			
Nasals					
glottalized		’n			
plain		n			
Glides					
glottalized	’w		’y		
plain	w		y		

Yuchi features a large consonant inventory. All glottalized obstruents are ejective consonants, while glottalized sonorants are actually pronounced with creaky voice.

Yuchi's vowel inventory contains three oral front vowels /i e æ/ and three oral back vowels /a o u/. Yuchi also has a richer nasal vowel system, with at least four phonemic nasal vowels /ą ę į ǫ/. Linn (2000) mentions [æ̨], but notes that it is likely an allophone of /ę/. Wagner (1934) did not record a distinct /æ/, and wherever /æ/ is found in modern Yuchi, Wagner (1934) had written down /e/, /ę/ or /a/. Since a small number of minimal pairs can be found, Linn (2000: 44) argues /æ/ is a phoneme.

The inventories of both Proto-Siouan and Yuchi have much overlap, especially with respect to the abundance of postaspirated stops and glottalized stops and fricatives. However, accounting for the richness of the creaky-voiced sonorants is a daunting challenge. Let us begin by examining some potential cognates, and see if the two sound systems can be reconciled.

4.3 Some cognates

In looking at motion verbs in Proto-Siouan (PSi), Catawba (Cat) and Yuchi (Yu), a great similarity was found (Kasak 2013); Table 3.

Table 3: Verbs of motion in Proto-Siouan, Catawba, and Yuchi

PSi		Cat		Yu	
*rÉEh	'go there'	dáa	'go there'	ɬa	'go' (PROG)
*krÉEh	'go back there'	dukráa	'go back there'	—	—
*híi	'arrive there'	—	—	—	—
*kíi	'arrive back there'	—	—	ji	'go' (INCEPT)
*húu	'come here'	húu	'come here'	—	—
*kúu	'come back here'	dukhúu	'come back here'	gǫ	'come'
*rhíi	'arrive here'	—	—	ɬi	'arrive'
*kríi	'arrive back here'	—	—	—	—

Wherever PSi *r or *rh appears in Table 3, or wherever Catawba [d]~/r/ appears, Yuchi /ɬ/ is found. In addition, PSi *E[4] becomes /a/ in Catawba and Yuchi.

[4] This ablaut vowel sound was likely pronounced [e] in general but could become [a] or [į] under certain conditioned circumstances. The *Comparative Siouan Dictionary* reconstructs 'to go there' as *rée(he), as the ablaut vowel is not posited as a separate phoneme in Rankin et al. (2015).

Likewise, there appears to be a correspondence between PSi *k and Yuchi /g/, as well as PSi *uu and Yu /ǫ/ in *kúu ~ /gǫ/.

The form of the inceptive form of 'to go' in Yuchi /ji/ could stem from frication of the *k with the *i. The affrication of [k] before a front vowel is typologically well-attested in many different language families. Once frication occurred, the onset could have become voiced and the length distinction lost, giving a possible course of change *kíi > číi > či > *ji* 'to go.' Palatalization of a /k/ to /tʃ/ in the environment of a [+high, +front] segment is a typologically robust diachronic phenomenon: e.g., seen in the change from pre-Old English /dreŋk+j+an/ 'to cause to drink' > [drentʃan] > *drench*. Similarly, word-initial voicing was found in earlier forms of English, which is the cause for pairs like *fox::vixen*, and is still a distinctive feature of certain varieties of West Country English. This initial voicing lines up with the potential initial voicing in the change from *kúu to /gǫ/. In both Proto-Siouan and Yuchi, there is an obvious correspondence between *ii and /i/ in Table 3. See §4.4 for additional *ii and /i/ correspondences.

While not earth-shattering, the fact that all three language groups more or less appear to have retained a set of motion verbs with extremely similar semantics is suggestive of a more than casual relationship. Since these data alone are unlikely to sway anyone, additional correspondences are needed.

4.4 Correspondence with Proto-Siouan *ii

As shown earlier, the Proto-Siouan verbs *kíi 'arrive back there' and *ɾhíi 'arrive here' appear to have cognates in Yuchi: *ji* and *ɬi*. A few additional examples of *ii to /i/ correspondence appear below in Table 4:

Table 4: Correspondences between PSi *ii and Yuchi /i/

Proto-Siouan		Catawba		Yuchi	
*síi(-re)	'yellow'	siri	'clear (as water)'	ti	'yellow'
*(wa-)'íi(-re)	'blood'	iit	'blood'	we'i	'blood'
*a^{h}píi	'liver'	hipíiyą	'his liver'	y'ǫpič^{h}i	'liver'
*kíi	'arrive back there'	—	—	ji	'go' (INCEPT)
*rhíi	'arrive here'	—	—	ɬi	'arrive'

With the examples from Table 3 and Table 4, there are a total of five cognates with *ii to /i/. Seeing as how Yuchi lacks phonemic long vowels, it is unsurprising that any long vowels in Proto-Siouan would correspond to short vowels in Yuchi.

4.5 Correspondence with Proto-Siouan *y

Carter (1980) suggests that there is a relationship between PSi and Proto-Catawban *y and /'y/ and /š/ in Yuchi, along with additional cognates in Biloxi and Lakota (Lak), as seen in Table 5.[5]

Table 5: Correspondences between PSi *y and Yuchi /'y/ and /š/

Woccon		Catawba		Yuchi		Siouan	
yonne	'tree'	yana	'tree'	'ya	'tree'	*wi-yą́ą (PSi)	'tree'
yau	'fire'	ya	'fire'	'yati	'fire'	čʰą́ka (Lak)	'match'
yah-testea	'black, blue'	yači	'ashes'	'yaše	'ashes, coal'	čʰaxota (Lak)	'ashes'
yau-huk	'snake'	ya	'snake'	ša	'snake'	*yeka (PSi)	'leg(?)'
yauh	'road'	yą	'road'	'yušt'æ	'road(?)'	čʰą́ku (Lak)	'road'
—	—	-yo	'flesh, meat'	šo	'body, waist	*i-yóo (PSi)	'flesh'
—	—	čapi	'beaver'	šapa	'fox'	*wi-yáape (PSi)	'beaver'

PSi *y has a reflex of /čʰ/ in Lakota, while having reflexes of either /'y/ or /š/ in Yuchi. For 'flesh' and 'beaver' on Table 5, it is possible that Yuchi /š/ occurs instead of /'y/ due to the fact that Proto-Siouan has *y in an intervocalic environment. This could have given rise to frication. However, the word for 'tree' likewise has a *y in the environment of two vowels, so if the initial hypothesis about *y > š when in an intervocalic environment is correct, that would mean that the *wi- prefix was lost for Proto-Yuchi before the *y > š sound change took place. Otherwise, some other factor could be at work, such as a partial sound change that only affected a certain set of lexical items in Yuchi. A third possibility is that some or all of these items could be due to borrowings, and a fourth is coincidence. The underlying theme of this paper investigates a genetic connection between Yuchi and Siouan-Catawban. It is true that there is the possibility that some of these cognate sets could be due to borrowing rather than genetics, but as seen in some of the data above, and certainly more below, there are some very basic lexical items that one expects to have a low instance of borrowing: e.g., numerals, organs, highly-functional non-lexical verbs like 'to be,' etc.

One additional Proto-Siouan form was added, *yeka 'leg, thigh.' While this was not included in Carter's (1980) original list of cognates, it would be consistent with the correspondences seen previously from the motion verbs, where PSi *e

[5] Carter (1980) does not distinguish Yuchi /y/ from /'y/. All his data are represented as /'y/ under Linn's (2000) analysis.

can correlate to Catawban *a and Yuchi /a/. Furthermore, the presence of a velar in the second syllable of the Woccon word *yau-huk* 'leg' might lend some support to the relatedness of *yeka to Woccon *yau-huk* and Yuchi *ša*. I have also replaced Carter's original correspondence between Siouan-Catawban *y and Yuchi /š/ by swapping out *šagę* 'beaver' with *šapa* 'fox,' since *šapa* is a clearer lexical cognate.

4.6 Correspondence with Proto-Siouan *uu and *ųų

Given the possible time depth between Modern Yuchi and the Siouan-Catawban languages, it is understandably difficult to find large sets of data that demonstrate relatedness. As shown in Table 3, *kúu 'come back here' and Yuchi /gǫ/ appear to be cognates. If this is so, then we would expect to see other lexemes where *uu is realized as /ǫ/. In seeking out additional cognates, another correspondence between Yuchi /ǫ/ and Proto-Siouan arises. In particular, PSi *ųų appears to be associated with /ǫ/, as shown in the data below in Table 6:

Table 6: Correspondences between PSi *uu and *ųų and Yuchi /ǫ/

Proto-Siouan		Catawba		Yuchi	
*kúu	'come back here'	dukhúu	'come back here'	gǫ	'come'
*ų́ų̨ke	'hand'	iksa	'hand'	(di')ǫkʰi	'(my) arm'
*'ų́ų	'be, do'	—	—	'ǫ	'be'
*rų́ų̨pa	'two'	nąpre	'two'	nǫwę	'two'

The Yuchi word *(di)'ǫkhi* '(my) arm' resembles PSi *ų́ų̨ke. This is a likely cognate if the *uu merged with *ųų, and then lost its length distinction and lowered to /ǫ/. A common process shown on Tables 4 and 6 is that long vowels in Proto-Siouan invariably shorten in Yuchi.

While the first two items on Table 6 could conceivably come through borrowing due to contact, the bottom two items are less likely. The likelihood of borrowing a verb like 'to be' or a numeral are much lower than borrowing a lexical verb like 'come.' Body parts are also less likely to be borrowed, but not outside the realm of possibility. Cognates among these kinds of words is evidence in favor of inheritance through common genetic descent instead of borrowing through contact.

4.7 Miscellaneous cognates

In the interest of time and space, any other cognates that do not neatly fit into a specific sound correspondence appear below in Table 7. The data that follow combines suggested cognates from the *Comparative Siouan Dictionary* (Rankin et al. 2015). I have also added a possible cognate from Mandan *ke'mį'* 'vomit,' which shares a remarkably similar shape to the Yuchi *k'wǽ* 'vomit.' To the best of my knowledge, no other connections have been made in other Siouan languages to this Mandan word, and as such, no Proto-Siouan form is immediately possible. Any false assumptions and leaping to conclusions are my fault alone.

Table 7: Miscellaneous Siouan-Yuchi cognates

Proto-Siouan		Catawba		Yuchi	
*Wa-	'by cutting'	—	—	pʰa	'cut'
*ʰpą́ąhe	'bag, sack'	pą'	'hold, contain'	p'ę	'grip, squeeze'
*pą́he	'call, shout'	wǫǫ	'cry out'	p'æ	'call for'
*pʰų	'nostril'	hipį́suu'	'his nose'	dąp'i	'nose'
*k'į́(-re)	'carry on back'	kida	'carry and go'	k'ǫ	'carry'
ke'mį' (Mandan)	'vomit'	—	—	k'wǽ	'vomit'
*wiʰté	'bison, cattle'	widée	'bison'	wedi	'cow'
*íte	'face'	neen	'face'	da	'face'
*táati	'father'	nane	'father'	t'ę	'father'
*ki-si	'good, heal'	—	—	'sę	'good'
*isáapE	'black'	—	—	'ispi	'black'
*išáapE	'dark'	—	—	'išpi	'dirty'
*riih-ší	'dance'	bari	'dance'	štǰči	'dance'
*(i-)šíipe	'intestines'	—	—	čʰi	'guts'
*waRóo	'potato, ground nut'	witakii	'potato'	tʰo(biło)	'potato'
*ši-(r)-ą́te	'knee, lap'	—	—	š'ǽtʰo	'knee, lap'
*(wa-)į́į(-re)	'rock'	įįti	'stone'	ti	'rock'
*rį́į-ha(-he)	'breathe'	—	—	dihæ'e	'my breath, life'
*wįhe	'female'	į́iya	'woman'	wǽ	'woman, female'
*hą́ą	'night, darkness'	—	—	f'ą	'night'

Each of the first four items in Table 7 contains a bilabial consonant in Proto-Siouan, Catawba, and Yuchi. Three out of four of these correspond to /p'/ in Yuchi with the Siouan counterparts all having word-initial /ʰp/, /p/, and /pʰ/. It

is possible that all three of these sounds collapsed into /p'/ in Yuchi, with the plain *p corresponding to Yuchi /p/ intervocalically and in consonant clusters. Additional investigation into other cognates is needed before any kind of regular sound change can be posited.

4.8 Summary of phono-lexical evidence

There are several items that clearly look to be common to all three language groups found within the data, such as *wi^{h}té~*widée~wedi* 'bison, cattle.' A small set of sound correspondences can be argued. Namely, *r and *rh can map to /ɬ/, *y can map to /'y/ and /š/, *ii maps to /i/, and *uu and *ųų map to /ǫ/.

Further study of older vocabulary from Wagner's (1931) texts may shed additional light on additional cognates. The lexical items used thus far are from Linn (2000) and Rankin et al. (2015). It is extremely difficult to find an adequate list of cognates, much less postulate sound correspondences that might lead to the recreation of hypothetical Proto-Siouan-Yuchi forms. This difficulty arises from Linn's pervasive use of paradigms to illustrate Yuchi morphophonology, and as such, it is not in itself as rich a source of lexical data as a dictionary might be. Additional work is certainly needed in this area. However, the fact that even some regular sound changes can be discerned, as well as the sharing of certain high-frequency lexical items like body parts, numerals, copulas, and verbs of motion, provide evidence of a connection that is not inherently one of language contact.

5 Morphological evidence

Lexicostatistics can only carry one so far before running into a dead end, in addition to being controversial in itself. Regardless of whatever half life a word may have in a language, a language's morphology is much more resistant to decay. The fact that the humble English word *am* can trace its roots directly back to Proto-Indo-European *esmi shows that morphology has the kind of staying power that open category words simply cannot match.

To date, the most prolific and explicit analysis of Sapir's (1929) idea of a special relationship between Siouan and Yuchi has been undertaken by Robert Rankin (1996; 1998). Through side-by-side comparisons of particular affixes and verbal paradigms, Rankin carefully argues that while certain lexical similarities may be chalked up to borrowings, it is difficult for certain morphological idiosyncrasies to be wholesale borrowed as well. Much of his treatment of the issue assumes

the possibility of cognacy through genetic inheritance rather than borrowing, a treatment I have adopted throughout this paper.[6]

His argument rests on four key points. He shows that there is a strong similarity in the classificatory systems of Proto-Siouan, Catawba, and Yuchi, as well as in their pronominal prefix morphology. That prefix morphology also has a marked interaction with preverbs for first person plural forms. A somewhat lesser but noteworthy point is that Siouan and Yuchi both feature fricative sound symbolism to add gradience to a verb.

5.1 Siouan, Catawban, and Yuchi classifiers

One particularly productive affix in Siouan languages comes from PSi *ko-, which is found on kinship terms that are possessive. The examples in Table 8 are from Mandan (MA), Tutelo (TU), Dakota (DA), Ofo (OF), Dhegiha (DH), and Ho-Chunk (HC).

Mandan appears to have the most robust use of *ko-, using it as a third-person possesion marker for family members. In several other languages, as seen above, *ko- seems to have simply melded onto the noun. Compare PSi *ko- with *ku-*, a prefix in Catawba that has a similar distribution. Like PSi *ko-, Catawba *ku-* appears as part of words denoting people. Table 8, as well as Tables 9 through 13 on the following pages are adapted from Rankin (1998).

Yuchi *go-* is described as being a human-specific prefix, often used for people who are Yuchi, with a different marker used for non-Yuchi. Proto-Siouan similarly has analogous prefixes, as shown above. The *go-* prefix is extremely productive in Yuchi. There is a different third-person singular marker on verbs that is homophonous that marks an impersonal subject, or a subject who is not Yuchi. It is quite possible that the forms are related.

A second classifier is PSi *wi-, which marks animals, food, and items in nature, as shown on Table 11. An identical distribution can be found in Catawba, as seen in Table 12, where morphology bearing the shape /wi-~wį-/ appears on animate nouns, certain foods, and natural phenomena. The overlap in both the phonological shape and semantics of PSi *wi- and Catawba /wi-~wį-/ strongly suggest that they are inherited from a common ancestor.

[6] There is the possibility for language contact to have influenced the presence of cognates across these language families. Given that this paper does not purport to be an exhaustive treatment of all scenarios under which a particular lexical item can have passed into one language to or from another, I propose cognacy through common descent as my main hypothesis. Fleshing out individual borrowings and determining the directionality of their borrowings is outside the scope of this paper and remains the subject of further study.

Table 8: Siouan personal classifiers

košųka	'his younger brother' (MA)	**koo**mihąą	'girl' (TU)
sųka**ku**	'younger brother' (DA)	hų́**ku**	'his mother' (DA)
hkóra	'friend' (DH)	hiča-**kó**ro	'friend' (HC)
kosíke	'woman's brother-in-law' (MA)	əkifhų**tku**	'little brother' (OF)
		ee-**ko**weei	'their chief' (TU)

Table 9: Catawba personal classifiers

kurii	'son'	**ku**koo	'girl'
kotóne	'host'	**kə**neyana	'his father'
ya **ku**re nanéwa	'her father'	**ka**tiyíise	'youngest son'

Table 10: Yuchi personal classifiers

golaha	'grandmother'	**go**jiɬą	'giant'
gotané	'brother'	**go**k'alá	'relatives'
go'ę	'baby'	**go**t'e	'husband, man'

Table 11: Siouan non-personal classifiers

*__wi__yáape	'beaver'	*__wi__^{h}té	'bison'
*__wi__šų́ke	'dog'	*__wi__^{h}tóxka	'fox'
*__wi__ráa	'fire'	*__wi__ŕ̨	'water'

Table 12: Catawba non-personal classifiers

*__wi__dée	'bison'	*__wi__mba	'barred owl'
*__wi__tka	'owl'	*__wí__dyu	'meat'
*__wi__tą'	'rat'	*__wi__čawa	'night'

Table 13: Yuchi non-personal classifiers

wedi	'cow'	**we**'ya	'deer'
weła	'hawk'	**weš**i	'sofki, soup'
weč^{h}æ'læ	'lightening'	**wet**'æ	'rainbow'

Both Proto-Siouan and Catawba have some reflex of /wi-/ as their non-personal classifier. Catawba and Proto-Siouan both use *wi- not only to mark animals, but also common foods and weather- and nature-related words. Yuchi, on the other hand, has *we-*; see Table 13.

Rankin (1998) notes that while the overlap of these two classificatory affixes is striking, it is conceivable that a classifier system that is two-pronged in nature could have been borrowed. Additional support that this bit of morphology is likely via common genetic descent appears below.

While these similarities in nominal affixation may be chalked up to coincidence, further investigation is needed into verbal and other morphology to determine if any deeper connection between Siouan and Yuchi is plausible.

5.2 Pronominal morphology

All three languages at hand are active-stative languages with SOV word order. Often, overt subjects are omitted if clear in the discourse or from the verb. This is easier to do in Yuchi, as Yuchi has a rather robust system of third-person marking, which stands in stark contrast to the zero-marking of third-person that Siouan and Catawban languages generally have.

Rankin (1996; 1998) posits that any non-Siouan language whose pronominal morphology is similar and has comparable idiosyncrasies is a likely candidate to be related to Siouan-Catawban. Unlike individual words, morphology is not so easily borrowed. In Proto-Siouan, there are four reconstructed person markers: first person singular, second person, inclusive first person plural, and exclusive first person plural. Rankin (1998) offers up the following set of correspondences, providing two different pronominal series for Yuchi: a *di-* series for verbs referring to activities, processes, or motion (Linn 2000: 130), and a *do-* series for transitive verbs with specific objects (Linn 2000: 178). Table 14 below lists subject prefixes in Proto-Siouan, Catawba, and Yuchi to illustrate the similarity between all three languages, particularly between Catawba and Yuchi.

At first glance, the similarity between all three languages is not close enough to conclusively demonstrate genetic affiliation. The second person singular seems

Table 14: Proto-Siouan, Catawba, and Yuchi subject prefixes

	PSi	Catawba	Yuchi	
1sg	*wa-	**dV-**	**di-**	**do-**
2sg	***ya-**	**ya-**	ne-	**yo-**
3sg	*∅-	**hi-**	**hę**-/se-/we-	hǫ-/syo-/yǫ-
1pl.incl	***’ų-**	ha-	**’ǫ-**	**’ǫ-**
1pl.excl	***rų-**	ha-	**nǫ-**	**nǫ-**
2pl	*ya-	wa-	’ane-	’ąyo-
3pl	*∅-	a-/**i-**	hǫ-/’o-/**’i**-/we-	’hǫ-/’o-/’yǫ-

to be the most promising, with all three languages beginning their second person with a [y] sound. Rankin (1998) points out in the endnotes that there is some variation between *r and *y in second person verbal paradigms in some languages, which could be the reasoning behind the variation between /n/ and /y/ in the second person Yuchi pronominals.

The Proto-Siouan and Yuchi first person plural markers are extremely similar, with *r usually being realized as [n] when followed by a nasal. In addition, the correspondence between *ų and /ko/ goes along with the correspondence previously done for *uu and *ųų and /ǫ/ in Yuchi. Catawba also has an independent first person pronoun *inu*, which seems analogous to the first person plural, as well as having the first person plural object marker *nų-* (Voorhis 1984). The only other evidence for an exclusive pronoun in Siouan comes from Mandan, which has *rų-* as its sole first person plural marker.

The first person singular for Catawba and Yuchi is extremely similar. Often, [d] in Catawba is the word-initial allophone for /r/, with [r] surfacing intervocalically (Rudes 2007). The environments in which Yuchi [d] surfaces are unclear; Gatschet (1885), Wagner (1934), and Linn (2000) do not spend substantial portions of their works dealing with variation and allophony, preferring to dive right into the meat of morphology. If Catawba /d/ does indeed stem from an underlying /r/, then it would mean some opaque path to explain how Proto-Siouan wound up with *w- for first person, while the other two languages had [d-].

Rankin (1998) wonders if some ancient system of allophony could be at work here, but it is entirely possible that all three languages started out with the same segment, and there has been a radical change in one or more groups. One needs only to look at the famous example of PIE *dw- > Armenian *erk-* to be reminded

that sound changes can move very far from their original source, given enough time and innovations. It is also possible that all three languages display a reflex of some earlier sound no longer reflected. I am tempted to think about Proto-Algonquian *n- for first person, and how few steps one would need to take to turn [n] into [d], or [n] into [m] and then [w]. While I am not suggesting that is what happened, I am noting that such a particular series of sound changes would not be completely beyond the realm of possibility.

Rankin (1998) does not address third person marking or second person plural marking.[7] Using Rudes's (2007) description of the Catawba verbal template along with Linn's (2000) Yuchi grammar, we can add to previous analyses, which does add further evidence of a closer affinity between Catawba and Yuchi than to Proto-Siouan, namely in the third person singular, where Catawba has *hi-*, while masculine third person subjects in Yuchi are marked with *hę*. Additional evidence of similarity in the third person marking of Catawba and Yuchi is in Linn's (2000: 133) description of an alternate third person plural marker *'i-*, which is identical to one of the Catawba allomorphs for third person plural subjects. There seems to be no clear candidate for a shared second person plural marker among the languages above. This point notwithstanding, Catawba and Yuchi appear to have cognate first person singular and third person plural marking, while Yuchi and Proto-Siouan share first person plural marking for both inclusive and exclusive. Second person singular marking is shared across all three groups, and the Yuchi *ne-* may also be shared with Proto-Siouan second person patient marker *yį-, given that there could have been nasal assimilation onto the *y where *y > n /__ [+nasal]. We see this process yield Mandan and Ho-Chunk *nį-* and Lakota *ni-*, which are quite similar in shape to Yuchi *ne-*.

5.3 Templatic morphology

Both Siouan-Catawban and Yuchi follow a templatic system of morphology. Most notably, Siouan and Catawban first-person and second-person pronominals tend to get trapped between the verb and preverbal morphology, while the first person plural is to the left of any other morphology, as the data on Tables 15 and 16 from Rankin (1998) show.

Table 15 demonstrates the morphologically marked behavior of the preverb *ée with respect to pronominals. Only the first-person plural can appear before the

[7] Yuchi has a wider array of third person markers, depending on whether referring to the gender of the referent and whether the referent is a Yuchi or not. The three-way split for third person on Table 14 refers to male Yuchi subjects, female Yuchi subjects, and non-Yuchi subjects.

Table 15: Pronominals within the verb 'to say' in Siouan

	PSi	Dakota	Mandan
1.SG	'éep**h**e	é**p**ha	éep**e**'š
2.SG	'éeše	éha	éet**e**'š
3.SG	éehe	éya	éehero'š
1.PL	'**ų**'éehe	'**ų**'éya	**r**éehero'š

preverb. Every other pronominal is trapped between the preverb and the verb, as demonstrated clearly in Proto-Siouan and Mandan. Mandan, in particular, is prone to creating portmanteaux with the first person plural marker and the preverb (Hollow 1970). This same process of portmanteau creation can be seen at work in Yuchi, as seen below.

Siouan has four main preverbs: *íi-, *áa-, *óo-, and *ée-. The preverbs all attract primary stress or high tone, depending on the language. This prominence attraction is also found in the Yuchi instrumental preverb *hi-*. Linn (2000) argues that *hi-* is simply a third-person non-agent marker, but her analysis does not explain how it is that such a preverb could have the same phonetic idiosyncrasy of high tone attraction in addition to causing the first-person plural marker to move out of the expected spot next to the verb and migrate all the way to the left of the verb.[8] The exact same scenario appears for the preverbal element *k'ą-* 'something.' For both *k'ą-* and *hi-* to yield this marked position for the first person plural prefix would be an extraordinarily rare occurrence, but for this construction to also be found in Siouan and Catawban languages would be highly improbable.

Another idiosyncrasy shared by Yuchi and Siouan is the tendency to create portmanteaux out of first person plural markers and preverbs. On Table 16, both *hi-* and *'ǫ-* and *hi-* and *nǫ-* combine their vowels to create new, more complex affixes: *hi-* plus *'ǫ-* become *'ę-*, and *hi-* plus *nǫ-* become *nę*. The prefix *k'a-* 'something' itself resembles a portmanteau of the Siouan *áa- preverb, which denotes a comitative action and the Siouan reflexive marker *ki-. The prefix *k'a-* is also the reciprocal marker in Yuchi, which affixes immediately to the left of the verb

[8] The presence of high tone on these preverb-like elements in Yuchi is reminiscent of the placement of high tone in Mandan when a word contains a preverb. If *hi-* truly is a third person singular marker of some sort, then it bears even more resemblance to Catawba, which also has *hi-*, but for third person singular subjects. If *hi-* is related to the Proto-Siouan preverbs, perhaps it is related to *íi-.

Table 16: Pronominals within the verb with Yuchi Preverbs

1SG	k'ą**d**a	hiča
	'I carry something'	'I find something'
2SG	k'ąša	hiša
	'you carry something'	'you find something'
1PL.INCL	'**ǫ**k'ąła	'**ę́**ła
	'we carry something'	'we find something'
1PL.EXCL	'**nǫ**k'ąła	**nę́**ła
	'we carry something'	'we find something'

before any pronominals are added on. Siouan reciprocals and reflexives have the same distribution.

(1) a. ***Nǫk'at^{h}et^{h}e.*** Yuchi
***nǫ-k'a**-t^{h}e~t^{h}e*
1PL.EXCL.AGT-RECP-hit~ITER
'We beat each other up/We hit each other repeatedly.' (Linn 2000: 250)

b. ***Nųkiruškapo'š.*** Mandan
***rų-ki**-ru-škap-o'š*
1PL.AGT-RECP-by.hand-pinch-IND.MASC
'We pinch each other.' (Hollow 1970: 440)

The similarity between Yuchi and Siouan – represented here with Mandan – is that both languages have a set of inner and outer pronominals: inner pronominals are first person singular and second person markers, which appear closer to the verb than preverbal elements like the Proto-Siouan applicative preverbs *áa-, *íi, etc. and morphology in Yuchi that intrinsically bears high tone and appears closer to the left edge of the word than these inner pronominals, such as *hi-*, *k'a-*, and *k'ą-*. First person plurals in both Mandan and Yuchi are treated as outer pronominals, meaning they appear further to the left than a preverb.

The data above in (1) demonstrate the similarity in not only the phonetic realization of the first person plural for both Mandan and Yuchi, but also the similarity in sound and semantics for both of their reciprocal markers. An additional potential cognate with Siouan is the inceptive marker in Yuchi, which is also *k'a*, while the inceptive marker in Mandan is *ka*. This similarity could be coincidental, but Siouan non-pronominal affixes have a somewhat large degree of polysemy.

This strong similarity could be due to ancient borrowing or some old areal feature, but when taken as a whole, the summation of these similarities begins to beg the question of whether we are looking at ancient borrowings or ancient features inherited from an ancestral language.

Further analysis of non-pronominal morphology in Yuchi is needed, but the preliminary look taken by Rankin (1998) plus the amount added here points to the conclusion that Sapir (1929) might not have been far off the mark in declaring a genetic relationship between Siouan and Yuchi.

5.4 Sound symbolism and ablaut

The final component in Rankin's (1998) analysis of the relationship between Siouan and Yuchi is that both languages share a fricative sound symbolism. A sound symbolism is a relationship between the place of articulation and some kind of scalar contrast. For example, in Mandan, the sounds [s š x] are in a sound symbolism relationship. In Mandan, two examples of such words are *seroo* 'to jingle' and *xeroo* 'to rattle,' or *sąsi* 'slick' and *šąši* 'smooth.' In Yuchi, such examples are *'ispi* 'black' and *'išpi* 'dirty,' and *č^{h}ała* 'pink' and *ts^{h}yała* 'red.'

In addition to this sound symbolism, Yuchi appears to have some vestiges of a Siouan-type ablaut system. In Siouan languages, there are certain vowels, usually marked with a capital letter in dictionaries, that will change their qualities under the influence of following morphology. In Mandan, there is a class of vowels that typically are realized as [e], but when followed by certain affixes are realized as [a]. Such vowels are marked as /E/ in underlying representation. One suffix that triggers ablaut is the second person pluralizer /-rįt/:

(2) a. *rareeho'š*
ra-rEEh-o'š
2.AGT-go.there-IND.MASC
'you (SG) went'

b. *raraahinįto'š*
ra-rEEh-rįt-o'š
2.AGT-go.there-2PL-IND.MASC
'you (PL) went'

In the Dakotan languages, one such element that triggers ablaut is the future marker *kte*. In Lakota, there are three ablaut grades: a-grade, e-grade, and į-grade. A-grade is the default form, while e-grade is triggered when the word is the last

in a sentence, or various other morphology is present. The last grade is what is most interesting here, as it causes [a] to become [ą] (Ullrich 2008).

In Yuchi, Linn (2000) states that nasalizing the final vowel in the stem denotes future tense. A similar process occurs in Lakota.

(3) a. *Weda.* Yuchi
'I'm going (now).' (Linn 2000: 279)

b. *Wedą.* Yuchi
'I'm going to go (soon).' (Linn 2000: 279)

c. *Blé.* Lakota
'I'm going.' (Ullrich 2008: 75)

d. *Mnį́ kte.* Lakota
'I'm going to go.' (Ullrich 2008: 75)

Though there is no overt future marker in Yuchi, the nasality added to the final vowel is very reminiscent of ą-ablaut in Siouan. To date, there has been no mention of a connection between future marking in Yuchi and Siouan ablaut, so this potential morphological cognate is deserving of further study. While not conclusive, this nasalization does raise further questions as to what other morphologically conditioned sound changes are taking place, and how they tie into Siouan-Catawban as a whole.

6 Conclusion

Proving an ancient genetic relationship is no easy task. If Parks & Rankin (2001) are hitting near the mark in estimating the split of Siouan and Catawba from each other around four thousand years ago, how much deeper would we have to go in order to account for the massive lexical and grammatical difference from Yuchi? Though my labors have just scratched at the surface of similarities between Yuchi and Siouan, some small signs of hope can yet be found for proponents of a Siouan-Yuchi family. There are indeed lexical cognates that can be used to create sound correspondences, and there are ample records showing that Catawbans and Siouans were close to the Yuchi during the early days of colonization in North America. Thus, a relationship is plausible, given the close proximity in which these peoples lived and their sharing several very elementary words. With the location of Catawba and the varieties of Ohio Valley Siouan and Yuchi firmly recorded in the Southeast, the center of gravity would suggest a homeland somewhere near the Ohio River Valley.

While Rankin (1996) expresses tempered pessimism regarding the depth at which linguists are capable of probing for relatedness, the morphological data gathered by Rankin (1998), in addition to a few affixes of interest, only cause more questions to be raised. It is quite possible that there are many more correspondences to be found, but that is a task best saved for a different paper.

My principal goal was to investigate the idea of a Siouan-Catawaban-Yuchi family, given Rankin; Rankin's (1996; 1998) previous efforts to investigate deeper genetic relationships between Siouan and other languages of North America. While these results are not definitive, there is room for some optimism on this front. I believe that there is substance to the idea that Siouan and Yuchi are distant cousins, and I encourage further study into the topic of Siouan's relationship with neighboring languages and groups with a particular eye towards Yuchi. Since Rankin's attempts to connect Yuchi with Siouan-Catawban, similar connections have been made by Vajda (2010), who puts forth a strong case for a distant genetic relationship between the Na-Dené languages of the Americas and the Yeniseian languages of Siberia. While lexical cognates are not overly common, they do exist, but the most compelling evidence is in the similarity of both the inflectional morphology and the sequencing of affixes within the verbal template. A similar argument occurs in §5 of this paper, by which I conclude that there is some validity to Rankin's claim that Yuchi is distantly related to Siouan-Catawban.

One particular avenue for study is to go through the lexical information from the few extant Yuchi language grammars, field notes, and attempted dictionaries housed at the American Philosophical Society to put together a Yuchi database. With a larger repository of Yuchi vocabulary, the attempt to make lexical connections between Yuchi and Siouan might be more fruitful. Another important task is to continue the work on Catawba started by Rudes (2007) before his passing to investigate whether there are additional similarities to be found between Yuchi and Catawba, since they share more pronominal morphology with each other than they do with Proto-Siouan.

Acknowledgment

I would like to thank Steve Anderson for his comments on a earlier draft of this paper, as well as my two anonymous reviewers for their insightful comments. Ultimately, I am grateful to the late Bob Rankin for all his insight on this topic over the years and for sharing his "tempered" enthusiasm about the deeper genetic relationships between Siouan and other languages of North America.

Abbreviations

1, 2, 3	first, second, third person	MA	Mandan
		MASC	masculine
AGT	agent	OF	Ofo
Cat	Catawban	PCa	Proto-Catawban
DA	Dakota	PIE	Proto-Indo-European
DH	Dhegiha	PL	plural
EXCL	exclusive	PROG	progressive
HC	Ho-Chunk	PSi	Proto-Siouan
INCEPT	inceptive	RECP	recipient
INCL	inclusive	SG	singular
IND	indicative	TU	Tutelo
ITER	iterative	Yu	Yuchi
Lak	Lakota		

References

Adelung, Johann C. & Johann S. Vater. 1816. *Mithridates oder allgemeine Sprachenkunde, mit dem Vater Unser als Sprachprobe in bey nahe fünfhundert Sprachen und Mundarten.* Vol. 3, part 3. Berlin: Vossische Buchhandlung.

Bloomfield, Leonard. 1925. On the sound-system of Central Algonquian. *Language* 1(4). 130–156.

Booker, Karen M., Charles M. Hudson & Robert L. Rankin. 1992. Place name identification and multilingualism in the sixteenth-century Southeast. *Ethnohistory* 39(4). 399–451.

Campbell, Lyle. 2000. *American Indian languages: The historical linguistics of Native America.* Oxford: Oxford University Press.

Campbell, Lyle & Marianne Mithun. 1979a. Introduction. In Lyle Campbell & Marianne Mithun (eds.), *The languages of Native America: Historical and comparative assessment*, 3–69. Austin: University of Texas Press.

Campbell, Lyle & Marianne Mithun (eds.). 1979b. *The languages of Native America: Historical and comparative assessment.* Austin: University of Texas Press.

Carter, Richard T., Jr. 1980. The Woccon language of North Carolina: Its genetic affiliations and historical significance. *International Journal of American Linguistics* 46(3). 170–182.

Chafe, Wallace L. 1976. *The Caddoan, Iroquoian, and Siouan languages* (Trends in Linguistics 3). Berlin: de Gruyter Mouton.

Crawford, James M. 1970–1973. *Field notebooks*. Philadelphia: Library of the American Philosophical Society.

Crawford, James M. 1979. Timucua and Yuchi: Two language isolates of the Southeast. In Lyle Campbell & Marianne Mithun (eds.), *The languages of Native America: Historical and comparative assessment*, 327–354. Austin: University of Texas Press.

Day, Gordon M. 1968. Iroquois: An etymology. *Ethnography* 15(4). 389–402.

Dorsey, James Owen. 1893. The Biloxi Indians of Louisiana. In Frederic W. Putman (ed.), *Proceedings of the Association for the Advancement of Science*, vol. 43, 267–287. Salem, MA: The Salem Press.

Dunn, Michael. 2009. Contact and phylogeny in Island Melanesia. *Lingua* 119. 1664–1678.

Gallatin, Albert. 1836. *A synopsis of the Indian tribes within the United States east of the Rocky Mountains, and in the British and Russian possessions of North America* (Archæologia americana: Transactions and collections of the American Antiquarian Society 2). Cambridge, UK: Cambridge University Press.

Gatschet, Albert Samuel. 1885. The Yuchi tribe of Indians, and its language. *Science* 5(112). 253.

Gatschet, Albert Samuel. 1900. Grammatic sketch of the Catawba language. *American Anthropologist* 2(3). 527–549.

Goddard, Ives. 1975. Algonquian, Wiyot, and Yurok: Proving a distant genetic relationship. In Marvin D. Kinkade, Kenneth L. Hale & Oswald Werner (eds.), *Linguistics and anthropology in honor of C. F. Voegelin*, 249–262. Lisse, NL: Peter de Ridder Press.

Haas, Mary. 1951. The Proto-Gulf word for 'water' (with notes on Siouan-Yuchi). *International Journal of American Linguistics* 17(2). 71–79.

Haas, Mary. 1952. The Proto-Gulf word for 'land' (with notes on Proto-Siouan). *International Journal of American Linguistics* 18(4). 238–240.

Hale, Horatio. 1883. The Tutelo tribe and language. *Proceedings of the American Philosophical Society* 21(114). 1–47.

Hanna, Charles A. 1911. *The wilderness trail, or the ventures and adventures of the Pennsylvania traders on the Allegheny path with some new annals of the old West, and the records of some strong men and some bad ones*. Vol. 2. New York: G. P. Putnam's Sons.

Hollow, Robert C. 1970. *A Mandan dictionary*. Berkeley: University of California – Berkeley (Doctoral dissertation).

Jackson, Jason B. 2012. *Yuchi Indian histories before the removal*. Lincoln: University of Nebraska Press.

Jennings, Francis. 1990. *The ambiguous Iroquois empire: The covenant-chain confederation of Indian tribes with English colonies from its beginnings to the Lancaster Treaty of 1744*. 1st edn. New York: W. W. Norton & Company.

Kasak, Ryan M. 2012. *Treatment of vertitivity in Siouan*. Paper presented at the 31st annual meeting of the Siouan and Caddoan Languages Conference, University of Kansas.

Kasak, Ryan M. 2013. *A reconstruction of verbs of motion in Proto-Siouan*. Paper presented at the annual meeting of the Society for the Study of Indigenous Languages of the Americas, Boston.

Kaufman, David V. 2014. *The lower Mississippi Valley as a language area*. Lawrence: University of Kansas (Doctoral dissertation).

Koontz, John E. 1995a. *Michigamea as a Siouan language*. Paper presented at the 15th annual Siouan and Caddoan Languages Conference, University of New Mexico – Albuquerque.

Koontz, John E. 1995b. *Notes on Michigamea for Emanuel Drechsel*. Boulder: University of Colorado.

Latham, Robert G. 1856. On the languages of northern, western, and central America. *Transactions of the Philological Society* 3(1). 57–114.

Linn, Mary S. 2000. *A grammar of Euchee (Yuchi)*. Lawrence: University of Kansas (Doctoral dissertation).

Mann, Charles C. 2006. *1491: New revelations of the Americas before Columbus, 2nd ed*. New York: Random House.

Mithun, Marianne. 1999. *The languages of Native North America* (Cambridge Language Surveys). Cambridge: Cambridge University Press.

Morgan, Lewis Henry. 1871. *Systems of consanguinity and affinity of the human family*. Washington: Smithsonian Institution.

Parks, Douglas R. & Robert L. Rankin. 2001. Siouan languages. In Raymond J. DeMallie & William C. Sturtevant (eds.), *Plains*, vol. 13 (Handbook of North American Indians), 94–114. Washington: Smithsonian Institution.

Rankin, Robert L. 1996. Deeper genetic relationships in North America: Some tempered pessimism. In David Librik & Roxane Beeler (eds.), *Proceedings of the 22nd annual meeting of the Berkeley Linguistics Society: Special session on historical issues in Native American languages*, 117–128. Berkeley: University of California – Berkeley.

Rankin, Robert L. 1998. *Siouan-Catawban-Yuchi genetic relationship: With a note on Caddoan*. Paper presented at the 18th annual Siouan and Caddoan Languages Conference, Indiana University – Bloomington.

Rankin, Robert L. 2004. *The Ofo language of Louisiana: Philological recovery of grammar and typology*. Paper presented at the 3rd Language and Variety in the South Conference, University of Alabama – Tuscaloosa.

Rankin, Robert L. 2010. *The place of Mandan in the Siouan language family*. Paper presented at the 30th annual Siouan and Caddoan Languages Conference, Northeastern Illinois University, Chicago.

Rankin, Robert L., Richard T. Carter & A. Wesley Jones. n.d. *Proto-Siouan phonology and grammar*. Unpublished manuscript.

Rankin, Robert L., Richard T. Carter, A. Wesley Jones, John E. Koontz, David S. Rood & Iren Hartmann (eds.). 2015. *Comparative Siouan dictionary*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://csd.clld.org (25 September, 2015).

Reed, Swanton John. 1952. *The Indian tribes of North America* (Bureau of American Ethnology Bulletin 145). Washington: Government Printing Office.

Rudes, Blair A. 2007. *Catawba grammar; Word-order frequency*. Unpublished manuscripts. Charlotte.

Sapir, Edward. 1920. The Hokan and Coahuiltecan languages. *International Journal of American Linguistics* 1. 280–290.

Sapir, Edward. 1929. Central and North American languages. *Encyclopaedia Britannica* 5. 138–141.

Swanton, John Reed. 1909. A new Siouan dialect. In Franz Boas (ed.), *Putman anniversary volume: Anthropological essays presented to Frederic Ward Putnam in honor of his seventieth birthday*, 477–486. New York: G. E. Stechert.

Ullrich, Jan F. (ed.). 2008. *New Lakota dictionary*. 1st edn. Bloomington, IN: Lakota Language Consortium. http://www.lakotadictionary.org/nldo.php.

Vajda, Edward J. 2010. A Siberian link with Na-dene languages. In Kari James & Ben A. Potter (eds.), *The Dene-Yeniseian connection*, vol. 5 (Anthropological Papers of the University of Alaska), 33–99. Fairbanks: University of Alaska – Fairbanks Department of Anthropology.

Voorhis, Paul. 1984. Catawba morphology in the texts of Frank Speck and of Matthew Red Thunder Cloud. In Letta Strantzali (ed.), *Studies in Native American languages III*, vol. 9 (Kansas Working Papers in Linguistics 9), 1–30. Lawrence: University of Kansas.

Wagner, Günter. 1931. *Yuchi tales*. Franz Boas (ed.) (Publications of the American Ethnological Society 13). New York: G. E. Stechert.

Wagner, Günter. 1934. Yuchi. In Franz Boas (ed.), *Handbook of American Indian languages*, vol. 3, 293–384. New York: Columbia University Press.

Wolff, Hans. 1950a. Comparative Siouan I. *International Journal of American Linguistics* 16(2). 61–66.
Wolff, Hans. 1950b. Comparative Siouan II. *International Journal of American Linguistics* 16(3). 113–121.
Wolff, Hans. 1950c. Comparative Siouan III. *International Journal of American Linguistics* 16(4). 168–178.
Wolff, Hans. 1950d. Comparative Siouan IV. *International Journal of American Linguistics* 17(4). 197–204.

Chapter 2

Two Siouan languages walk into a sprachbund

David Kaufman

In this paper, I examine two Siouan languages, Biloxi and Ofo, and how they have been influenced by their participation in the Lower Mississippi Valley (LMV) language area, or *sprachbund*, which I previously analyzed in-depth in my dissertation. The LMV sprachbund shows the convergence of eight languages of different language families, including four isolates: Atakapa, Biloxi, Chitimacha, Choctaw-Chickasaw, the Mobilian Trade Language (MTL), Natchez, Ofo, and Tunica, from ca. 500 CE to 1700 CE. This sprachbund involves moderate levels of copying, not only of lexical items but also of grammatical elements. As members of this sprachbund, Biloxi and Ofo share several phonetic and phonological, morphological, and lexical features with other LMV languages, which are examined here.

1 Introduction

In this paper, I examine two Siouan languages, Biloxi (ISO 639-3: bll) and Ofo (ISO 639-3: ofo), and how they have been influenced by their participation in the Lower Mississippi Valley (LMV) language area, or SPRACHBUND[1] (Kaufman 2014: 3). As members of this sprachbund, Biloxi and Ofo share several phonetic, phonological, morphological, and lexical features with other LMV languages, which are Atakapa, Chitimacha, Choctaw-Chickasaw,[2] Mobilian Trade Language (MTL; also called "Mobilian Jargon"), Natchez, and Tunica. All of these languages, with the exception of Biloxi and Ofo (Siouan), and Choctaw-Chickasaw and MTL (Muskogean), are isolates with no known living linguistic relatives.

I define the Lower Mississippi Valley (LMV) as an area extending from about 260 miles (418 km) west of the Mississippi River eastward to Mobile Bay on the Gulf of Mexico, a total of about 380 miles (612 km), and about 425 miles (684

[1] *Sprachbund* is a German term literally meaning 'language union'.

[2] Since Choctaw and Chickasaw are generally mutually comprehensible, I combine them here into one unit.

David Kaufman. 2016. Two Siouan languages walk into a sprachbund. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 39–63. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.121 DOI:10.17169/langsci.b94.121

km) northward from the Gulf of Mexico toward the vicinity of the Tombigbee and Arkansas Rivers, an area encompassing 144,600 square miles (496,600 square km). This area encompasses what is now northern Arkansas, Mississippi, and Alabama, southeastern Oklahoma and eastern Texas over toward central Alabama, and includes all of the modern states of Louisiana and Mississippi; see Figure 1. My examination of the LMV reveals this region to be a language area on par with the Balkans (Eastern Europe), South Asia (India), the Amazon Basin, and other such *Sprachbünde* around the world.

Biloxi and Ofo, along with Tutelo, form part of the Ohio Valley, or Southeastern,[3] branch of the Siouan language family. While it is unknown exactly when Biloxis and Ofos reached the LMV, we do have evidence that the Ofos (Mosopeleas) migrated into the LMV in the seventeenth century. Biloxis are harder to pin down, but given the scraps of language data available to us based on toponyms, it is likely that ancestral Biloxis once occupied the southern Appalachian mountain region, probably in the Cumberland Plateau and areas of modern eastern Tennessee near the Tennessee River (see Rankin 2011 and footnote 4) from where they likely migrated southward to the Gulf coast.[4]

Linguists have long used the Stammbaum ('family tree') model of linguistic ancestral descent, which is usually described with a biological metaphor: the "genetic" origins of languages, which insist on a "single-parent source and its belief that practically all language change resulted from internal causes" (Winford 2003: 7). In this case, Proto-Siouan would be the "single-parent source," while the modern Siouan languages, including Biloxi and Ofo, would be its descendants. However, language change can also arise from external causes through language contact, where similarities arise not through genetic affiliation but through close cultural and linguistic contact. Language areas arise when languages, which may or may not be "genetically" related, come into close contact through such things as trade, alliance, intermarriage, and intergroup gatherings, thereby encouraging "diffusion of linguistic features across geographically adjacent languages"

[3] I use the term *Southeastern* rather than "Ohio Valley" for this branch of Siouan, since habitation for all members, with the exception of Ofos (Mosopeleas), of this branch in the Ohio Valley is uncertain.

[4] Further language evidence, based on toponyms, indicates that the Biloxi word for 'salt', *waasi*, may occur in a couple of place names in this region: Ouasioto (*Waasi-oto*?) and Guasile (*Waasi-le*?). The first is the old name for Cumberland Gap, which was indeed situated near a salt-producing mound town (Meyer 1925). However, I have no good linguistic explanations for the suffixes *-oto* and *-le* in these names, which do not immediately appear to be Biloxi based on extant data, so that, though intriguing, a definite correlation cannot be made with Biloxis or their ancestors.

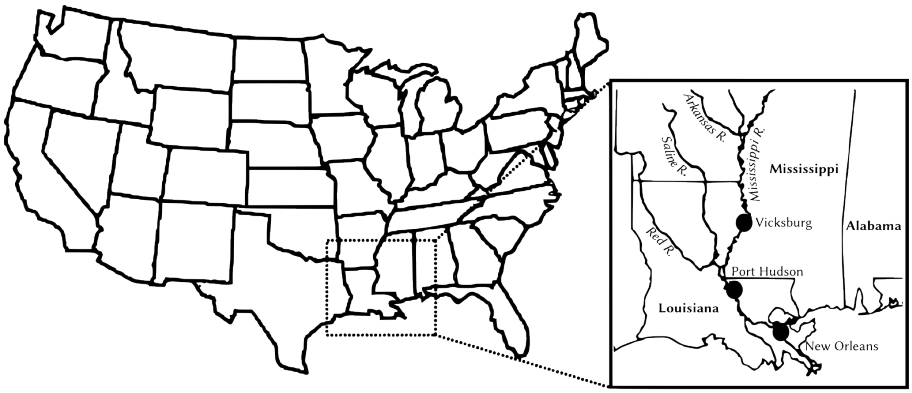


Figure 1: Lower Mississippi Valley

(Winford 2003: 7). The LMV was a major hub of trade and contact between many different ethnolinguistic groups, enabling contact among speakers of various languages.

2 Internal versus external language developments

While the bulk of this paper will focus on external, or contact-driven, change, I should mention certain internal developments which make the Southeastern branch unique within the Siouan language family. Among the shared phonological innovations of Southeastern Siouan are common Siouan **š* to Southeastern *č* (e.g., Biloxi *čǫki*, Ofo *ačǫki*, Tutelo *chǫ:ki* 'dog'[5]) and the merger of glottalized and non-glottalized stops (Rankin 2011). Shared lexical innovations include innovative terms for 'road' (Biloxi *natkhohi*, Ofo *nakhó•hi*, Tutelo *hątkóx*; 'prairie' (Biloxi *takohǫ*, Ofo *akhó•hi*, Tutelo *lata:hkoi*, *oni:i*); and 'squirrel' (Biloxi *ąsaki*, Ofo *tó•staki*, Tutelo *hista:xkai*); and fusion of the terms for 'grizzly' and 'black bear' (Biloxi *ǫti*, Ofo *ųthi*, Tutelo *hamǫ:thi*, *mǫ:ti*) (Rankin 2011). Shared morphosyntactic innovations include the auxiliation of *yukê* 'be (PL)' and 'durative aspect', collapse of the 'here/there,' or 'home base/apogee' (Cumberland 2005: 125), distinction in verbs of arrival, collapse of active/stative argument marking, and split negation (Cumberland 2005). These innovations are internal developments that likely occurred before the Biloxi and Ofo migrations into the LMV and the contact-related developments that happened after that.

[5] Biloxi terms are based on Dorsey & Swanton (1912), Ofo terms on Rankin's reanalysis (2002) of Dorsey & Swanton (1912), and Tutelo terms on Oliverio (1996).

External, as opposed to internal, language developments arise through languages coming into contact with each other, usually over an extended period of time. The depth of contact between two or more languages can generally tell us how long those groups were in contact. Lexical and phonetic features, which are easily recognizable surface features in languages, can be borrowed between groups with minimal contact and are thus weighted lower in determining the overall strength of a sprachbund (Kaufman 2014). Morphological features, which are more deeply embedded in the grammatical structure of languages, are more difficult to borrow and require more intimate contact to develop. Thus, morphological features are weighted more highly (Kaufman 2014).

For this paper, I address only those features I weighted more highly in Kaufman (2014) – those given a score of 2 (the features most indicative of an LMV sprachbund), and only if they occur in the LMV Siouan languages.[6] Phonetic and phonological features discussed are: (1) nasalized vowels; (2) voiceless labiodental fricative /f/; (3) alternation of /i/ and /u/; and (4) alternation of word initial /h/ ~ /Ø/. Morphological features discussed are: (1) focus and topic (discourse) marking, (2) valence-reducing prefix, (3) positional verb auxiliaries and (4) verb number suppletion.

I will then discuss lexical items that appear to have been shared among LMV languages, particularly those involving Biloxi and Ofo. Although lexical features were scored differently from phonetic/phonological and morphosyntactic features (see Kaufman 2014) and are weighted less overall, it has been long noted that certain lexical items appear broadly diffused in the region.

3 Phonetic and phonological features

3.1 Nasalized vowels

Nasalized vowels are a feature of Siouan and Muskogean languages. All Siouan languages, with the exception of Hidatsa and Crow, have vowel nasalization, including Biloxi and Ofo. Nasal vowels also occur in the LMV languages Atakapa, Choctaw-Chickasaw, MTL, and Natchez. In Natchez, however, nasal vowels oc-

[6] In Kaufman (2014), I weighted features on a tripartite scale of 0, 1, and 2. A score of 0 indicates that the feature in question does not exist in the area I delimited as the LMV. A score of 1 indicates that the feature exists in the area but is so common crosslinguistically that its presence in the LMV is not distinctive and thus not deemed relevant to supporting the LMV as a sprachbund. A score of 2, the highest weighting, indicates that the feature is either geographically limited to the LMV and its immediate periphery, or is so unusual crosslinguistically as to be especially relevant in supporting the LMV as a sprachbund (Kaufman 2014).

cur only in phrase- or sentence-final position and are thought to be based on underlying final /n/, which acts as a type of declarative marker (Geoffrey Kimball 2013, p.c.). Vowel nasalization in Atakapa is at times uncertain, perhaps being an allophone of the phoneme /ŋ/. Vowel nasalization in Atakapa and Natchez may be due to contact with LMV Siouan and Muskogean languages, although such nasalization may also be due to internal impetus.

3.2 Voiceless labiodental fricative /f/

Only one Siouan language, Ofo, has this phoneme, although all Muskogean languages, including MTL, have it. Haas postulated Muskogean /f/ as the modern reflex of Proto-Muskogean /xw/ (1969: 36). Biloxi may have had at least a dialectal reflex of /xw/ pronounced as /f/, as evidenced by Mrs. Jackson's pronunciation of *nixuxwi* (*nišofeˀ*) 'ear' (Haas & Swadesh 1968: 79), a pronunciation that correlates with the probable change of Proto-Muskogean /xw/ to /f/. (It is unclear whether this was a dialectal feature of Biloxi at the time data were elicited or whether this was an idiosyncratic pronunciation based on possible personal influence of Choctaw-Chickasaw.) This phoneme is also found in Atakapa, though rare and usually in word-final position, and may be due to internal impetus such as through fricativization of word-final labiodental velar /w/.

3.3 Alternation of /i/ and /u/

The alternation of /i/ and /u/ occurs in Biloxi, Natchez, and Tunica. This alternation appears to be a feature of Siouan languages, particularly of Biloxi but also of Dhegiha Siouan languages. The transition of /u/ to /i/ in Siouan is most apparent in Kanza (Kaw), wherein /u/ is pronounced like German *ü* (/y/), apparently midway in transition between /u/ and /i/. (Dorsey & Swanton 1912 also occasionally note the phoneme /y/ in Biloxi pronunciation, though it was apparently infrequent.) Examples include Biloxi *ci* and *cu* 'put, place, plant'; Natchez *išuš* and *ušuš* 'back'; and Tunica *tahišini ~ tahišuni* 'sieve'; *hiši ~ hišu* 'sift'.

This feature is crosslinguistically rare and is not likely a genetic or internally developed feature. It is likely that this feature's occurrence in Natchez and Tunica arose through contact with Siouan languages, although it could also be the result of vowel harmony.

3.4 Alternation of word initial /h/ ~ Ø

The alternation of word initial /h/ ~ Ø (zero marking) is a feature of the LMV area that occurs in Biloxi as well as in Atakapa and MTL. Examples include Atakapa *hipa ~ ipa* 'husband' (Gatschet & Swanton 1932: 42), *hikat ~ ikat* 'foot' (Gatschet & Swanton 1932: 40), *himatol ~ imatol* 'four' (Gatschet & Swanton 1932: 41) and *huket ~ uket* 'mother' (Gatschet & Swanton 1932: 46); Biloxi *hane ~ ane* 'find', *hamihi ~ amihi* 'heat' and *hasne ~ asne* 'thief' (Dorsey & Swanton 1912: 3); and MTL *hat(t)ak ~ atak* 'man' (Crawford 1978: 88; Drechsel 1996: 295) and *hoyba ~ oyba* 'rain' (Drechsel 1996: 306). This feature appears to be a Siouan-language-internal development, since "glottal stop is often inserted before word-initial vowels in Siouan sentences as a *Grenzsignal* — a boundary marker — so it is possible that the Biloxi initial *h-* that comes and goes in these words is the local reflex of [ʔ]" (Rankin 2011: 3). Regarding MTL, the alternation appears "to be instances of an *h-* that was present etymologically in Western Muskogean that was lost among certain users of Mobilian" (Rankin 2011: 3). Since the change from [ʔ] to *h-* appears to be an internal Siouan development, it is possible that this feature was copied from Siouan (Biloxi) into Atakapa and MTL.

4 Morphological features

The ranking of morphological features is a bit trickier than for phonetic and phonological features, since data on morphological features for languages in and around the LMV are often lacking in specific features. For example, MTL totals very low on the morphological-features scale simply because the language, typical of pidgins, is largely isolating and contains few morphological features. Ofo also scores low, simply because extant data on the language is scanty, not because it did not participate more fully in the LMV language area.

Morphological features that have been determined most relevant in analyzing the LMV as a sprachbund (Kaufman 2014: 3) are:

1. Focus and topic marking.
2. Valence-reducing prefix.
3. Positional verb auxiliaries.
4. Verbal number suppletion.

These features have been determined most relevant in the analysis of an LMV sprachbund partly because of their limited overall distribution beyond the LMV

and their relative rarity among the world's languages. Such limited distribution indicates a comparatively confined area probably once having a high volume of ongoing contact.

4.1 Discourse marking

Pragmatic or discursive affixation such as focality and topicality marking is fairly common among Native American languages. I use the term DISCOURSE-MARKING to include speaker-centered emphatic marking, often labeled *focus*, *topic* and *assertion*, as well as evidentiality and reference tracking. These markers, in each language in which they occur, are discussed below.

4.1.1 Focus

I use the term FOCUS to refer to new information (what Prague School linguists call "rheme") (Payne 1997: 271). LMV focus-marking suffixes can occur on both nouns and verbs.

Biloxi, along with Atakapa, Chitimacha, Choctaw-Chickasaw, and Natchez, has focus-marking suffixation. Atakapa and Chitimacha appear to share a focus-marking suffix *-š* while Choctaw-Chickasaw and Natchez appear to share the suffix *-ook*. Unfortunately, focus and topic marking cannot be discerned in Ofo from extant data.

In Biloxi, the marker *-di* is often suffixed to nouns in texts, particularly with nouns newly introduced into the narrative or discourse (Kaufman 2011: 3). The suffix *-di* descends directly from Proto-Siouan **-ri*, a focus marker also found in Hidatsa and Mandan (Boyle 2007: 3, p.c.). This suffix is sometimes used at first mention when objects or characters are first introduced into a story, thus signaling new information.

(1) *Skakana-**di*** *ewite-xti* *eyąhi* *yuhi* *yohi-yą.*
Ancient.of.Opossums-FOC early-INTENS 3SG.arrive 3SG.think pond-TOP
'The Ancient of Opossums thought he would reach a certain pond very early in the morning.' (Dorsey & Swanton 1912: 26)

(2) *Ąyaa-**di*** *wax* *ni* *yukê.*
person-FOC hunt walk MOVE
'Some people were hunting.' (Dorsey & Swanton 1912: 65)

4.1.2 Topic

I use the term TOPIC to refer to old, previously mentioned, or known information (what Prague School linguists call "theme") (Payne 1997: 271). Biloxi and Choctaw-Chickasaw have suffixes that serve as types of definite article, indicating previous mention. Biloxi *-yą* is a form of definite article that tends to occur most frequently when the noun to which it is suffixed has already been introduced into a story, thus marking old or already given information, as the following examples show:

(3) *Ątatka-**yą** khu-ni ǫǫni e-tu xa.*
child-TOP 3.give-NEG PST 3.say-PL always
'Always she did not give him the child.' ('She never gave him the child'?)
(Dorsey & Swanton 1912: 43)

(4) *"Yamą na," e-di ąyaa-xohi-**yą**.*
no DECL.M 3sg.say-ASRT person-old-TOP
' "No," the old woman said.' (Dorsey & Swanton 1912: 67)

In the above examples, 'child' and 'old woman' were previously mentioned in the discourse.[7]

The Choctaw-Chickasaw suffix *-aaš* indicates previous mention, in essence acting as a type of definite article:

(5) *Hattak-Ø-**aaš**-at čaaha-h.*
man-COP-PREV-NOM tall-TNS
'The previously mentioned man is tall.' (Broadwell 2006: 89)

4.1.3 Assertive marking

Biloxi, along with Atakapa, Chitimacha, and Natchez, has assertive markers, with which a speaker may choose to add particular emphasis or immediacy to a verb.

We have seen the Biloxi focus marker *-di* attached to nouns, but the suffix *-di* also attaches to verbs. With verbs, *-di* shows more emphasis or immediacy and has been glossed as an "assertive" marker (Kaufman 2011: 3), as the following examples demonstrate:

[7] In example 1 above, *-yą* appears on *yohi* 'pond', though the pond is not previously mentioned in the text. However, since this certain pond is already known to the Ancient of Opossums, it seems to be treated as previous knowledge, or a previously known location that can take the definite article marker.

(6) *Sǫǫnitǫǫni-k ǫha ąyaa ǫǫni ustax kanê-**di**.*
tar-ACC with man make stand.up EVID-ASRT
'He made a tar baby [person] and stood it up there.' (Dorsey & Swanton 1912: 13)

(7) *Kąkǫǫni dǫhi tê dê-**di** ê-tu-xa.*
trap see want go-ASRT they-say-always
'They say that he departed, as he wished to see the trap.' (Dorsey & Swanton 1912: 184)

4.2 Valence-reducing prefix

All languages have operations that adjust the relationship of semantic roles and grammatical relations in languages, using a range of structures for accomplishing this (Payne 1997: 169). In the LMV, a preverb or prefix is used as a valence-reducing operation. Atakapa, Biloxi, Chitimacha, Choctaw-Chickasaw, Natchez, and Ofo all have valence-reducing prefixation.

Siouan languages have a prefix *wa-* (reduced to *a-* in Biloxi and Ofo[8]), whose actual translation is murky, though it often can be translated as 'thing' or 'something' (i.e., an indefinite object prefix) and acts as a type of valence reducer (Robert Rankin, p.c.):

(8) ***a**-duska*
thing-bite
'rat' (Dorsey & Swanton 1912: 186)

In Atakapa, the valence-reducing prefix is *šok-*:

(9) ***šok**-koi*
INDF.OBJ-speak
'chief' ('speaking things') (Gatschet & Swanton 1932: 9)

The Chitimacha valence-reducing preverb is *ni*:

(10) ***ni** katš hamtši:k*
thing fortune having
'having (good) luck' (Daniel Hieber, p.c.)

[8] Biloxi and Ofo normally lose word-initial labial resonants, or most reflexes of **w*, **m*, and **W* (Rankin 2002: 19).

The Choctaw valence-reducing prefix is *naa-* or *nąn-*:

(11) ***naa**-hóoyo-´*
INDF.OBJ(SUBJ)-hunt-NZR
'hunter' or 'prey' (Broadwell 2006: 53)

Example 11 demonstrates that Choctaw *nan-* or *naa-* can be ambivalent, since the preverb *naa-* can represent either the actor (hunter) or the patient (prey) (Broadwell 2006: 53). The Western Muskogean prefixes *nąn-* and *naa-* likely derive from the word *nąta* 'what, something, someone.'

The Natchez valence-reducing prefix is *kin-*:

(12) Nokkinhantawąą.
*nok-**kin**-han-ta-w-aa-n*
PVB-INDF.OBJ-make-1SG-AUX-INC-PHR.TRM
'I can work.' (Kimball 2005: 405)

4.3 Positional verb auxiliaries

Classificatory verbs of the LMV signal position classification of noun referents: SIT, STAND, LIE, and MOVE, which occur as markers of continuative aspect in most if not all of the Siouan languages (Rankin 2004: 203). Positional verbs have been grammaticized in the Siouan languages as continuative aspect markers and proximal demonstrative determiners (Mithun 1999: 116). Biloxi and Ofo, along with Atakapa, Chitimacha, Choctaw, and Tunica, all use positionals in a similar manner, indicating possible borrowing between them.

(13) *Nihọ ani dêxtowê **nê**.* (Biloxi)
cup water full STAND
'The cup is full of water.' (Dorsey & Swanton 1912: 166)

(14) *B-ashě **nąki**.* (Ofo)
1-sit SIT
'I am sitting down.' (Rankin 2002: 20)

Positional verbs are also used for continuative aspect in other LMV languages, as these examples show:

(15) *Keu kam-š-kin-**tu**.* (Atakapa)
sit protrusion-DEF-LOC-STAND
'I am [seated] paddling.' (Gatschet & Swanton 1932: 61; Watkins 1976: 27)[9]

(16) wekt kas tuhjyi:kʔ peʔanki (Chitimacha)
we-t-k kas tuhjte-:ikʔ ***pe****-ʔe-nk-i*
DEM-REFL-LOC back stoop.down-PRTP be(horizontal)-3SG-LOC-NZR
'when he had stooped down' (Swadesh, unpublished notes)

(17) *Bill-at ma* ***binɨli****.* (Choctaw-Chickasaw)
SUBJ there sit.ANIM
'Bill is (sitting) over there.' (Watkins 1976: 21)

(18) ya· potkop kaʔašup kaʔepe·nakiyaku·š (Natchez)
*ya· potkop kaʔašup-Ø ka·-****ʔepe****·-na-ki-ya-ku·š*
that mountain blue-ABS PVB-lie-3PL-AUX-ART-ALL
'(where) that blue mountain is (lying)' (Kimball 2005: 438)

(19) *T-uruna-tʔe-ku* ***ʔuna****.* (Tunica)
DEF-frog-large-M.SG sit
'There is the (sitting) bullfrog.' (Watkins 1976: 26)

In many languages of the world the same lexical item can express both actual physical stance and can be used as an auxiliary, as is demonstrated in the Chitimacha, Choctaw-Chickasaw, Natchez, and Tunica examples above. In Biloxi and Ofo, however, physical stance and locative-existential predicates/verbal auxiliaries generally form two different sets of lexemes. The stance verbs used as independent verbs in Biloxi are *toho* 'lie', *xêhê* 'sit', *sɨhɨ* 'stand', and *hine* and *ni* 'move'. In Ofo the independent verbs are *čáftu* 'lie', *áshě* 'sit', and *askho(le)* 'stand' (there is no data for 'move' in Ofo). Their grammaticized auxiliary counterparts are *mąki* 'lie' and *nąki* 'sit' in both Biloxi and Ofo, while *nê* 'stand' and *ąde* and *hine* 'move' occur in Biloxi but are unattested in Ofo. The Biloxi form *hine* is used for both singular and plural while *ąde* has a suppletive plural form, *yukê*. *Ąde* is used for general movement and running while *hine* is for walking only (Kaufman 2013: 3).

These verbs form a discrete set of auxiliary verbs that often no longer specify actual physical position or movement but, rather, are used to express nuanced aspectual meanings. Biloxi *mąki*, *nąki*, and *nê* are used for both animates and inanimates, while *ąde* and *hine* are confined to use only with animates. *Mąki*, *nąki*, and *nê* share a common plural form *(h)amąki*, apparently a form of *mąki* 'lie'.

[9] Watkins (1976) identified *kamškintu* only as 'paddle.' I have analyzed it into its component parts.

4.4 Verbal number suppletion

For this section, the definition of suppletion includes cases that satisfy either of the following criteria: (1) exceptions to very productive derivational patterns, and (2) exceptions to established agreement patterns (Veselinova 2003: 3). The verbal suppletion treated here relates to nominal arguments of the verb, where the verb agrees with its arguments. All languages of the LMV, except MTL and Natchez, have verbal number suppletion in relation to nominal arguments. This feature is further limited in the region by being primarily used in relation to the positional auxiliaries STAND, SIT, LIE, MOVE (see above). In Tunica, only these auxiliary verbs show suppletion, while other verbs in the language do not (Haas 1946: 40). While not displaying direct borrowing of the suppletive terms between the languages, the fact that verbal number suppletion occurs primarily or only in positional auxiliaries makes this a distinguishing feature of the LMV. While the suppletive verb forms may be unique to each language, the underlying pattern of such deviating forms across LMV positional auxiliaries would seem to indicate a deeper-level pattern influence among multilingual speakers of this sprachbund.

Verbal number suppletion in each language is shown below:

Table 1: Biloxi (Dorsey & Swanton 1912: 3)

	singular	plural
STAND	*nê*	
SIT	*nąki*	*(h)amąki*
LIE	*mąki*	
MOVING	*ąde*	*yukê*

Table 2: Atakapa (Gatschet & Swanton 1932: 3)

	singular	plural
STAND	*to/tu* or *ta*	*tsot*
SIT	*ke*	*nul*
LIE	*tixt*	*yoxt*

[9]

Chitimacha, like Biloxi, neutralizes the singular auxiliary forms to a single plural form, *na(h)*.

Table 3: Chitimacha (Swadesh 1939: 32)

	singular	plural
STAND	*ci(h)*	
SIT	*hi(h)*	*na(h)*
LIE	*pe(h)*	

Choctaw-Chickasaw has both animate and inanimate forms for SIT.

Table 4: Choctaw (Broadwell 2006: 3)

	singular	dual	plural
STAND	*hikiya*	*hiili*	*(hi)yoh-*
SIT (ANIM.)	*binili*	*chiiya*	*binoh-*
SIT (INANIM.)	*talaya*	*taloha*	*taloh-*
LIE	*ittola*	*kaha*	*kah-*

In Tunica, suppletion is "a process not used by any other word-class of the language" (Haas 1946: 40). Thus, Tunica suppletion appears to be a borrowed feature from contact with other LMV languages.

Table 5: Tunica (Haas 1946: 40)

	singular	dual	plural
STAND	*kali*	*?*	*?*
SIT	*ʔuna*	*ʔunana*	*ʔukʔɛra*
LIE	*ʔura*	*ʔurana*	*naʔara*

It should be noted that Dhegiha Siouan languages, such as Kanza (Kaw), also show some suppletion in positional verbs (e.g. Kaw *yį́khé* 'sitting animate/inanimate singular object' and *yąkhá* 'sitting animate plural object)'. Whether this is due to contact between Dhegiha Siouan and LMV languages is debatable and remains

a possibility to be further studied. The Dhegiha Siouan language Quapaw, for example, was on the LMV periphery.

Unfortunately, in Ofo, only the positional forms *mąki* and *nąki* are attested, so determination of verbal number suppletion is not possible.

5 Lexical features

Lexical borrowing, due to the easy surface-level recognition of lexical items, is considered less important for establishing a sprachbund. Word borrowings operate according to a certain set of probabilities. Languages are more likely to borrow nouns than verbs (Tadmor, Haspelmath & Taylor 2010: 231). Adjectives and adverbs are almost as hard to borrow as verbs, and words with grammatical meanings (function words) are harder to borrow than verbs (Tadmor, Haspelmath & Taylor 2010: 231). Basic vocabulary is borrowed before structure and is indicative of more intense contact, while non-basic vocabulary is easiest to borrow (Thomason 2001: 69) and gets borrowed under conditions of even casual contact (Tadmor, Haspelmath & Taylor 2010: 231). Intensity of contact is, however, "a vague concept, and it cannot be made much more precise because it interacts with speakers' attitudes as well as with more easily specified factors, such as the level of fluency of the borrowers and the proportion of borrowing-language speakers who are fully bilingual in the source language" (Tadmor, Haspelmath & Taylor 2010: 231).

5.1 Basic vocabulary

The concept of basic vocabulary is important to the analysis of lexical borrowings. Several lists have been created to reflect basic concepts that are considered to be universal and culturally independent, such as basic kinship (e.g., *mother*, *father*), general animal terms (e.g., *fish*, *bird*), and basic verbs (e.g., *make*, *go*). The stability of the resulting list of "universal" vocabulary has been brought into question, however, and multiple lists of basic vocabulary have been published. The first was the Swadesh list of 100 basic words.

The Swadesh list was assembled by the linguist Morris Swadesh (Swadesh 1971). Swadesh "determined a priori what constituted basic vocabulary based on his intuitions, and then proceeded to refine his list by trial and error" (Tadmor, Haspelmath & Taylor 2010: 230). A newer list, which I used in analyzing LMV lexical items, is the Leipzig-Jakarta (L-J) 100-word list (Haspelmath & Tadmor 2009). This list (see Table 6) is based on systematic empirical data from

40 different languages. An advantage of the Leipzig-Jakarta list is that it "has a strong empirical foundation and is thus a more reliable tool for scientific purposes" (Tadmor, Haspelmath & Taylor 2010: 230). However, as with acceptance of any word list, things are not always perfect and certain questions remain unaddressed, such as why *black* is considered a basic color but not *white*.

Table 6: Leipzig-Jakarta list of 100 basic words Haspelmath & Tadmor (2009)

ant	eye	leg/foot	small
arm/hand	to fall	liver	smoke
ash	far	long	soil
back	fire	louse	to stand
big	fish	mouth	star
bird	flesh/meat	name	stone/rock
to bite	fly	navel	to suck
bitter	to give	neck	sweet
black	to go	new	tail
blood	good	night	to take
to blow	hair	nose	thick
bone	hard	not	thigh
breast	he/she/it/him	old	this
to burn (intrans)	to hear	one	to tie
to carry	heavy	rain	tongue
child (recip of parent)	to hide	red	tooth
to come	to hit/to beat	root	water
to crush/to grind	horn	rope	what?
to cry/to weep	house	to run	who?
to do/to make	I/me	salt	wide
dog	in	sand	wind
drink	knee	to say	wing
ear	to know	to see	wood
to eat	to laugh	shade/shadow	yesterday
egg	leaf	skin/hide	you (sg)

5.2 Semantic classes of borrowings

The number of borrowings between LMV languages can tell us something about the prior location and migration patterns of LMV groups. For example, the sheer volume of borrowings between Atakapa and Biloxi suggests that these languages were heavily in contact at one time. This seems extraordinary given the post-contact geographic locations of these groups, being on opposite sides of the Mississippi River. It is also notable that there are fewer borrowings between Chitimacha and Biloxi than between Atakapa and Biloxi, even though the Chitimachas, at least given their post-contact location, were in between. This could indicate, however, that Atakapas and Biloxis were geographically much closer to each other at one time. Biloxis may once have been located west of the Mississippi River before migrating eastward to the Pascagoula River region along the Gulf of Mexico where they encountered the French in 1699.

Table 7 is a list of LMV borrowings by semantic category (L-J basic vocabulary in **bold**):

Table 7: LMV borrowings by semantic category

Agricultural	(2)	seed, turn (soil?)
Body parts	(9)	anus/back, **arm/hand**, belly, **breast**, elbow, face, **knee**, **mouth**, **tooth**
Botanical	(9)	berry, cedar, corn, cotton, cypress, oak, peach, pepper, pumpkin/turnip
Color	(2)	black, **white**
Drink	(1)	**water**
Food	(2)	tortilla, bread
Kin	(1)	brother
Transport	(1)	canoe
Weapon	(1)	bow
Zoological	(19)	bee, **bird**, bison/buffalo, blackbird, bullfrog, buzzard, cow/calf, crane, deer, **dog**, duck, **fish**, flying squirrel, raccoon, robin, skunk, snake, wildcat, woodpecker

Several basic words appear to have been shared between Biloxi, Ofo, and other LMV languages; see Table 8.

Table 8: Shared basic vocabulary

	Atakapa	Biloxi	Chitimacha	Natchez	Ofo	Tunica
hear	*nak*	*naxe*				
laugh	*hayu*	*xahaye*				
blow	*pun*	*po*	*puuh(te)*	*puuh-hoo'iš*		
cord/ rope		*įką*				*yúnka*
cry		*wahe*				*wáha*
knee	*timak*	*cinąki*				*cina(hki)*
mouth		*ihi*	*i 'tooth'*	*ihi*	*ihi*	
wind		*xux(we)*	*howi*			*húri*

5.3 Widespread lexical borrowings in the LMV and Southeast

Certain nouns, and at least three verbs, are fairly widespread throughout the LMV and Southeast in their diffusion: those for 'bison/buffalo', 'bullfrog', 'cut', 'deer', 'goose', 'metal', 'robin', 'split', 'town', 'turn', 'water', and 'woodpecker'.

(20) 'Bison/buffalo': Similar terms for 'bison/buffalo' are of particularly widespread diffusion, ranging from Caddoan in the western Plains to Catawba near the eastern seaboard, including in the LMV: Bi. *yinisa*, *yanasa*, Choc.-Chic. *yanaš*, MTL. *yanaš*, Nat. *yanašah*, and Tun. *yanši*, *yanškaši*. The Ofo term *naf* 'cow' is likely also derived from this widespread 'bison' term. While the source of the borrowing is unknown, Taylor (1976: 166) suggested the possibility of its origin in an Athabaskan language. I concur with him that the Apache *iyaná ɬa'* (with loss of the initial *i* and the second element being the enclitic for indefinite determiner) could indeed be the source of copying. Apaches were a Plains group who may have been in contact on a regular basis with buffalo hunting parties of other groups from the LMV and Southeast and were probably also involved in the buffalo fur trade. Totonac has the word *tiyaná* for 'ox,' raising the possibility of borrowing between this Mexican Gulf coastal language and the LMV for this similar bovine perhaps through Mobile Bay.

(21) 'Bullfrog': Similar terms for 'bullfrog' occur in At. *anenui*, Bi. *kǫninuhi*, MTL. *hanono*, Nat. *hánanai*, and Tun. *uruna(te)*. The source language of this borrowing is unknown.

(22) 'Cut': Similar terms for 'cut' occur in At. *kets* or *kuts*, Bi. *kutsi*, Nat. *keš*, and Tun. *kušu*. The Plains languages Comanche, Tonkawa, and possibly Caddoan have terms similar to the LMV form. The source language of the borrowing is unknown.

(23) 'Deer': Similar terms for 'deer' appear to have been borrowed in the LMV as well as in the Plains periphery. The Proto-Siouan form is **wi-htáa*, indicating possible borrowing from Siouan (possibly Biloxi *[i]tha*) into Natchez *ša*. Similar terms also appear in Pawnee (Caddoan) and Kiowa (Kiowa-Tanoan), possibly borrowed from Biloxi.

(24) 'Metal': Similar words beginning with *nasal + /a/ + fricative / lateral* occur in Bi. *maasa* or *maasi*, Choc.-Chic. *maała* 'kettle' and Nat. *naLkw*. Intriguingly, forms of this word also occur on the other side of the Gulf in Mayan (e.g., Proto-Yukatekan **mahskab'* 'metal,' Yukatek *maHskab'* 'machete' and Mopan *ma?aska?* 'metal' (Kaufman & Justeson 2003: 208).

(25) 'Robin': Similar terms for 'robin' occur in the LMV (e.g. Bi. *sįkuki*, Choc.-Chic. *biškoko*, MTL *beškoko*, Nat. *miškokw*, and Tun. *wišk?ohku*. The term also extends into Eastern Muskogean (e.g. Ala. *čiskokko*).

(26) 'Split': Similar terms for 'split' occur in At. *čal*, Bi. *ča*, Chit. *čap*, Choc.-Chic. *ču?alli*, MTL. *čolale*, and Tun. *čal*. It may be significant that the semantically similar verb 'cut' also has a fairly widespread distribution in the LMV.

(27) 'Town': Similar terms for 'town' occur in Western Muskogean – but not in Eastern Muskogean – and are widespread across Siouan languages. It is possible that the term was borrowed between the two families, though the direction of borrowing is uncertain. It is possible that the term was borrowed into Siouan from Algonquian, since the Lakota word for town *otȟúŋwahe* is strikingly similar to, for example, Ojibwe (Algonquian) *oodena* (Nichols & Nyholm 1995: 272). Even if the Siouan term was borrowed from Algonquian, the Choctaw-Chickasaw term may have its source in a Mexican Gulf coastal language: Totonac. The Totonacan term *tamawan* (*tamāhuan*) means 's/he buys' while *liitamaw* (*litamáu*) and *puutamawan*

(*putamahuán*) means 'plaza' or 'place to buy' (Aschmann 1973: 110). (The Totonac prefix *lii-* is an instrumental prefix while *puu-* is a locative prefix (MacKay 1999: 386,388).) Assuming that there may have been circum-Gulf navigation and trade, it is possible that this term entered Choctaw-Chickasaw and MTL as *tamaha* from Totonacan *tamawan* as a means of referring to a center for buying, selling, and trading (i.e., a plaza or town center).

(28) 'Turn': Similar terms for 'turn' occur in At. *miš*, Bi. *mixi*, Chit. *tamix*, and Tun. *maxsi*.

(29) 'Woodpecker': Similar terms for 'woodpecker' occur in Bi. *pakpakhayi*, Choc.-Chic. *bakbak*, Nat. *pukpúku* and Tun. *páhpahkana*, and extend into Eastern Muskogean.

Certain of the above terms (e.g., 'goose', 'woodpecker') may be due to onomatopoeia, or words mimicking the sounds of nature. Yet "some resemblances are remarkably precise even if one allows for onomatopoeia" (Haas 1969: 82), as in the above examples. It might also be noted that certain widespread terms may be cultural in nature. For example, the Redheaded Woodpecker has a particular association with the ball game in Chickasaw (Galvan 2011: 34–37); the cultural iconicity of this bird associated with this sport and its nomenclature could easily have been copied by other groups through the ritual of intergroup ball play. The significance of the diffusion of certain terms such as 'cut', 'split', and 'turn' is unknown, although 'cut' and 'split' may be related to such activities as communal hunting and feasting and the sharing of meat. 'Turn' may be related either to the turning of soil involved in agriculture or perhaps to communal dancing, though this currently can only be speculation on my part.

Calques are loan translations (word-for-word semantic translations) shared among languages. Rather than an individual term being copied, as in lexical borrowing, calques involve the copying of a semantic phrase, the concept behind the phrase being copied rather than just the individual words.

Table 9 lists calques that are found among LMV languages (some of which are found beyond the LMV in peripheral languages).

Some of the most widespread calques – 'butter', 'donkey', 'jail', 'sugar' – were likely diffused through the MTL pidgin, which also contains the calques. Since extant data is limited for MTL, it is now impossible to know if other borrowings and calques were diffused through this medium, though it seems likely.

Table 9: Calques

idiomatic gloss	calque gloss	languages sharing calque
‘bedbug’	‘flat bug’	Biloxi, Caddoan
‘butter’	‘cow / milk grease’	Atakapa, Biloxi, MTL, Natchez
‘cologne’	‘smell good water’	Biloxi, Natchez
‘corn crib’	‘corn house’	Atakapa, Biloxi, Natchez, Tunica
‘donkey / mule’	‘long ear’	Atakapa, Biloxi, Caddoan, Choctaw, MTL, Natchez
‘jail’	‘strong house’	Atakapa, Biloxi, Choctaw, Creek, MTL
‘nostril’	‘nose hole’	Atakapa, Biloxi, Caddoan, Comanche, Kiowa, Natchez, Nahuatl
‘ocean’	‘big water’	Biloxi, Comanche, Nahuatl, Natchez
‘rattlesnake’ 1	‘big snake’	Biloxi, Tonkawa, Tunica
‘rattlesnake’ 2	‘chief / king snake’	Biloxi, Natchez, Tunica, Yukatek (Mayan)
‘stable [horse]’	‘horse house’	Atakapa, Biloxi, Comanche, Nahuatl
‘sugar’	‘sweet salt’	Atakapa, Biloxi, Choctaw, MTL, Natchez
‘thumb’	‘big / old hand’	Atakapa, Biloxi, Comanche, Natchez, Tunica
‘vein’	‘blood house’	Atakapa, Biloxi

6 Summary and conclusion

In my dissertation (Kaufman 2014: 3), I concluded that the LMV was a sprachbund on par with other well-known language areas, such as the Balkans of Eastern Europe, South Asia (India), and the Amazon basin. The strength of the LMV as a language area lies in the phonetic, phonological, morphological, and lexical features delineated above. Two Siouan languages – Biloxi and Ofo, members of the Southeastern, or Ohio Valley, branch of the Siouan language family – participated in the LMV sprachbund after their migrations into the region. In this paper, we have seen that several features typical of the LMV language area, and largely absent from other Siouan languages, are present in Biloxi and Ofo. Data on the latter language are admittedly sparse, leaving many aspects of the language inconclusive, though Ofo still seems to have participated to a great degree in the LMV sprachbund.

I have discussed the following LMV phonetic, phonological and morphological features, which received the highest weighting in Kaufman (2014): nasalized vowels, voiceless labiodental fricative /f/, alternation of /i/ and /u/, alternation of word initial /h/ and Ø, focus and topic marking, valence-reducing prefixes, positional verb auxiliaries, and suppletive verbal number agreement. We have also seen that several lexical items appear to have been shared in the LMV, including among Biloxi and Ofo. While the direction of borrowing is often unclear, it appears that borrowing involving Biloxi and Ofo went in both directions.

Dhegiha Siouan languages may have participated to some degree in, and been influenced by, the LMV sprachbund as well, especially in the area of positional verbal auxiliaries and verbal auxiliary suppletion. The extent of Dhegiha Siouan participation in the LMV sprachbund remains to be further studied.

Language contact has been less studied than "genetic," or family tree, linguistics, especially in regards to Native North American languages. The LMV is another of several Sprachbünde that have arisen around the world in response to the mingling of two or more languages and cultures. As we have seen, Biloxi and Ofo, though genetically Siouan, have been moderately influenced by contact with other LMV languages. These two Siouan languages were essentially subsumed into a broader cultural area that was centuries, if not millennia, in the making.

Abbreviations

1, 2, 3	first, second, third person	M	masculine
		MTL	Mobilian Trade Language (Mobilian Jargon)
ABS	absolutive		
ALL	allative		
ANIM	animate	Nat.	Natchez
ART	article	NEG	negative
ASRT	assertive	NOM	nominal
At.	Atakapa	NZR	nominalizer
AUX	auxiliary	OBJ	object
Bi.	Biloxi	PHR.TRM	phrase terminal
Chit.	Chitimacha	PL	plural
Choc.-Chic.	Choctaw-Chickasaw	PREV	previous (mention)
		PVB	preverb
DECL	declarative	RECIP	reciprocal
DEM	demonstrative	REFL	reflexive
EVID	evidential	SG	singular
FOC	focus	SUBJ	subject
INC	incompletive	TNS	tense
INDF	indefinite	TOP	topic
INTENS	intensifier	Tu.	Tunica
INSTRANS	intransitive		

References

Aschmann, Herman Pedro. 1973. *Vocabulario Totonaco de la Sierra*. México, D. F.: Instituto Lingüístico de Verano.

Boyle, John P. 2007. *Hidatsa morphosyntax and clause structure*. Chicago: University of Chicago (Doctoral dissertation).

Broadwell, George Aaron. 2006. *A Choctaw reference grammar*. Lincoln: University of Nebraska Press.

Crawford, James M. 1978. *The Mobilian trade language*. Knoxville: University of Tennessee.

Cumberland, Linda A. 2005. *A grammar of Assiniboine: A Siouan language of the northern Plains*. Bloomington: Indiana University (Doctoral dissertation).

Dorsey, James Owen & John Reed Swanton. 1912. *A dictionary of the Biloxi and Ofo languages, accompanied with 31 Biloxi texts and numerous Biloxi phrases*

(Bureau of American Ethnology Bulletin 47). Washington: Government Printing Office.
Drechsel, Emanuel. 1996. An integrated vocabulary of Mobilian Jargon, a Native American pidgin of the Mississippi Valley. *Anthropological Linguistics* 38(2). 248–354.
Galvan, Glenda. 2011. *Shared spirit* (Chickasaw Stories 1). Sulphur, OK: Chickasaw Press.
Gatschet, Albert Samuel & John Reed Swanton. 1932. *A dictionary of the Atakapa language, accompanied by text material* (Bureau of American Ethnology Bulletin 108). Washington: Government Printing Office.
Haas, Mary. 1946. A grammatical sketch of Tunica. In Harry Hoijer, Leonard Bloomfield & Mary Haas (eds.), *Linguistic structures of Native America*, 337–366. New York: Johnson Reprint Corporation.
Haas, Mary. 1969. *The prehistory of languages*. Berlin: de Gruyter Mouton.
Haas, Mary & Morris Swadesh. 1968. The last words of Biloxi. *International Journal of American Linguistics* 34. 77–84.
Haspelmath, Martin & Uri Tadmor (eds.). 2009. *Loanwords in the world's languages: A comparative handbook*. Berlin & New York: Mouton de Gruyter.
Kaufman, David V. 2011. Biloxi realis and irrealis particles. *Kansas Working Papers in Linguistics* 32. 1–7.
Kaufman, David V. 2013. Positional auxiliaries in Biloxi. *International Journal of American Linguistics* 79(2). 283–299.
Kaufman, David V. 2014. *The lower Mississippi Valley as a language area*. Lawrence: University of Kansas (Doctoral dissertation).
Kaufman, Terrence & John Justeson. 2003. *A preliminary Mayan etymological dictionary*. Los Angeles: Foundation for the Advancement of Mesoamerican Studies. http://www.famsi.org/reports/01051/index.html.
Kimball, Geoffrey. 2005. Natchez. In Heather Hardy & Janine Scancarelli (eds.), *Native languages of the southeastern United States*, 385–453. Lincoln: University of Nebraska Press.
MacKay, Carolyn Joyce. 1999. *A grammar of Misantla Totanac*. Salt Lake City: University of Utah Press.
Meyer, William. 1925. *Indian trails of the southeast* (Annual Report of the Bureau of American Ethnology to the Secretary of the Smithsonian Institution 42). Washington: Government Printing Office.
Mithun, Marianne. 1999. *The languages of Native North America* (Cambridge Language Surveys). Cambridge: Cambridge University Press.

Nichols, John D. & Earl Nyholm. 1995. *A concise dictionary of Minnesota Ojibwe*. Minneapolis: University of Minnesota Press.
Oliverio, Giulia. 1996. *A grammar and dictionary of Tutelo*. Lawrence: University of Kansas (Doctoral dissertation).
Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
Rankin, Robert L. 2002. *Quapaw dictionary*. Unpublished manuscript. Lawrence: University of Kansas.
Rankin, Robert L. 2004. The history and development of Siouan positionals. *Sprachtypologie und Universalienforschung* 57. 202–227.
Rankin, Robert L. 2011. *Who were the Tomahitans?* Paper presented at the Siouan and Caddoan Languages Conference, White Cloud, KS.
Swadesh, Morris. 1939. *Chitimacha grammar, texts, and vocabulary*. Manuscript 497.3.B63c G6.5, Franz Boas Collection of Materials for American Linguistics. Philadelphia: American Philosophical Society.
Swadesh, Morris. 1971. What is glottochronology? In Joel Sherzer (ed.), *The origin and diversification of language*, 271–284. Chicago: Aldine-Atherton.
Tadmor, Uri, Martin Haspelmath & Bradley Taylor. 2010. Borrowability and the notion of basic vocabulary. *Diachronica* 27(2). 226–246.
Taylor, Allan. 1976. Words for 'buffalo'. *International Journal of American Linguistics* 42. 165–166.
Thomason, Sarah G. 2001. *Language contact: An introduction*. Washington: Georgetown University Press.
Veselinova, Ljuba N. 2003. *Suppletion in verb paradigms: Bits and pieces of a puzzle*. Stockholm: Stockholms universitet (Doctoral dissertation).
Watkins, Laurel. 1976. Position in grammar: Sit, stand, lie. *Kansas Working Papers in Linguistics* 1. 16–41.
Winford, Donald. 2003. *An introduction to contact linguistics*. New York: Wiley-Blackwell.

Chapter 3

Regular sound shifts in the history of Siouan

Rory Larson

The contributors to the Comparative Siouan Dictionary (CSD) reconstructed a phonemic set for Proto-Siouan, together with the necessary reflexes to produce the actual speech sounds found in the various daughter languages. Until recently, this system was common knowledge within the Siouanist community, since the participants in the CSD project were active as the leaders of that community, and were available to explain the predicted sound shifts. With the passing, retirement, or disappearance of most of the CSD team, however, it seems that it might be useful to document the reconstructed system and its most important regular reflexes, as an aid to comparative studies. This paper will rely primarily on the CSD, edited until his passing by Dr. Robert Rankin, to summarize the regular sound shifts known to have occurred in Siouan. It will prioritize sound shifts in which separate phonemes or clusters have collapsed together to become indistinguishable in the daughter languages, since this is where interesting confusion is most likely to occur.

1 Introduction

In 1984, a group of linguists studying Siouan languages began a project under NEH and NSF sponsorship to assemble a comparative dictionary of the Siouan language family. The principal investigator was David S. Rood. The team included Richard T. Carter, A. Wesley Jones and Robert L. Rankin as senior editors, along with Rood and John E. Koontz. Together with Willem de Reuse, Randolph Graczyk, Patricia A. Shaw and Paul Voorhis, the dictionary team began their project at the Comparative Siouan Workshop held at the University of Colorado in 1984. A number of other scholars, including Louanna Furbee, Jimm Goodtracks, Jill Hopkins Greer, Kenneth Miner, Carolyn Quintero, Kathleen Shea and Mark Swetland also contributed information.

Rory Larson. 2016. Regular sound shifts in the history of Siouan. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 63–83. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.122 DOI:10.17169/langsci.b94.122

This undertaking was huge.[1] The present writer regrettably turned down an offer by Richard Carter in about 1997 of an advance copy of the dictionary manuscript, on the assurance that it would be published within a year. In fact, it was never completed to the editors' satisfaction or published in book form. Carter himself retired from active work in the Siouan field after around 2002, and Robert Rankin became the principal steward of the project. In 2006, Rankin distributed a .pdf file of the manuscript as it stood so far to interested members of the Siouanist community, on condition that any further requests be submitted to himself or David Rood. The file runs to nearly a thousand pages, and is full of working notes and comments about the various words and their relationships, mostly by Rankin.

Rankin, Carter, and their colleagues developed a sophisticated understanding of the phonological and phylogenetic relationships among the various groups of Siouan. Until recently, they formed a body of respected linguistic "elders" who freely shared this lore on request with more junior scholars. With the untimely passing of Robert Rankin in February of 2014, however, and the retirement, disappearance, or focus shift of most of the other leading members of the team, the framework they developed seems in danger of being forgotten by the Siouanist community. This paper is intended to address that concern. Drawing on working notes found throughout the CSD, as well as years of discussions on the Siouan List, it will attempt to summarize the model of Siouan phonology and its standard sound shifts built by Rankin and the CSD team, with occasional comments and additions from the writer. All references to the CSD are to the 2006 version. (In May 2015, after this paper was completed, the most recent version was made available online at csd.clld.org (Rankin et al. 2015) where most of the original notes and comments can now be found.)

2 The Siouan family tree

The CSD recognizes four major branches of Siouan. In the far northwest is Missouri Valley, or Crow-Hidatsa, consisting of the Crow and Hidatsa languages. Next is Mandan, an isolate within Siouan. Third is Mississippi Valley Siouan, or "MVS", which itself has three branches. Fourth is Southeastern, or Ohio Valley Siouan, at the southeastern end of the Siouan span.

[1] For an enlightening and wryly humorous history of the project, see "The Comparative Siouan Dictionary Project" (Rood & Koontz 2002) written by two of its principal participants.

MVS branches into Dakotan, which includes the five "Sioux" dialects of Santee-Sisseton (Dakota), Yankton-Yanktonai, Teton (Lakota), Assiniboine and Stoney;[2] Winnebago-Chiwere, composed of Ho-Chunk and the Iowa, Oto and Missouria languages; and Dhegiha, comprising Omaha-Ponca, Kaw, Osage and Quapaw. Southeastern Siouan contains Biloxi and Ofo as one branch, and Tutelo and Saponi as another.

Catawba is the language most closely related to Siouan. Though sound-historical relationships are not very clear, Catawba examples are often included in an entry's word list. The language probably next most closely related is Yuchi (cf. Kasak 2016 (in this volume)), but few examples of Yuchi are given.

3 The reconstructed Proto-Siouan phoneme set

The CSD team recognizes eight vowels for Proto-Siouan, five oral and three nasal, which were distinguished also by length (Rankin, Carter & Jones 1998).

(1) i u į ų
e o
a ą

Basic stops are *p, *t, *k, and the glottal stop. Proto-Siouan had a series of alveolar, palatal and velar fricatives: *s, *š, and *x, as well as *h. It also had three resonants, *w, *r and *y. Minimally, its consonant structure was as follows:

Table 1: Consonants

	Labial	Alveolar	Palatal	Velar	Glottal
Stops	p	t		k	ʔ
Fricatives		s	š	x	h
Resonants	w	r	y		

[2] Parks & DeMallie (1992) present the results of a major dialect survey undertaken to clarify the relations among the various Dakotan dialects, a field in which considerable confusion had prevailed in the literature prior to their work. The CSD itself is rather deficient in Dakotan material other than Lakota and Dakota, although it contains a few words from four other categories: Stoney, Assiniboine, Yankton, and "Sioux Valley". The latter presumably refers to the Sioux Valley reservation in southwestern Manitoba, which Parks & DeMallie classify along with the Minnesota "Dakota" within the broad Santee-Sisseton dialect group. The present paper will follow Parks & DeMallie in sub-classification of Dakotan.

Many of the consonants of Siouan have occurred in clusters, however, so the actual historical picture is more complex than this. Stops can be adjoined to other stops in almost any order, all non-glottal stops and fricatives can be glottalized, aspiration (h) can occur either before or after stops, and combinations can occur involving fricatives, stops and resonants. In particular, there exist two historical phonemes that manifest as either stops or resonants in the daughter languages, called "funny w" and "funny r". We symbolize these sounds as *W and *R. *R was found in Proto-Siouan, and *W in MVS only. Rankin believed that *R was originally a combination of *r with a laryngeal, either *h or the glottal stop.

Notably, Siouan had no distinct nasal consonant series. When *w or *r occurred in the environment of a nasal vowel, they usually manifested as |m| or |n|, respectively.

Accent in Proto-Siouan was normally on the second syllable of a word.

4 Historical Siouan sound shifts

One of the first sound shifts affecting Siouan was a process called "Carter's Law'. Wherever a simple stop, *p, *t or *k, occurred before the vowel of an accented syllable, the stop itself was more prominently "marked", either by lengthening it or by preaspirating it. In the CSD, these are considered to be preaspirated. Thus, *p, *t and *k become *hp, *ht and *hk before an accented syllable. Since accent normally was on the second syllable of a word, these preaspirated stops and their derivatives are usually found inside the word rather than at the beginning. When they are found at the beginning of a word, it may be an indication of a lost initial syllable.

(2) Carter's Law:

*pV́	>	*hpV́
*tV́	>	*htV́
*kV́	>	*hkV́

4.1 Missouri Valley (Crow-Hidatsa) reflexes

In Missouri Valley Siouan, loss of historical aspiration, loss of nasal vowels and the merger of *y with *r are the most sweeping transformations of the Proto-Siouan phonemic inventory. Several other changes also occur.

- As in Mandan, Proto-Siouan aspiration is lost.[3] This notably includes the preaspirate series produced by the operation of Carter's Law.

[3] Rankin et al. 2006: 50, 85.

Loss of aspiration:

*hp	>	*|p|	<	*p
*ht	>	*|t|	<	*t
*hk	>	*|k|	<	*k

- Phonemic nasalization is completely lost. The three Proto-Siouan nasal vowels merge with their oral counterparts, and neither vowels nor consonants are distinguished by nasality.[4]

  Loss of nasal vowels:

*ą	>	*|a|	<	*a
*į	>	*|i|	<	*i
*ų	>	*|u|	<	*u

- As in Mandan and Hoocąk, Proto-Siouan *y merges with *r.

  *y/*r merger: *y > *|r| < *r

- Between vowels at the end of a word, *h is lost .

  Loss of intervocalic *h: $*V_1hV_2 > *|V_1V_2|$

- Rightward vowel exchange, in which the first two vowels of a word are swapped.[5] Both Crow and Hidatsa show this feature, but not necessarily in the same words, which suggests that this change was spreading at the time Crow and Hidatsa separated.

  Rightward vowel exchange: $*CV_1CV_2 > |CV_2CV_1|$

4.1.1 Hidatsa reflexes

Few changes are specific to Hidatsa. There may be a few vowel shifts and cluster changes. Proto-Siouan *w generally manifests as |m|.

- Short *o is raised to |u|:[6] *o > |u|
- *xk becomes |hk|:[7] *xk > |hk|
- *w becomes |m|: *w > |m|

[4] Rankin et al. 2006: 109.
[5] Rankin et al. 2006: 193, 788.
[6] Rankin et al. 2006: 137, 922.
[7] Rankin et al. 2006: 193.

4.1.2 Crow reflexes

Crow is more innovative. The biggest change is complete loss of glottals, usually with lengthening of the following vowel. Proto-Siouan *x manifests as |xš|. Proto-Siouan *t becomes |s|.

- Glottalization is lost, but is reflected in the lengthening of the following vowel, usually with rising pitch.[8]

  Loss of glottals: *(C)ʔV > |(C)VV́|

- *x becomes |xš|:[9] *x > |xš|

- *t becomes |s|: *t > |s|

4.2 Mandan reflexes

In Mandan, loss of historical aspiration[10] and the merger of *y with *r are the most notable sound shifts, as well as a peculiar reversal of sibilants.[11]

- As in Crow and Hidatsa, historical aspiration is lost, including the preaspirate series.

  Loss of aspiration: *hp > |p| < *p
  *ht > |t| < *t
  *hk > |k| < *k

- As in Crow, Hidatsa, and Hoocąk, Proto-Siouan *y merges with *r.

  *y/*r merger: *y > |r| < *r

- Proto-Siouan *s and *š swap phonetic value. *s becomes |š| and *š becomes |s|.

  *s/*š reversal: *s > |š|
  *š > |s|

- The cluster *sp metathesizes to become |ps|. More generally, there seems to be a usual, but not quite complete, constraint against having |p| as the second element of a cluster.[12]

  *sp metathesis: *sp > |ps|

[8] Rankin et al. 2006: 232.
[9] Rankin et al. 2006: 124.
[10] Rankin et al. 2006: 50.
[11] Rankin et al. 2006: 126.
[12] Rankin et al. 2006: 275.

- Before a consonant, the absolutizing or generalizing *wa- prefix loses its vowel through syncopation, and the *w becomes |p|.[13]

  *wa- syncopation: *waC > |pC|

4.3 MVS reflexes

In Mississippi Valley Siouan (MVS), the fricatives are divided between a voiceless series and a voiced series. This is also the only branch of Siouan in which the preaspirates are clearly distinguishable. Another major transformation is the loss of short, unaccented vowels in the initial syllable,[14] and the production of clusters that result from this syncopation. This frequently involves the absolutizing or generalizing *wa- prefix, as well as the first person subject pronoun *wa^{1}- prefix. Also, the *hr cluster becomes *ht,[15] merging with the original preaspirate *ht. For this group, we may restate the basic consonant set as follows:

Table 2: Basic consonant set

		Labial	Alveolar	Palatal	Velar	Glottal
Stops:						
	Simple:	p	t		k	ʔ
	Preaspirate:	hp	ht		hk	
	Postaspirate:	ph	th		kh	
	Glottalized:	pʔ	tʔ		kʔ	
Fricatives:						
	Voiceless:		s	š	x	h
	Voiced:		z	ž	ɣ	
	Glottalized:		sʔ	šʔ	xʔ	
Resonants:						
	Normal:	w	r	y		
	“Funny”:	W	R			

- The Proto-Siouan fricatives are divided between a voiced and a voiceless set, possibly according to phonological conditions.

[13] Rankin et al. 2006: 793.
[14] Rankin et al. 2006: 10.
[15] Rankin et al. 2006: 199.

Voiced/voiceless fricative split: *s > *|s| and *|z|
*š > *|š| and *|ž|
*x > *|x| and *|ɣ|

- Proto-Siouan *pr merges with syncopated *w-r to become MVS *br.

  *pr/*w-r syncopation: *w-r > *|br| < *pr

- Syncopated Proto-Siouan *w-w usually becomes MVS *W.[16]

  *w-w syncopation: *w-w > *W

- Syncopated Proto-Siouan *wa^1- used as the first person affixed pronoun 'I', however, becomes MVS *m when it precedes *w or the glottal stop.[17]

  I-*wa^1-w syncopation: *wa^1-w > *m < *wa^1-ʔ

- Syncopated Proto-Siouan *w-C, where C is a voiceless contoid, becomes MVS *pC.[18]

  *w-C syncopation: *w-h > *ph
  *w-t > *pt

- Proto-Siouan *hr merges with preaspirate *ht to become *|ht|.[19]

  *hr/ht merger: *hr > *|ht| < *ht

4.3.1 Dakotan reflexes

In Dakotan, vowel length is lost. Proto-Siouan *y manifests as aspirated |čh|. So too do cases in which *r is preceded by *i. Many inalienably owned nouns beginning with |čh| in Dakotan are explained as *r-initial stems preceded by the *i- of inalienable possession. When *r stands alone without an adjacent consonant, it manifests as |y|. When *k is preceded by a front vowel, it palatalizes to |č| in Dakotan. Otherwise, the main sound shifts involve clusters. In particular, Proto-Siouan or MVS preaspirates become postaspirates, merging with that series.[20] The cluster *rh, which is important in a few words, becomes plain |h|.[21] In Dakotan, clusters of two stops are frequent, and the cluster *wR becomes *|br|,

[16] Rankin et al. 2006: 164, 193, 213
[17] Rankin et al. 2006: 10.
[18] Rankin et al. 2006: 793.
[19] Rankin et al. 2006: 199.
[20] Rankin et al. 2006: 199, 269, 818.
[21] Rankin et al. 2006: 165.

merging this with the MVS *br series. In all Dakotan languages, *br manifests as |mn| before a nasal vowel. Stoney and Assiniboine manifest *br and *kr the same regardless of environment, but these sounds alternate in the other three dialects according to whether the following vowel is oral or nasal.

- vowel length is lost: *VV > |V| < *V
- *y- and *ir- merge as |čh-|: *y- > |čh-| < *ir-
- *rV becomes |yV|: *rV > |yV|
- *k after front vowel becomes |č|: *ik > |ič|
  *ek > |eč|
- ates merge with postaspirates:
  *hp > |ph| < *ph
  *ht > |th| < *th
  *hk > |kh| < *kh
- *rh becomes |h|: *rh > |h| < *h
- *wR merges with MVS *br: *wR > *|br| < *br
- Dakotan *br then manifests as |mn| before a nasal vowel: *brṰ > *|mnṰ|

4.3.1.1 Santee-Sisseton reflexes

- *br alternates by nasality: *brV > |md| or |bd|[22]
  *brṰ > |mn|
- *kr alternates by nasality: *krV > |hd|
  *krṰ > |hn|
- *R manifests as |d|: *R > |d|

[22] In the CSD, "Sioux Valley" seems to agree with "Dakota" in practically everything except that the one case recorded of a Sioux Valley word with a *br reflex before an oral vowel shows this as |bd|, rather than the usual |md| for Dakota.

4.3.1.2 Yankton-Yanktonai reflexes

- *br alternates by nasality: *brV > |bd| or |md|
  *brV̨ > |mn|
- *kr alternates by nasality: *krV > |kd| or |gd|
  *krV̨ > |kn| or |gn|
- *R manifests as |d|: *R > |d|[23]

4.3.1.3 Teton (Lakota) reflexes

- *br alternates by nasality: *brV > |bl|
  *brV̨ > |mn|
- *kr alternates by nasality: *krV > |gl|
  *krV̨ > |gn|
- *R manifests as |l|: *R > |l|
- *tp becomes |kp|:[24] *tp > |kp| < *kp

4.3.1.4 Assiniboine reflexes

- *br always manifests as |mn|: *br > |mn|[25]
- *kr always manifests as |kn|: *kr > |kn|
- *R manifests as |n|: *R > |n|

[23] The CSD records very few words of Yankton, none of which are useful here. Parks & DeMallie (1992) stress that the long-repeated claim that Yankton-Yanktonai is "Nakota", is false; their self-designation, when not misled by confused linguists, is "Dakota", which means that *R manifests as |d|, not |n|, in their dialect. The true "Nakotas" are the Assiniboines and the Stoneys. The authors clearly illustrate the *kr clusters for this group on pages 245-6, but do not include any *br clusters. Words listed on the Yankton Reservation pedagogical website, http://www.nativeshop.org/learn-dakota.html, show that *br before an oral vowel normally manifests as |bd|, or perhaps sometimes as |md| or |mbd|.

[24] Rankin et al. 2006: 253, 265, 865.

[25] The CSD contains only a few words of Assiniboine, or Nakota. Parks & DeMallie (1992) demonstrate that *R becomes |n|, and that *kr manifests as |kn| before both oral and nasal vowels. The preliminary Assiniboine text developed by Shields (2012) contains words with *br clusters showing that these manifest as |mn| regardless of the nasality of the following vowel.

4.3.1.5 Stoney reflexes

- Fricatives tend to shift forward: *s > |θ|
  *š > |s|
- Free simple stops are voiced: *p > |b|
  *t > |d|
  *k > |g|
- *br always manifests as |mn|: *br > |mn|
- *kr always manifests as |hn|: *kr > |hn|
- *R manifests as |n|: *R > |n|
- *tk becomes |kt|: *tk > |kt| < *kt

4.3.2 Winnebago-Chiwere reflexes

Hoocąk and IOM share a number of innovations. The cluster *pt merges with preaspirate *ht. Proto-Siouan simple stop *p before vowels becomes |w|. Generally, it appears that the postaspirate stop series merges with the simple stop series. The *rh cluster also merges with the simple stop *t. As in Dhegiha, the presumed cluster *wR always seems to reduce to simple *|R|.

Both languages show a sporadic tendency to nasalize vowels that are not nasal in other MVS languages.[26] Both of them also sometimes replace a glottal stop with a glottalized |t^{ʔ}| following *i. This could be interpreted as an epenthetic |y| being naturalized as *|r|, and then converted to |t| before the glottal stop. The problem is that the glottal stop itself would seem to be in the way of obtaining the epenthetic |y| in the first place. Rankin suggests that in verb paradigms, the glottal stop is lost in conjugated forms, and that the conjugated form was recast back into the main verb.

- *pt becomes |ht|: *pt > *|ht| < *ht
- *rh becomes *|d|: *rh > *|d| < *t
- *wR merges with MVS *R: *wR > *|R| < *R
- *p becomes |w| before a vowel: *pV > *|wV|
- *i^{ʔ}V verbs become |it^{ʔ}V|: *i^{ʔ}V > *|it^{ʔ}V|

[26] Rankin et al. 2006: 50.

4.3.2.1 Hoocąk reflexes

Hoocąk shows quite a number of sound shifts of its own. One of its biggest is that it levels vowel length in monosyllables: the vowel of all monosyllabic words is long.[27] Further, it creates many new monosyllabic words by dropping the trailing final vowel, especially *-e. On top of this, it creates an extra syllable within an obstruent-sonorant cluster, by inserting the vowel that follows the cluster into the spot between the two consonants as well.[28] As in Mandan, Crow and Hidatsa, Proto-Siouan *y merges with *r. The *t series, except for glottalized t *t^{ʔ}, is affricated into a |č| series. An *r^{ʔ} cluster may become either |t^{ʔ}| or |k^{ʔ}|.[29]

- The vowels in monosyllables are always long.

  Long monosyllables: *CV(C) > *|CVV(C)| < *CVV(C)

- Trailing final vowels are often dropped, making even more monosyllables.

  Trailing vowels dropped: *CVCe > *|CVVC| < *CVVCe

- Obstruent plus sonorant clusters are broken up by insertion of the following vowel between the obstruent and the sonorant.

  Back insertion of vowel: *$C_{obst}C_{son}V_1$ > *|C_{obst}V1$C_{son}V_1$|

- As in Crow, Hidatsa, and Mandan, Proto-Siouan *y merges with *r.

  *y/*r merger: *y > |r| < *r

- *t series affricatizes: *t > |ǰ|

  *ht > |č|

- *r^{ʔ} becomes |t^{ʔ}| or |k^{ʔ}|: *r^{ʔ} > |t^{ʔ}| or |k^{ʔ}|

- *R manifests as |d|:[30] *R > |d|

[27] Rankin et al. 2006: 303, 797.

[28] Helmbrecht (2011: 123–124). This Hoocąk pattern of back-filling an obstruent-sonorant cluster with the following vowel is known as 'Dorsey's Law'.

[29] Rankin et al. 2006: 816-817.

[30] Helmbrecht (*Phonetics and Phonology*) and personal communication. This sound is written 't' in the CSD and in modern Wisconsin Hooąk orthography. But the 't' is voiced in prevocalic and intervocalic position, where it is the reflex of *R.

4.3.2.2 IOM reflexes

A distinctive feature of IOM is its forward shifting of the fricatives. Siouan *s becomes |θ|, and *š becomes |s|.[31] In clusters of *k before a fricative, the |k| is replaced by a glottal stop.[32] As in Kaw and Osage, the *t-series, including *t$^{?}$, is affricatized before a front vowel *i or *e. Initial *o- regularly becomes |u-|.[33]

One of the most interesting features of IOM is its treatment of the Proto-Siouan *y phoneme. As in several other Siouan languages, Proto-Siouan *y merges with another phoneme. Uniquely to IOM, however, the *y words are split about evenly between which other phoneme they merge with. Some of them merge with Siouan *r, as in Hoocąk, Mandan, Crow and Hidatsa. Others remain |y|, but these are joined by MVS *ž, which itself becomes |y| in IOM. The fact that many IOM *y fail to merge with *r is mentioned in the CSD, but the significance of the counter-merger of these *y with MVS *ž seems not to have been noticed. For IOM only, we must consider the *y phoneme to be two distinct phonemes, *y_1 and *y_2.

- Fricatives shift forward:

  *s > |θ|
  *š > |s|

- *k before fricative becomes |$^{?}$|: *kS > |$^{?}$S|
- Initial *o- becomes |u-|: *o- > |u-|
- *R manifests as |d|: *R > |d|
- *y_1 merges with *r as |r|: *y_1 > *|r| < *r
- *y_2 merges with *ž as |y|: *y_2 > |y| < ž
- *t-series affricates before *i/*e: *ti > *|či|
  *te > *|če|
  *hte > *|hče|
  *t$^{?}$e > *|č$^{?}$e|
  etc.

[31] Rankin et al. 2006: 245.
[32] Rankin et al. 2006: 857.
[33] Rankin et al. 2006: 893.

4.3.3 Dhegiha reflexes

Dhegiha is characterized by substantial shifts and mergers in its vowel structure. The nasal Proto-Siouan vowel *ų merges with *ą, producing a variably pronounced low back vowel with minimal rounding. The oral vowel *u also shifts forward to become |ü|. In Dhegiha, Siouan *y merges completely with MVS *ž. Unlike the other MVS languages, the preaspirate stops do not merge with another stop series. In most Dhegiha languages, these manifest as 'tense', or double-long unaspirated stops, but in Osage they manifest as preaspirates. Proto-Siouan *rh becomes |th|.[34] MVS stop clusters collapse into a single stop, of the preaspirate series. The clusters *ks and *ps become |s|, and the clusters *kš and pš become |š|.[35] Siouan *xw becomes |ph|.[36] As in Winnebago-Chiwere, the presumed cluster *wR always seems to reduce to simple *|R|.

- *ų merges with *ą: *ų > |ą| < *ą
- *u becomes *|ü|: *u > |ü|
- *y merges with MVS *|ž|: *y > |ž| < *ž
- *rh merges with *th: *rh > |th| < *th
- *xw merges with *ph: *xw > |ph| < *ph
- *ps and *ks merge with *s: *ps > |s| < *s
  *ks > |s| < *s
- *pš and *kš merge with *š: *pš > |š| < *š
  *kš > |š| < *š
- *wR merges with MVS *R: *wR > *|R| < *R
- Stop clusters merge with preaffricate stops (general pattern):

*pt	>	*|ht|	<	*ht
*pk	>	*|hk|	<	*hk
*tp	>	*|ht|	<	*ht
*tk	>	*|ht|	<	*ht
*kp	>	*|hp|	<	*hp
*kt	>	*|ht|	<	*ht

[34] Rankin et al. 2006: 165.
[35] Rankin et al. 2006: 64, 123, 222, 849.
[36] Rankin et al. 2006: 180.

4.3.3.1 Omaha-Ponca reflexes

Omaha and Ponca carry the vowel reorganization begun in Dhegiha even further. Dhegiha *ü, from Siouan *u, now loses its rounding and merges completely with Siouan *i. Behind it, the Siouan *o vowel is raised to |u|. Siouan *R manifests as |n|, thereby merging with the |n| from Siouan *r before a nasal vowel. The plain Siouan glottal stop disappears, while the glottalized velar clusters *k^{ʔ} and *x^{ʔ} both reduce to |^{ʔ}| as a neo-glottal stop. The preaspirate stop series manifest as tense, while simple stops are voiced. The postaspirate *ph usually, but not always, reduces to |h|. The Siouan *r phoneme manifests as what I call 'ledh', a quick, smooth, flip of the tongue from an apical |l| to edh and off the back of the front teeth. Linguists generally indicate it with the edh symbol, ð, though l and r would be equally reasonable choices. Additionally, an entire series of new stops is being generated from a custom of affricating the t-series stops as a "baby talk" method of suggesting smallness or cuteness.

- Dhegiha *ü merges with *i: *ü > |i| < *i
- *o becomes |u|: *o > |u|
- *R manifests as |n|: *R > |n| < *n < *r
- *^{ʔ} disappears: *V^{ʔ}V > |VV|
- *k^{ʔ} and x^{ʔ} become |^{ʔ}|: *k^{ʔ} > |^{ʔ}| < *x^{ʔ}
- *ph usually becomes |h|: *ph > |h|
- Free simple stops are voiced: *p > |b|
  *t > |d|
  *k > |g|
- Preaspirate stops are tense: *hp > |pp|
  *ht > |tt|
  *hk > |kk|
- Diminutive t-series transform: |d| dim.> |ǰ|
  |t| dim. > |č|
  |tt| dim. > |čč|
  |th| dim. > |čh|
  |t^{ʔ}| dim.> |č^{ʔ}|

4.3.3.2 Kaw-Osage reflexes

Kaw and Osage share a characteristic of dropping the velar stop from the *kr cluster and replacing the cluster with |l|. It seems that both of them also merge the glottalized fricatives *s^{ʔ} and *š^{ʔ} into a glottalized dental/alveolar affricate |c^{ʔ}| (|ts^{ʔ}|).[37] As in IOM, the *t-series, including *t^{ʔ}, is affricatized before a front vowel *i or *e.

- *kr drops the velar stop: *kr > |l|
- *s^{ʔ} and *š^{ʔ} merge as |c^{ʔ}|: *s^{ʔ} > |c^{ʔ}| < *š^{ʔ}
- *t-series affricates before *i/*e: *ti > *|či|
  *te > *|če|
  *hte > *|hče|
  *t^{ʔ}e > *|č^{ʔ}e|
  etc.

4.3.3.2.1 Kaw reflexes

Kaw agrees with Omaha and Ponca in voicing the free simple stops and in pronouncing the aspirated stops as tense. In Kaw, Siouan free *r manifests as |y|.

- Free *r manifests as |y|: *r > |y|
- *R manifests as |d|: *R > |d|

4.3.3.2.2 Osage reflexes

In Osage, the preaspirate series is pronounced with preaspiration, and the free simple stops are voiceless. Siouan free *r manifests as edh or ledh (ð). *ph manifests as |pš|.[38]

- Free *r manifests as |ð|: *r > |ð|
- *R manifests as |t|: *R > |t|
- *ph manifests as |pš|: *ph > |pš|

[37] Rankin et al. 2006: 856.
[38] Rankin et al. 2006: 64.

4.3.3.3 Quapaw reflexes

In Quapaw, free Siouan *r manifests as |d|. It seems that simple stops sometimes become tense.[39] The Siouan cluster *p$^{?}$ is reduced to plain glottal stop.[40]

- Free *r manifests as |d|: *r > |d|
- Simple stops may become tense: *t > |tt|
- *p$^{?}$ becomes |$^{?}$|: *p$^{?}$ > |$^{?}$|

4.4 Southeastern Siouan reflexes

Very few systematic sound shifts characterize Southeastern Siouan as a whole. One mentioned in the CSD is the loss of glottalized fricatives. Also, it seems that *š usually affricatizes to |č|.

- Fricatives lose glottalization and merge with the corresponding plain form. Thus, *S$^{?}$ > *|S|.[41]

  Fricatives deglottalize:

*s$^{?}$	>	*|s|	<	*s
*š$^{?}$	>	*|š|	<	*š
*x$^{?}$	>	*|x|	<	*x

- *š then usually becomes |č|:[42] *š > *|č|

4.4.1 Tutelo reflexes

Tutelo seems conservative. The only significant change noted involves the Proto-Siouan *š and *s phonemes.

- *š normally becomes |č|: *š > |č|
- Sometimes, *š becomes |s|:[43] *š > |s| < *s
- *s is indifferently pronounced:[44] *s > |s| or |š|

[39] Rankin et al. 2006: 833.
[40] Rankin et al. 2006: 831.
[41] Rankin et al. 2006: 856.
[42] Rankin et al. 2006: 99, 126, 167, 827, 931.
[43] Rankin et al. 2006: 912.
[44] Rankin et al. 2006: 54, 931.

4.4.2 Ofo-Biloxi reflexes

In Ofo and Biloxi, initial Proto-Siouan *w or *h before a vowel is lost.[45]

- *wV becomes plain |V|: *wV > |V|
- *hV becomes plain |V|: *hV > |V|

4.4.2.1 Biloxi reflexes

Biloxi is fairly conservative. Final *-i and *-e merge as |-i|,[46] and the glottal stop often appears as |h|.[47]

- Final *-e merges with *-i: *-e > |-i| < *-i
- The glottal stop becomes |h|: *ʔ > |h|

4.4.2.2 Ofo reflexes

Ofo is much more innovative. Proto-Siouan *y becomes aspirated |čh|,[48] as in Dakotan. The CSD suggests that Proto-Siouan *š before an accented syllable may have become aspirated |čh| as well.[49] Notably, the *s fricative changes to |f|, while Proto-Siouan *x shifts forward to become a neo-|š|.[50] Several of the Proto-Siouan clusters do interesting things as well. In the case of a glottalized stop consonant, the glottal stop seems to shift forward so that it releases prior to the stop. This phenomenon is suggested in Ofo transcriptions as a neutral vowel appearing epenthetically in front of the stop that in other languages is known to be glottalized. The stop consonant is then aspirated as well.

- *y becomes |čh|: *y > |čh|
- Accented *|š| becomes |čh|: *šV́ > *|čV́| > |čhV́|
- *s becomes |f|: *s > |f|
- *x becomes |š|: *x > |š|

[45] Rankin et al. 2006: 7, 223, 817, 929.
[46] Rankin et al. 2006: 901.
[47] Rankin et al. 2006: 103.
[48] Rankin et al. 2006: 85, 242.
[49] Rankin et al. 2006: 827.
[50] Rankin et al. 2006: 174, 299.

- *hs becomes |fh|:[51] *hs > |fh|
- *Cr becomes |l|:[52] *Cr > |l|
- *C^{ʔ} becomes |əCh|:[53] *C^{ʔ} > |əCh|

Abbreviations

CSD	Comparative Siouan Dictionary 2006
IOM	Iowa-Otoe-Missouria
MVS	Mississippi Valley Siouan

References

Helmbrecht, Johannes. *Phonetics and phonology*. Manuscript.

Helmbrecht, Johannes. 2011. The accentual system of Hocąk. In Jan Wohlgemuth & Michael Cysouw (eds.), *Rara and rarissima: Documenting the fringes of linguistic diversity* (Empirical Approaches to Language Typology 46), 117–143. Berlin: de Gruyter Mouton.

Kasak, Ryan M. 2016. A distant genetic relationship between Siouan-Catawban and Yuchi. In Catherine Rudin & Bryan James Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 5–37. Berlin: Language Science Press.

Parks, Douglas R. & Raymond J. DeMallie. 1992. Sioux, Assiniboine and Stoney dialects: A classification. *Anthropological Linguistics* 34(1). 233–255.

Rankin, Robert L., Richard T. Carter & A. Wesley Jones. 1998. Proto-Siouan phonology and grammar. In Xingzhong Li, Luis López & Tom Stroik (eds.), *Papers from the 1997 Mid-America Linguistics Conference*, 366–375. Columbia: University of Missouri Press.

Rankin, Robert L., Richard T. Carter, A. Wesley Jones, John E. Koontz, David S. Rood & Iren Hartmann (eds.). 2006. *Comparative Siouan dictionary*. http://csd.clld.org.

Rankin, Robert L., Richard T. Carter, A. Wesley Jones, John E. Koontz, David S. Rood & Iren Hartmann (eds.). 2015. *Comparative Siouan dictionary*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://csd.clld.org (22 May, 2015).

[51] Rankin et al. 2006: 174, 299.
[52] Rankin et al. 2006: 90.
[53] Rankin et al. 2006: 229, 232.

Rood, David S. & John E. Koontz. 2002. The comparative Siouan dictionary project. In William Frawley, Kenneth C. Hill & Pamela Munro (eds.), *Making dictionaries: Preserving indigenous languages of the Americas*, 259–281. Berkeley: University of California Press.

Shields, George, Sr. 2012. *Stories told by George Shields, Sr.* Douglas R. Parks, Raymond J. DeMallie & Linda A. Cumberland (eds.) (Assiniboine Narratives from Fort Belknap, Montana). Bloomington: American Indian Studies Research Institute, Indiana University.

Chapter 4

Ba-be-bi-bo-ra: Refinement of the Ho-Chunk syllabary in the nineteenth and twentieth centuries

Kathleen Danker

In 1885, a man from the Nebraska Winnebago Reservation learned the Great Lakes syllabary writing system from the Sac and Fox in Iowa and began adapting it to Ho-Chunk, the only Siouan language to be written in this syllabary. During the first half of the twentieth century, the Ho-Chunk syllabary continued to be refined by writers who created new symbols necessary for the additional Ho-Chunk vowels and consonants and discarded unnecessary Sac and Fox characters. Increasing correspondence to the phonemic characteristics of the Ho-Chunk language can be seen by comparing the version of it used by its original adapter in Nebraska published in 1890, a text composed in 1938 by Sam Blowsnake in Wisconsin, and one written in the 1970's by Felix White, Sr., of Winnebago, Nebraska, who referred to the Ho-Chunk syllabary as the *ba-be-bi-bo-ra*.

1 Introduction

The members of the Ho-Chunk, or Winnebago, tribe, with reservations in Wisconsin and Nebraska, are the only speakers of a Siouan language to have developed a phonemic written language system. Generally referred to as a syllabary, this type of orthography is more specifically termed an *abugida* as defined by Peter T. Daniels in his 1990 typology of writing systems. Rather than employing separate symbols for all possible syllabic combinations of consonants and vowels in a language, as does a classic syllabary, an abugida consists of a phoneme-specific consonant symbol followed by a secondary, also phoneme-specific, vowel symbol. In the case of the Ho-Chunk system, the vowel /a/ was usually inherent or unmarked, but all other vowels were marked and connected to the preceding consonant in a syllable. In this article, I will use the more familiar term syllabary for the Ho-Chunk system. The Ho-Chunk people were introduced to this type of

Kathleen Danker. 2016. Ba-be-bi-bo-ra: Refinement of the Ho-Chunk syllabary in the nineteenth and twentieth centuries. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 83–103. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.123 DOI:10.17169/langsci.b94.123

writing in the 1880's by members of the Sac and Fox tribe, with whom they had historical ties.

Little is known about the origins of the Great Lakes Algonquian Syllabary developed by the Algonquian-speaking Sac and Fox, Potawatomi, Kickapoo, and possibly Ottawa tribes in the late nineteenth century. First mentioned in print in 1880, its characters were based on cursive European handwriting, likely French (Walker 1996: 169). In 1884 or 1885, a man from the Nebraska Winnebago Reservation learned it while visiting the Sac and Fox in Iowa and adapted it to Ho-Chunk (Fletcher 1890b: 354). According to the ethnographer Alice Fletcher, by 1888 knowledge of the syllabary writing system had "spread rapidly among the Winnebagoes of Nebraska, and also to that part of the tribe living in Wisconsin, so that at the present time the principal correspondence of the tribe takes place by means of these characters" (1890a: 299). Considerable adaptation was necessary to convert the Sac and Fox system to Ho-Chunk. Sac (Sauk) and Fox (Mesquakie) have four vowels, all oral, while Ho-Chunk has five oral vowels and three nasalized ones. There are nearly twice as many consonants in Ho-Chunk as in Sac and Fox. The Sac, however, do have one consonant, /θ/, not found in Ho-Chunk, and the Fox included a single character in their syllabary for a consonant cluster, /kw/ (Jones 1906). (See Tables 1, 2 and 3 below.) Examples of the Ho-Chunk syllabary from the nineteenth and twentieth centuries show that it took some time for its users to discard unnecessary Sac and Fox characters and to develop the new ones required for their Siouan language.

Initial enthusiasm for the syllabary may have declined among the Ho-Chunk by 1912 when anthropologist Paul Radin reported that it "was known to very few Indians and was used only for the writing of letters" (Radin 1954: 21; cited in Walker 1981: 161). However, this writing of letters between Nebraska and Wisconsin did continue until correspondence in Ho-Chunk became outmoded because of general English literacy in the second half of the twentieth century. Up to that time, tribal members in Nebraska and Wisconsin worked on altering the orthography and the syllabification conventions of their writing system to better represent the Ho-Chunk language. Because this took place primarily within families, the syllabary inevitably developed into somewhat different versions. Nonetheless, an overall evolution of the Ho-Chunk syllabary can be seen by comparing and contrasting a version of it published by Alice Fletcher in 1890 (Tables 4 and 5 below); a syllabary text composed in 1938 by Sam Blowsnake in Wisconsin that was included in an unpublished manuscript written by Amelia Susman in 1939 (Table 6 below); and a fragment of a text written by Felix White, Sr., of Winnebago, Nebraska, in the late 1970's when he taught me his version

of the syllabary, which he called the *ba-be-bi-bo-ra* (Table 7 below). The published 1890 version of the Ho-Chunk syllabary contained additions to its Sac and Fox model as well as unnecessary retentions of some of its features. The texts of Blowsnake and White in the twentieth century conformed more closely to the phonemic characteristics of the Ho-Chunk language. The few ways in which these later texts continued to differ from these characteristics are of particular interest because of what they may reveal about how Ho-Chunk speakers conceptualize the phonemes of their language.

2 The Fletcher publications

Alice Fletcher published the first account of the Ho-Chunk syllabary, actually two slightly different accounts, both in 1890, one in the *Proceedings of the American Association for the Advancement of Science* and the other in the *Journal of American Folklore*. By chance, Fletcher had been on the Nebraska Winnebago reservation to witness the earliest inception of the writing system. In May 1883, she had been moved there from the adjacent Omaha Reservation to recuperate from a serious illness incurred while she was apportioning Omaha land allotments (Mark 1988: 90). Fletcher lay bedridden at the Winnebago Agency in the winter of 1883–1884 when some fifteen to twenty members of the Sac and Fox tribe living in Iowa came to visit the reservation. The Fox and the Ho-Chunk had been neighbors in Wisconsin since at least 1600 as on-again, off-again allies and enemies – although primarily allies after 1737, when the Ho-Chunk helped persuade the French not to exterminate the Fox at the end of the Fox Wars (Bieder 1995: 75). Fletcher wrote that the travelers from Iowa "were in full Indian costume and bent on enjoying old-time pleasures. I met these visitors several times; they came not infrequently to see me when I lay upon my sick bed" (1890b: 354). They told her that someone in the tribe had recently invented a writing system for their language, but none of them then visiting in Nebraska had learned it (1890a: 299).

In 1884 or 1885, a group of Ho-Chunk paid a return visit to the Sac and Fox, and one of them studied the Sac and Fox writing system. Once back at the Nebraska Winnebago reservation, he began to adapt the Sac and Fox orthography to the Ho-Chunk language and to teach it to others. According to Fletcher, he "became before long quite expert in its use, to his own amusement and that of his friends" (1890a: 299). Fletcher received a letter from the Winnebago agent in August 1885 stating:

> The tribe have suddenly taken to writing their own language, and people who have never learned English have acquired this art. The people claim they took the basis of it from the Sauk and elaborated it themselves. It is a very suggestive sight to see half a dozen fellows in a group with their heads together working out a letter in these new characters; it illustrates the surprising facility with which they acquire what they want to learn. (1890a: 299)

In 1887–1888, Fletcher returned to the Winnebago Reservation, having been hired to oversee allotment there as she had done on the Omaha Reservation (Mark 1988: 162). The Ho-Chunk man who had brought the Sac and Fox syllabary back to Nebraska arranged a table of his Ho-Chunk syllabary characters for Fletcher, which she then published in her 1890 articles about the writing system. Unfortunately, she did not report her informant's name in these articles (Fletcher 1890a: 300).

3 The phonemes of the Ho-Chunk and Sac and Fox languages

Table 1 shows the vowels and consonants of the Ho-Chunk language as organized by Amelia Susman in her dissertation on the Ho-Chunk accentual system (1943: 15). Her listing does not distinguish between long and short vowels or glottalized and unglottalized consonants as do more recent descriptions of the language such as that in *Hocąk Teaching Materials, Volume I* by Johannes Helmbrecht and Christian Lehmann (2010: 5–7). However, Susman's arrangement is particularly useful for this article because of the way it divides Category I consonant stops, affricatives, and spirants into two columns: voiced (sonant) in column A and voiceless (surd) in column B.

Tables 2 and 3 show the vowels and consonants found in the closely related Mesquakie (Fox) and Sauk (Sac) languages (Mesquakie-Sauk). To the right of each vowel in Table 2, and of each consonant in Table 3, I have added the corresponding Fox syllabary characters as described by William Jones in 1906. Length is phonemic in both Ho-Chunk and Sac and Fox vowels, something indicated in Table 2, but not marked in either the Ho-Chunk or the Sac and Fox syllabaries. Jones did not record a syllabary symbol for the consonant /h/ found in Sac and Fox as well as in Ho-Chunk, nor one for the consonant /θ/ found only in Sac. As can be seen by comparing Tables 1, 2 and 3, Ho-Chunk has the four oral vowels /a/, /e/, /i/ and /o/ used in Sac and Fox as well as a fifth Ho-Chunk oral vowel /u/. There are also nasalized /ą/, /į/ and /ų/ vowels in Ho-Chunk. Ho-Chunk shares

the consonants /č/, /h/, /k/, /m/, /n/, /p/, /s/, /š/, /t/, /w/ and /y/ with this Algonquian language, but in addition has the consonants /b/, /g/, /j/, /r/, /x/, /ɣ/, /z/, /ž/ and a glottal stop.

Table 1: Ho-Chunk phonemes (Susman 1943: 15).

Ho-Chunk Vowels						
Oral	a	i	u	e		o
Nasal	ą	į	ų			
Ho-Chunk Consonants						
	(Voiced)	(Voiceless)				
I	A	B	II		III	
stops	b	p	nasals	m	stops	t
	g	k		n		’
affricative	ǰ	č	trill	r	breath	h
spirants	z	s	semi-vowels	w		
	ž	š		y		
	ɣ	x				

Table 2: Mesquakie and Sauk vowel phonemes (adapted from Native Languages of the Americas 1998–2015; Susman 1939; Jones 1906).

Alphabetic	Syllabic	Americanist
a	ᓂ	a, ʌ
â		aˑ
e	ᒐ	e, ɛ
ê		eˑ, æˑ
i	ᓭ	i, ɪ
î		iˑ
o	ᐅ	o, ʊ
ô		oˑ

Table 3: Mesquakie and Sauk consonant phonemes (adapted from Native Languages of the Americas 1998–2015; Susman 1939; Jones 1906).

Alphabetic	Syllabic	Americanist
ch		č
h		h
k		k
m		m
n		n
p		p
s		s (only in Mesquakie)
sh		š
t		t
th		θ (only in Sauk)
w		w
y		y
kw		kw

4 The Fletcher syllabary

An early stage of how Ho-Chunk letter writers went about adapting the Sac and Fox syllabary to their own language can be seen in Tables 4 and 5, the Ho-Chunk syllabary table published by Fletcher in 1890. For easier display on the page, Fletcher's original table has been divided into two halves, one consisting of four vertical columns (one each for the vowels /a/, /e/, /i/ and /o/) of CV syllables and the other of four column of the same vowels in CVC syllables having the last consonant of /k/. This table appears odd to anyone familiar with the Ho-Chunk language, but its anomalies can be explained as resulting from a combination of the following factors: the arrangement of the table by syllables rather than by vowels and consonants; the retention of two Sac and Fox syllabary characters not corresponding to Ho-Chunk phonemes; the lack of syllabary characters for four Ho-Chunk vowels and seven Ho-Chunk consonants (counting the glottal stop); the difficulty of representing handwritten symbols in print; the misreading

Table 4: Fletcher's representations of CV characters in the Ho-Chunk syllabary with English pronunciations (adapted from Fletcher 1890a: 300).

Syll. = Engl.	Syll. = Engl.	Syll. = Engl.	Syll. = Engl.
Ka = gah	Ke = gay	Ki = gee	Ko = go
da = jah	de = jay	di = g	do = jo
wa = wä	we = we̲	wi = wï	wo = wo
xa = xä	xe = xe̲	xi = xï	xo = xo
ta = tdä	te = tde	ti = tdï	to = tdo
ma = mä	me = me	mi = mï	mo = mo
na = nä	ne = ne	ni = nï	no = no
La = Rä	Le = Ray	Li = Ree	Lo = Row
ga = gwar	ge = Gway	gi = gwee	go = gwo
ra = Sah	re = say	ri = see	ro = So
Tha = Thä	The = They	Thi = The	Tho = Tho
Ya = yä	Ye = yea	Yi = Ye	Yo = Yo
ba = pah	be = pay	bi = pee	bo = po
a = ä	e = e̲	i = ï	o = o
đa = shar	đe = shay	đi = shee	đo = sho
Aa = hah	Ae = hay	Ai = hee	Ao = ho

of handwriting; the use of a separate character for the consonant /k/ when it followed a vowel to end a syllable; and the ad hoc spelling of the English sound equivalents employed.

The arrangement of Tables 4 and 5 into syllables is the system used by the Sac and Fox and the Ho-Chunk, who put spaces between these syllables when they combined them into words and sentences. The two characters retained from the Sac and Fox syllabary, although not phonemes of the Ho-Chunk language, were a *Th*-shaped syllabary character standing for the Sac phoneme /θ/, which was not represented in Jones' Fox syllabary; and a *g*-shaped character used to represent a consonant cluster – although not the Fox consonant cluster /sk/, but one combining /g/ and /w/ into /gw/. Although it includes these unnecessary characters, Fletcher's table lacks characters for the Ho-Chunk oral vowel /u/, its three nasalized vowels /ą/, /į/ and /ų/, and its glottal stop /'/. The table also omits six other Ho-Chunk consonants, because it gives syllabary characters for only one of each of the paired voiced and voiceless stops, affricatives, and spirants

Table 5: Fletcher's representations of CVC characters in the Ho-Chunk syllabary with English pronunciations (adapted from Fletcher 1890a: 300).

Syll. = Engl.	Syll. = Engl.	Syll. = Engl.	Syll. = Engl.
Ka*m* = gark	Ke*m* = gake	Ki*m* = geek	Ko*m* = goke
da*m* = jark	de*m* = jake	di*m* = geek	do*m* = joke
wa*m* = wärk	we*m* = we̲rk	wi*m* = week	wo*m* = woke
xa*m* = xärk	xe*m* = xe̲rk	xi*m* = xeek	xo*m* = xörk
ta*m* = tdärk	te*m* = tde̲rk	ti*m* = tdeek	to*m* = tdörk
ma*m* = märk	me*m* = make	mi*m* = meek	mo*m* = moke
na*m* = närk	ne*m* = nake	ni*m* = neek	no*m* = noke
La*m* = Rark	Le*m* = Rake	Li*m* = Reek	Lo*m* = Roke
ga*m* = Gwark	ge*m* = Gwake	gi*m* = Gweek	go*m* = Gwooke
ra*m* = sark	re*m* = sake	ri*m* = seek	ro*m* = soke
Tha*m* = Thark	The*m* = Thake	Thi*m* = Theek	Tho*m* = Thoke
Ya*m* = Yark	Ye*m* = Yake	Yi*m* = Yeek	Yo*m* = Yoke
ba*m* = park	be*m* = pake	bi*m* = peek	bo*m* = poke
a*m* = ark	e*m* = ake	i*m* = eek	o*m* = oke
đa*m* = shark	đe*m* = shake	đi*m* = sheek	đo*m* = shoke
Aa*m* = hark	Ae*m* = hake	Ai*m* = heek	Ao*m* = hoke

shown in columns A and B of Category I in Table 1. There is a character for the voiceless /p/ but not the voiced /b/; the voiced /g/ but not the voiceless /k/; the voiced /ǰ/ but not the voiceless /č/; the voiceless /s/ but not the voiced /z/; the voiceless /š/ but not the voiced /ž/; and the voiceless /x/ but not the voiced /ɣ/.

As far as can be told from the printed letters used in the table to stand for cursive handwriting, the Ho-Chunk syllabary in 1890 used the same or similar characters as the Fox syllabary for the shared vowels and consonants /e/, /i/, /o/, /p/, /m/, /n/, /š/, /w/, /y/, and /t/. The Ho-Chunk symbol for the vowel /a/ looked like a handwritten lower case <a> rather than the handwritten lower case <u> that stood for /a/ in the Fox syllabary. The Ho-Chunk character for the consonant /g/, an upper case <K>, seems to be the same as that used for the Sac and Fox consonant /k/. However, the Ho-Chunk consonant /ǰ/ was not represented by the <tt> used by the Fox for the consonant /č/, but by a printed <d>. Fletcher's informant's syllabary used another printed <d>, this one italicized, to stand for the consonant /š/. Fletcher's informant also deviated from the Sac and Fox model

by using a printed lower case <r> to stand for the consonant /s/ although the Fox syllabary used an <s>. Perhaps when the Ho-Chunk adapter of the syllabary studied Sac and Fox syllabary writing, he saw a lower-case handwritten <s> that was not closed at the bottom, mistaking it for an <r>, because some Fox users of their syllabary wrote <s> that way (Walker 1996: 170).

The last four vertical columns of Fletcher's table, here reproduced in Table 5, contain some of its oddest-looking material, in both the Ho-Chunk syllables to the left of the equal signs and the English sound equivalents to their right. The italicized lower case <*m*> in each CVC Ho-Chunk syllable appears to represent an experiment with using a special character to stand for the consonant /k/ when it follows a vowel to end a syllable, /k/ being the consonant found most frequently in this position in Ho-Chunk. Fletcher indicated that, when handwritten, this <*m*> character was "more like a wavy line" (1890a: 300). Of the English sound equivalents given for the Ho-Chunk syllables, the <ar> constructions in the first column of Table 4, the <ark> constructions in the first column of Table 5 and the three cases of <erk> in the second column of Table 5 are the hardest to make sense of, because the Ho-Chunk language has no /r/-controlled vowels. It seems to me that the only possible explanation for these constructions is that someone mistook someone else's handwritten <h> for <r>. That would make the top horizontal progression in the CVC table through the four oral vowels for a syllable beginning with /g/ and ending in /k/ read as follows: "Ka*m* = gahk" (rather than gark), "Ke*m* = gake", "Ki*m* = geek", and "Ko*m* = goke". The third English sound equivalent here would be pronounced like the word in English spelled the same way, and the second and fourth English sound equivalents would each end in a silent <e> (Fletcher 1890a: 300).

Despite its shortcomings, Fletcher's table reproduced in Tables 4 and 5 shows that by 1890 the Ho-Chunk adapter of the Sac and Fox syllabary had instigated two important innovations to make it more useful for writing his language. He had created a character based on an upper-case handwritten <L> to represent the Ho-Chunk consonant /r/ and one resembling an upper-case handwritten <A> for the consonant /h/. The idea for this <A> character came originally from the Fox, but rather than using it for the consonant /h/, the Fox had put it before the vowel /a/ to represent that vowel when used in the initial position (Walker 1996: 170). Further steps toward making the Ho-Chunk syllabary accurately reflect the language can be seen in the writings of Sam Blowsnake and Felix White, Sr., in the twentieth century. While Fletcher printed only a few words in syllabary characters in her 1890 articles, both Blowsnake and White wrote manuscripts which can be examined for the way the syllabary worked in extended discourse.

5 The Blowsnake syllabary

Table 6: Ho-Chunk syllabary characters used by Sam Blowsnake in 1938 (Susman 1939), added to Susman's (1943: 15) arrangement of Ho-Chunk phonemes.

Ho-Chunk Vowels Oral										
a										
e										
i										
o										
u										
Ho-Chunk Consonants										
	(Voiced)		(Voiceless)							
I	A		B		II			III		
stops	b		p		nasals	m		stops	t	
	g		k			n			'	
affricative	ǰ		č		trill	r		breath	h	
spirants	z		s		semi-	w				
	ž		š		vowels	y				
	ɣ		x							

Table 6 adapts Table 1 by adding the syllabary characters written by Sam Blowsnake in 1938 to Susman's Ho-Chunk phonemes. We can see in Table 6 that, although Blowsnake still did not differentiate nasalized from oral vowels, he did sometimes use the character <u> for the fifth oral Ho-Chunk vowel /u/. This may have been a recent innovation for him, because he still usually wrote /u/ with an <o> as he did the vowel /o/, conforming most of the time to the four-vowel syllabary systems of the Sac and Fox and of Fletcher's informant. Blowsnake did not mark glottal stops, but he had a character for every other Ho-Chunk consonant. Along with the symbols for /m/, /n/, /r/, /w/, /y/, /t/ and /h/ seen in Fletcher's syllabary, he used a regularized system for indicating the paired stops, affricatives, and aspirants that Susman divided into columns A and B of her Category

1. Syllables omitted by Blow Snake, accidentally.
2. Phonemically ᵛ· can be broken down into šᵛ.

Figure 1: Ho-Chunk syllabary text written by Sam Blowsnake in 1938 (Susman 1939).

I consonants. He wrote the voiceless consonants in Column B as combinations of the characters for the voiced consonants in Column A with the symbol for /h/ (labeled by Susman as "breath") that the Fletcher syllabary had printed as an <A>. For voiceless /k/, he used two different characters – one of them the upper-case character <K> for a voiced /g/ plus the symbol <A>, and the other a single upper-case <K> written in a slightly different way than the <K> representing /g/. Blowsnake wrote the Sac and Fox symbol <tt> for /ǰ/ rather than the <d> character employed by Fletcher's informant, and he did not use a wavy-line <*m*> to indicate a /k/ following a vowel in a CVC syllable. That aspect of the syllabary recorded by Fletcher had apparently been abandoned.

Figure 1 shows a page of a syllabary text written by Blowsnake that Susman appended to the end of her *Winnebago Syllabary* manuscript. It consisted of a description of traditional child-rearing instructions given to boys, entitled "Child Teaching." Apparently, Blowsnake wrote the syllabary portion of it, and Susman typed the transcription, the English translation, and the two footnotes at the bottom. I have numbered the lines. In this text, we can see that the vowel /a/ was omitted after consonants in Ho-Chunk syllabary writing except when it followed a glottal stop as it did at the end of line 10 in the iterative suffix *-s'a.* The glottal stop itself was not marked in Blowsnake's syllabary, but its presence in this morpheme was indicated by the writing out of the usually unmarked vowel character <a>. Blowsnake did not mark glottal stops occurring before other vowels. Fletcher's informant's syllabary table listed syllables containing /a/ along with those using the other three marked Ho-Chunk oral vowels. Without the evidence of Ho-Chunk syllabary texts from the nineteenth century, I do not know whether or not the earliest Ho-Chunk letter writers left out the vowel /a/ when writing syllables. The Sac and Fox syllabary most closely resembling the Ho-Chunk one did not omit /a/ or any other vowel, but other early forms of Sac and Fox syllabaries have been recorded, one of which omitted /a/ after a consonant and used a dot to indicate /e/ after a consonant, a raised dot to indicate /i/, and two dots to indicate /o/ (Walker 1981: 158–159).

The lower-case handwritten <r> that Blowsnake used to write the /s/ in *-s'a* at the end of line 10 in Figure 1 should be a <rA>, but he frequently wrote the voiced rather than the voiceless syllabary characters for voiceless consonants, especially when they were not phrase initial or were part of a consonant cluster. Halfway through line 4, we can see that Blowsnake used his *y*-shaped syllabary character for the glide sound in the phrase *čiiokisak* ('center of the lodge'), which he spelled <ttAiyo Ki rAK>. However, in the middle of line 9, he also used this syllabary character for the diphthong in the first syllable of the clause *nąąįreže*

('they slept, it is said', pronounced like the English word *nigh*), which he spelled <ny Le de>.

Another characteristic of Blowsnake's syllabary style seen in Figure 1 was his regular insertion of the same vowel between the two consonants of a consonant cluster as the vowel that followed the cluster. If the syllable ended in a consonant, he repeated the vowel after that consonant, too. This can be seen in *wąąkšik* ('Indians' or 'people'), the third word in line 1, which he spelled <w KidiKi>. In this word, he omitted, as he often did, the <A> from the symbol standing for /k/. When he did write the <A> to make a consonant voiceless, he sometimes inserted a vowel between the character for the consonant and the <A>. For example, in the middle of line 3, he spelled the morpheme *pįį'ų* ('make good') <biAi o>. At the ends of lines 1, 4, 5, 7 and 10, and elsewhere in Figure 1, we can see that Blowsnake employed punctuation in his writing, in that he marked the end of sentences with periods.

In her unpublished analysis of Blowsnake's syllabary text, Susman came to the conclusion that it employed "a fairly consistent method of representing syllables by symbols derived from English script." However, she noted three inefficiencies:

1. Duplication of symbols for a single sound, most notably the two symbols for /k/.

2. Ambiguity of symbols, such as in the use of <o> for both /o/ and /u/ sounds. She found especially troubling the ambiguity caused by the insertion of extra vowels in CCV and CVC syllables that sometimes made it impossible to tell them apart.

3. Some inconsistency of syllabification.

6 The White syllabary

The syllabary writing system used by Felix White, Sr., shown in Table 7 and Figure 2 was quite similar to that of Sam Blowsnake, but it addressed some of the inefficiencies mentioned by Susman. When Mr. White first began giving Ho-Chunk language classes at the Little Priest Community College in Winnebago, Nebraska, in the early 1970's, he used the syllabary in teaching the language. He also taught me the syllabary characters in the late 1970's when I expressed an interest in learning Ho-Chunk. He told me that he and his aunt Florence Mann had made several changes to the syllabary, including marking nasalization. He used his nasalization marks fairly often, but not always. They consisted of lines

extending off of the top of vowels to the right. He indicated a nasalized unmarked /a/ by drawing a line from one consonant to the next over the space where /a/ would have appeared had it been marked. (See the second syllable in the first word, *hoočąk*, in line 4 of Figure 2 below.)

Table 7: Ho-Chunk syllabary characters used by Felix White, Sr., in the 1970's, added to Susman's (1943: 15) arrangement of Ho-Chunk phonemes.

Ho-Chunk Vowels										
Oral			Nasal							
a			ą							
e										
i			į							
o										
u	o, oo		ų							
Ho-Chunk Consonants										
	(Voiced)		(Voiceless)							
I	A		B		II			III		
stops	b		p		nasals	m		stops	t	
	g		k			n			'	
affricative	ǰ		č		trill	r		breath	h	
spirants	z		s		semi-	w				
	ž		š		vowels	y				
	ɣ		x							

Unlike Blowsnake, White used only one character for the consonant /k/, but like Blowsnake, he employed a new symbol, in his case <oo>, to stand for the oral vowel /u/. However, also like Blowsnake, he almost never used it, much more frequently writing <o> for both /o/ and /u/. Figure 2 shows that White marked glottal stops, and that he wrote a more open form of the <A> character standing for /h/ than Blowsnake did. He said that this character was not supposed to represent an <A>, but a star shape. Like Susman, he also called it a "breath" symbol. He wrote the character he used for the semi-vowel /y/, not like Blowsnake's lower-case English <y>, but more like the MesquakieFox symbol recorded by Jones (1906) for this phoneme, which looked something like the front part of

Figure 2: Ho-Chunk syllabary text written by Felix White, Sr., in the 1970's (Danker papers).

a number <2>. One case in which White used less regularized syllabary characters than Blowsnake had to do with the voiced consonant /z/ and the voiceless consonant /s/. White wrote /s/ with the character <d> and /z/ with the character <dA> instead of the other way around.

These and other aspects of the syllabary as written by White can be seen in Figure 2. It is the first page of a transcription White made of an audio tape of a story told by a friend of his named Jim Frenchman. I remember that in English White called this story "The Killing of the Body Outright." White transcribed the story using the syllabary and annotated it himself based on his studies of linguistics. I found this single page in some of my papers a few years ago and have added the line numbers. Line 6 shows an example of nasalization marked on the second and third syllables of its last word *hočįčį*, which means 'boys'. White followed the practice of appending an <A> to a voiced Category I consonant to make it voiceless more consistently than Blowsnake did, though he did not always do it, and never in the case of /s/. Except in one instance on this manuscript page, White did not insert extra vowels between consonant clusters or after syllable-ending consonants. The exception occurs at the beginning of line 6, where he put an <o> with a nasalization line (standing for /ų/) between the consonant cluster beginning the first syllable of the dubitive enclitic *šgųnį*, which he spelled <dAoko ni>.

In the middle of line 8, in the iterative marker *-s'a*, we can see that, like Blowsnake, White wrote out the usually unmarked vowel <a> to indicate a glottal stop preceding it, but he also included a mark like an apostrophe for the stop itself. He also used this glottal-stop marker when the glottal stop occurred with vowels other than /a/. White employed his *2*-shaped syllabary character for the semi-vowel /y/ where it was appropriate at the beginning of the second syllable of *čiyo*, the first phrase in line 10. However, he did not use it at the beginning of line 9 to represent the diphthong pronounced much like the English word *why* in the second to the last syllable of the clause *čąqgrogiwaire* ('they went outside'), as Blowsnake would have done. Instead, White indicated this diphthong by writing out the vowels <a> and <i> and extending a curved line upward from the end of the <i>. In this way, he differentiated the semi-vowel /y/ from a diphthong that sounded nothing like it. Like Blowsnake, White marked the end of sentences in his syllabary writing with periods, but he also used commas to separate clauses within sentences, making the structure of these sentences easier to determine when read. Nonetheless, it remained difficult for the recipients of letters written using the syllabary to always tell how to group syllables into the words their writers intended. White told me that people would puzzle over some parts of a

letter they received because of this, but they could usually figure out the correct words and meanings in the context of the rest of the message.

7 Conclusion

Comparing the Ho-Chunk syllabary of Alice Fletcher's unnamed informant with the writings of Sam Blowsnake, and Felix White, Sr. reveals steady progress from the 1880's into the last half of the twentieth century toward correspondence with the phonemic characteristics of the Ho-Chunk language. The users of this writing system would no doubt have continued to refine it had it not become unnecessary for communication between Wisconsin and Nebraska when almost all tribal members learned to speak and write English. Examination of these syllabaries also reveals some instances in which their practitioners deviated in similar ways from strict phonemic fidelity to Ho-Chunk. Perhaps these reoccurring "errors" indicate something about how Ho-Chunk speakers conceptualized the phonemes of their language.

That Fletcher's informant wrote down syllables containing only the four oral vowels /a/, /e/, /i/ and /o/ in his table could be attributed to the influence of the Sac and Fox model from which he worked. However, it could also reflect a perception that the three nasalized Ho-Chunk vowels represented only a subset of their oral counterparts. Of the three syllabary writers, only White decided to mark nasalization. Perhaps there was also a sense that the vowels /o/ and /u/ were related in some way that made their differentiation unnecessary. It is interesting that while Blowsnake and White each developed a new character to represent /u/ as a separate vowel, they almost always conformed in their writing to the pattern of Fletcher's informant by writing down the same character for /u/ as for /o/. Even the word that White used to refer to the syllabary, *ba-be-bi-bo-ra*, is made up of syllables containing only the vowels /a/, /e/, /i/ and /o/, followed by the definite enclitic *ra*.

The manner in which the three syllabary versions examined in this paper handled the voiced and voiceless consonants listed by Susman under Category I in Table 1 also invites speculation. The Sac and Fox syllabary contained characters for all of the consonants that Susman listed in Categories II and III for Ho-Chunk in Table 1 except for /r/ and the glottal stop, which are not found in the Sac and Fox language, and /h/, which the Sac and Fox syllabary did not mark. Fletcher's informant virtually copied the Sac and Fox syllabary characters for these shared Category II and III phonemes (/m/, /n/, /w/, /y/ and /t/) in his Ho-Chunk syllabary, yet he treated Susman's Category I consonants differently. Sac and Fox

shares five of these consonants with Ho-Chunk, all of them voiceless (/č/, /k/, /p/, /s/ and /š/). Ho-Chunk has, in addition, the voiced counterparts of these consonants (/ǰ/, /g/, /b/, /z/ and /ž/) as well as the voiceless and voiced pair /x/ and /ɣ/. Fletcher's informant, however, did not simply retain in his Ho-Chunk syllabary the five voiceless consonants for which the Sac and Fox had invented syllabary characters. Instead, he kept three of these voiceless consonants (/p/, /s/ and /š/) and added two voiced ones (/ǰ/ and /g/) and the unvoiced /x/. In doing so, he devised a system which included a character for only one of each of the paired Ho-Chunk voiced and voiceless stops, affricatives, and spirants in Category I. I assume that in writing he used the same character for both the voiced and voiceless consonants in a pair. Perhaps we can infer from this that he was aware that these consonants formed linked pairs and thought each pair so closely related that one syllabary character would suffice for both of its parts. This system of having one syllabary symbol stand for two different consonants when they consist of voiced and voiceless pairs was also practiced by the Potawatomi in their Great Lakes Algonquian syllabary. The Potawatomi language has phonemes for all of the voiced and voiceless consonant pairs listed by Susman for Ho-Chunk except for /ɣ/ and /x/. The writers of the Potawatomi syllabary used the character <l> for /b/ and /p/, <s> for /z/ and /s/, <sh> for /ž/ and /š/, <tt> for /ǰ/ and /č/, and <K> for /g/ and /k/. Their language also has a voiced and voiceless pair /d/ and /t/ that they wrote with a <t> (Walker 1996: 172, Fig. 11).

There can be no doubt that Blowsnake and White knew of the relationship between the paired voiced and voiceless consonants in Ho-Chunk, because of the way they sometimes, at least, wrote them with a character for each voiced consonant to which they added the <A> symbol to indicate its voiceless counterpart. When White started teaching me Ho-Chunk in the 1970's, he had me make a set of flash cards with which to practice recognition that had the syllabary characters on one side and their English equivalents on the other. These cards clearly showed White's syllabary system for marking voiced and voiceless consonants as well the fact that White switched that system around in the case of /z/ and /s/. This reversal appears to be a reversion to the way Fletcher's informant wrote the symbol for /s/ as <r>. In a way, it goes along with Blowsnake's frequent, and White's own occasional, practice of writing voiceless consonants as voiced. This inconsistency seems to indicate a persistent perception, reminiscent of those of Fletcher's informant and the Potawatomi syllabary writers, that paired voiced and voiceless consonants are so closely united with each other that it is not crucial to differentiate them in writing, that they are interchangeable. Perhaps this sense of interchangeability comes from the way that voiceless

consonants do sometimes change to voiced ones, and vice-versa, at the intersection of morphemes in spoken Ho-Chunk. For example, the voiceless consonant at the end of the word *waač* ('boat') becomes voiced when the definite enclitic *-ra* ('the') is added to it, forming *waaǰra* ('the boat'). Conversely, the voiced consonant at the beginning of the causal suffix *-geǰįnį* becomes voiceless when it is added to the clause *hįwuus*, 'I was dry,' to become *hįwuuskeǰįnį* ('because I was dry') (Lipkind 1945: 33 & 41).

In their efforts to preserve and teach their language today, the Ho-Chunk in Wisconsin have collaborated with linguists to devise an orthography that is more completely phonemic, more easily printed, and more comprehensible to a wider range of readers than was their handwritten syllabary. However, the native-language syllabary the Ho-Chunk developed and refined from the late nineteenth to the mid twentieth centuries was important. It provided families and friends with a way to communicate and maintain ties between Nebraska and Wisconsin through years of forced separation, persistent poverty, and cultural suppression. It contributed to the Ho-Chunk tribe's survival as a people.

References

Bieder, Robert E. 1995. *Native American communities in Wisconsin, 1600–1960: A study of tradition and change*. Madison: University of Wisconsin Press.

Daniels, Peter T. 1990. Fundamentals of grammatology. *Journal of the American Oriental Society* 119(4). 727–731.

Fletcher, Alice C. 1890a. A phonetic alphabet used by the Winnebago tribe of Indians. *Journal of American Folklore* 3(11). 299–301.

Fletcher, Alice C. 1890b. The phonetic alphabet of the Winnebago Indians. In Fredric W. Putnam (ed.), *Proceedings of the American Association for the Advancement of Science, 38th meeting, held at Toronto, Ontario, August, 1889*, 354–357. Salem, MA: Salem Press Publishing & Printing Company.

Helmbrecht, Johannes & Christian Lehmann (eds.). 2010. *Elements of grammar, learner's dictionary* (Hocąk Teaching Materials 1). Albany: State University of New York – Albany Press.

Jones, William. 1906. An Algonquian syllabary. In Berthold Lanfer (ed.), *Boas anniversary volume: Anthropological papers written in honor of Franz Boas*, 88–93. New York: G. E. Stechert.

Lipkind, William. 1945. *Winnebago grammar*. New York: King's Crown Press.

Mark, Joan T. 1988. *A stranger in her native land: Alice Fletcher and the American Indians*. Lincoln: University of Nebraska Press.

Native Languages of the Americas. 1998–2015. *Mesquakie-Sauk (Sac and Fox).* http://www.native-languages.org/meskwaki-sauk.htm2015.

Radin, Paul. 1954. *The evolution of an American prose epic: A study in comparative literature*. Basel: Ethnographical Museum.

Susman, Amelia L. 1939. *The Winnebago syllabary*. Manuscript, Freeman guide no. 3903. Philadelphia: Library of the American Philosophical Society.

Susman, Amelia L. 1943. *The accentual system of Winnebago*. New York: Columbia University (Doctoral dissertation).

Walker, Willard. 1981. Native American writing systems. In Charles A. Ferguson & Shirley Brice Heath (eds.), *Languages in the USA*, 145–174. Cambridge: Cambridge University Press.

Walker, Willard. 1996. Native writing systems. In William C. Sturtevant & Ives Goddard (eds.), *Languages* (Handbook of North American Indians 17), 158–184. Washington: Smithsonian Institution.

Chapter 5

A forgotten figure in Siouan and Caddoan linguistics: Samuel Stehman Haldeman (1812–1880)

Anthony Grant

In the light of Bob Rankin's Dhegiha work, this paper examines some of the earliest recorded material on Kanza and Osage, collected and transcribed by the naturalist Samuel Stehman Haldeman in an alphabet of his own devising (Haldeman 1859; 1860). Although his transcriptions fail to capture many crucial phonetic and phonemic distinctions, they are useful as records of earlier and more conservative forms of these languages.

1 Introduction

Robert Rankin's examinations of earlier sources on Native American languages which have rarely been the subject of fuller description impel us to look at the work of other early collectors of data on Siouan and Caddoan languages. We may mention for instance his paper on Max von Wied's (1839–1841) brief vocabulary of Kaw, Kanza or Kansa (Rankin 1994), Nor should we overlook his splendid salvage work on Kanza (the name I will use henceforth in this paper) and Quapaw, and his pivotal role in the organization of the Siouan-Caddoan Conferences.

One researcher is almost overlooked nowadays (despite a memoir by Lesley 1886 which hymns his activities while getting its dedicatee's name wrong). The naturalist, sawmill manager and avocational linguist Samuel Stehman Haldeman (1812–1880) was mostly known to the linguists in the 19th century for his 'Analytic Orthography' (Haldeman 1859, also produced in book form as Haldeman 1860). This was a prizewinning attempt to construct a universal phonetic alphabet, based on Latin letters (and following some precepts of classical Ciceronian Latin pronunciation, for instance <C> for /k/ and <V> for /w/) but enhanced with some created symbols. It also added a number of diacritics, for documenting

Anthony Grant. 2016. A forgotten figure in Siouan and Caddoan linguistics: Samuel Stehman Haldeman (1812–1880). In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 103–114. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.124 DOI:10.17169/langsci.b94.124

phonetic data in the world's languages, especially from previously undescribed languages of North America and elsewhere.

This work represented a determined effort to describe and notate sounds of speech, for which it was awarded the Trevyllian Prize in London against eighteen contenders. And there its reception ground to a halt. The alphabetic system, based on what Haldeman assumed were Classical Latin letter-values, was well-adapted to indicate certain aspects of vowel quality and quantity and basic consonantal distinctions. But his pioneering work is one of several such pre-International Phonetic Association schemes proposed in the 19th century, of which Lepsius (1863) is the most famous and influential, and it is cumbrous. Because it required a large number of special fonts and diacritics it was difficult to reproduce, with the result that nobody save Haldeman ever adopted it.

Haldeman's system is elaborate but just how successfully or consistently he applied his own transcriptional system is moot. For instance in his Chinese data (actually Cantonese), he fails to indicate any tones for the numerals for Guangzhou Cantonese, although he makes an effort to do so for the nearby Macanese variety of Cantonese. Haldeman uses the phonetic terminology of his time, with *surd, sonant, lenis, asper*, employed where modern phoneticians talk about *voiceless, voiced, unaspirated* and *aspirated* sounds, and with *sigmal, lingual, cerebral, guttural* and *faucal* used for modern *alveolar, dental, retroflex, velar* and *uvular* respectively. He also talks about "pure" (non-nasalized) and nasal sounds, and arrays consonants according to their degree of "interruption" (plosives are the most "interrupted" consonants in this scheme); see Figure 1. He also divided consonants into *mutes* (plosives and nasals) and *liquids* (other consonants). Haldeman (1860: 83, 369) recognizes thirty-four vowel qualities, which he arranges in a dense A-shaped diagram, and he indicates vowel quantity with macrons and breves. Unlike Daniel Jones (1909) he does not propose a scheme in which the distinction between back [ɑ] and front [a] is crucial.

The concentration in this work is on Haldeman's Dhegiha-language data, although observations from his work on Caddo and Wichita will be added where relevant. (Unfortunately I lack sufficient modern lexical data to give as full an analysis with modern examples of the Caddo and Wichita data as I would wish.) Data from Haldeman's work are taken from Haldeman (1860), a corrected and book-length edition of Haldeman (1859). Haldeman divided his work into sections (often extremely short and usually corresponding to paragraphs), in addition to the book being paginated. Both modes of reference will be used here.

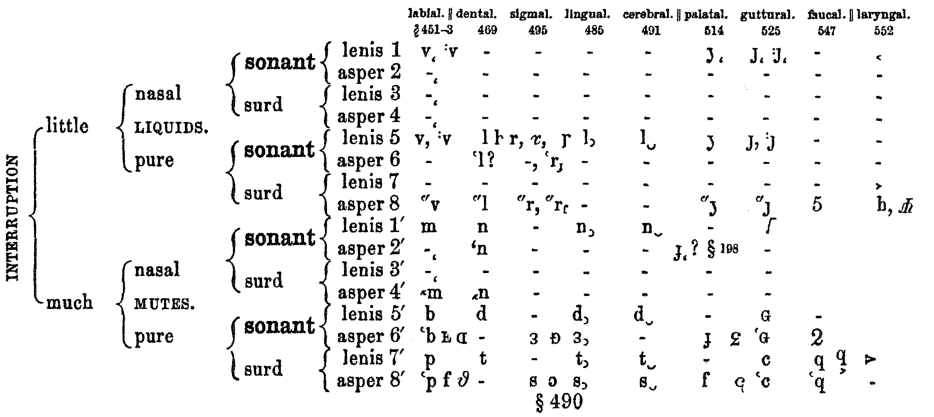
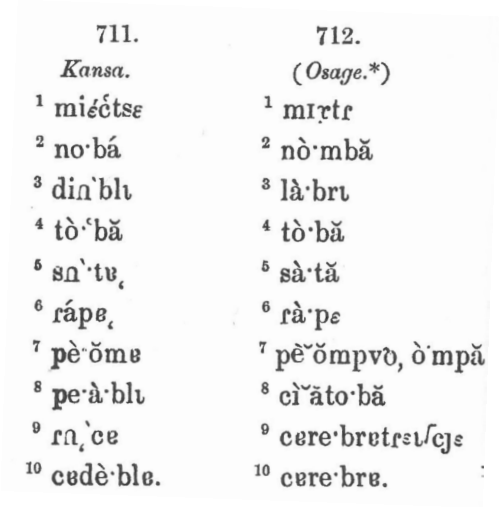


Figure 1: A digest of consonantal symbols in Haldeman 1860: 121: §576

2 Haldeman the Americanist and his work on Dhegiha

In addition to a number of versions of the Lord's Prayer (including those in Cherokee and Wyandot) Haldeman (1860) provided data in the form of 75 cardinal numeral sets from 1–10 from a wide range of languages of Europe, Asia and North America (plus Grebo from Liberia). A number of these were Algonquian, Muskogean and Iroquoian languages in addition to numerals from Makah, Chinook, Comanche, Jicarilla Apache and the Yuman language 'Iipay Aa. Among the languages on which Haldeman tried out his spelling system are the Dhegiha Siouan language Kanza (for which he also provided some Santee parallels for certain forms from Riggs 1852; see Figure 3). He also provided data from the Caddoan languages Caddo (which Haldeman referred to as "Nadaco") and Wichita (referred to as "Waco" but identical with the Wichita recorded in the 20th and 21st centuries, for instance in Rood (1975). In each case the data presented are cardinal numerals from 1–10 and some additional lexicon (over 40 such items in the case of Caddo and 10 from Wichita), evidently recorded by Haldeman from native speakers and not previously listed elsewhere. Haldeman also collected the numerals from 1–10 in Osage; see Figures 2 and 3.

As Rankin (1994) showed, Max von Wied (zu Wied-Neuwied 1839–1841) had described sounds in various Dhegiha languages (Wied documented Omaha, Ponca, Kanza and Osage) quite well within the limitations of his annotated Franco-German spelling system. This means that though his work is superb for its time



711. Kansa.	712. (Osage.*)
[1] miéćtsɛ	[1] mɪɾtɾ
[2] no·bá	[2] nò·mbă
[3] diɲ`blɪ	[3] là·brɪ
[4] tò·ʿbă	[4] tò·bă
[5] sɲ`·tu꜀	[5] sà·tă
[6] ɾápɐ꜀	[6] ɾà·pɛ
[7] pè·ŏmɐ	[7] pĕ˘ŏmpvò, ò·mpă
[8] pe·à·blɪ	[8] cì˘ăto·bă
[9] ɾɲ꜀`ce	[9] cɐre·brɐtɾɛɪ/cjɛ
[10] cɐdè·blɐ.	[10] cɐre·brɐ.

Figure 2: Numeral data from Kanza and Osage (Haldeman 1860: 146, §711, 712)

KANSA (=cɔnᵊɔ.)

634. *The vowel* Y, French *u*, is found here, although very rare in the aboriginal languages of North America. In our examples we add (in parentheses) the Dacòta equivalents, but placing Riggs' diacrits *after* instead of over the letters, as g· (which is compared with ghain,) s′, English *sh;* h·, 'a deep surd guttural;' c′, Eng. *tsh;* z′ Eng. *dzh;* *n*, as in English *sing*, and French *bon*, the two being confounded after eighteen years study by a number of missionaries. Probably both sounds occur, as in Kansa.

ear,	nɐtá (nog·e)	*forehead,*	pĭèsᵋsᵋ (ité)
eye,	ɪɾtátɐ˘/ɐ˘ (is´tá)	*fan,*	ĭcīlăɟɐ (ic´ádu)
brow,	ɟɾtáhɪɾăbă (is´táh·e, *eye-ridge.*)	*pipe,*	nŏnóbă (c´*a*ndúhupa, c´*a*ndí, tobacco.)
mouth,	ìhă (i)	*knife,*	mòhɛ (mi*n*ná)
tongue,	ɟéssɛ* (c´ez´í)	*warm,*	mɐ.ɾtjèa˙v
nose,	pɲh* (póg·e)	*leggins,*	hy`/ɢă (hu*n*ská)
nostril,	pɲɾʊ′ɾɐ (póg·e-oh·dóka, nose-hole.)	*shirt,*	ɔscɔ́scŏʊcɹdo꜀ʊdă (o*n*-h·doh·da.)

* Pronounce each *s*, and the *h*.

Figure 3: Lexicon from Kanza, with some Dakota parallels from Riggs (1852) (Haldeman 1860: 135, §634)

he missed many crucial details and failed to record other details consistently. Haldeman's system was theoretically more precise as far as it went (although there is little consistent coverage of tones and essentially none of consonants which are ingressive, velaric, or other kinds of clicks). But it was deployed less consistently and less accurately. His records of Kanza and Osage do show an ability to indicate primary stress using acute accents, while grave accents are used to indicate a variety of vowel qualities, short vowels are marked with breves, long vowels with postvocalic dots, and vowel nasalization is represented with the ogonek or Polish hook placed at the bottom of the line after the vowel.

Working on small amounts of material (often only the numerals from 1–10) from a large number of languages, Haldeman recognized that some sounds were problematic in terms of his descriptive criteria, as his discussion of the two ejectives Caddo /t'/ and Wichita /k'/ shows (Haldeman 1860: 99: §448; 131: §574). But he did not make the leap (as García 1760 had done for Coahuilteco) by discovering that what made these sounds distinctive from other speech sounds but similar to one another was their common possession of ejective quality, with the corollary that ejectives should be represented consistently. As a result he was unable to indicate the ejective quality of the final consonant in Caddo *wists'i* 'one'. In fact, Haldeman's attempts at transcribing Caddo (for instance Haldeman's <vátɐt'> for *waadat* 'earth', in which he fails to hear that the medial stop is voiced: Haldeman 1860: 135, §633) are scattershot enough to be unreliable. Even so he recognized that the Caddo word for 'cheek' used a dental or alveolar rather than a velar nasal.

His array of consonantal types was defective in other respects. Although Dhegiha languages contain ejectives, the small samples from Kanza and Osage which Haldeman cites happen not to include any of these sounds; Haldeman would probably have been unable to indicate these, as they are not provided for in his consonantal chart, and his encounters with them in Caddo and Wichita left him uncertain as to the nature and phonetic structure of the ejectives which he encountered there. He also lacked a consistent way of indicating the glottal stop, either initially, medially or finally, which is a special problem when recording Caddo data. Nor did Haldeman's system capture the three degrees of phonemic vowel length which are present in Wichita (although Haldeman 1860: 80, §353–355) provides the wherewithal to do this.

As a result of these and other shortcomings, Haldeman's work has received rather little attention from modern phonologists or indeed other linguists. Even Haldeman himself made no use of the system in his work on Pennsylvania German (Haldeman 1872). The discussion in Pilling (1887) and the brief account in

Kelly & Local (1989), written incidentally by the academics who taught phonetics to this author, are rare exceptions to this neglect.

Haldeman's data on Osage, comprising merely the cardinal numerals from 1 through 10, and the corresponding forms in Kanza, help us to get a better sense of his transcriptional techniques. Modern Osage data are from Quintero (2009) and Kanza data from Cumberland & Rankin (2012). Original transcriptional systems have been preserved. We note that the two languages, though very close, are represented differently in regard to orthographic conventions employed to indicate postalveolar sibilants, vowel length and nasalized vowels.

Dhegiha languages share a number of crosslinguistically marked features in their segmental phonology. These include the differentiation of nasalised from oral vowels, the differentiation of geminate and lengthened stops, of preaspirated and voiceless ejective stops, and the use of a high front rounded vowel. Modern forms are given below (Kansa <u> is /y/ and superscript <n> represents nasalization, indicated in Osage by an ogonek).

3 Modern counterparts of the data

In Tables 1 and 2 are given modern equivalents in Kanza and Osage for Haldeman's data in Figures 2 and 3.

Note that what are written as single plosives in the modern Kanza orthography are actually geminates, thus <k> is /kk/.

4 Remarks on the forms

The materials here represent examples of impressionistic phonetic transcriptions, which is what we would expect in a work from the pre-phonemic era. The Kanza and Osage words in Haldeman's material (especially the former) are recorded with comparatively greater detail than numerical data from some of the other languages are. Indeed the Kanza numerals are recorded with greater detail by Haldeman with respect to accent than they are in the present orthography. But the forms are not necessarily noted with greater accuracy, and neither system indicates the differences between the various voiceless stop series clearly. Tense stops in Kanza in Haldeman's transcription are represented by the use of bold consonantal characters, so that Haldeman's <p> is [p ∼ ph], <**p**> is [pp] while <ʿp> is [hp] in his Kanza work. (This fact is clearer in the version of Haldeman's work published in the *Transactions of the American Philosophical Society* than in

Table 1: Cardinal Numerals (Haldeman 1860: §711, 712)

	Kanza	Osage
1	mi^{n}xci	wįxce
2	no^{n}bá	ðǫǫpa
3	yábli^{n}	ðáabrįį
4	dóba	tóopa
5	sáta^{n}	sáhtą
6	shápe	šáhpe
7	péyo^{n}ba	hpéeðǫǫpa
8	kiadóba	hkietóopa
9	shá^{n}ka	lébrą hce wįįke
10	glébla	lébrą

Table 2: Additional Kanza lexicon (Haldeman 1860: §634)

	Kanza
ear	na^{n}tá
eye	ishtá (note ishtá toho 'iris' and ishtáka^{n}ha 'eyelid')
brow	ishtáhin
mouth	i (Haldeman's form iha is 'mouth-skin' or 'lips')
tongue	léze
nose	pa
nostril	pa xlóge
forehead	pe
fan	ijéayuzúbe (fan hung over baby's face)
pipe	nannónba
knife	mánhín
warm	moshcé
leggins [sic]	húyuyinge
shirt	ókiloxla

the acid-heavy and aged brown paper of the version of Haldeman 1860 available from archive.org.)

The forms are in general readily identifiable from recording of the languages over a century later, as the references from the Kanza and Osage dictionaries show (Cumberland & Rankin 2012 for Kanza, and Quintero 2009 for Osage). The few differences are instructive.

Most interesting in this regard are the numerals, especially 'nine' and 'ten'. In Osage 'nine' is a subtractive compound ('ten lacking one') involving 'ten' and an allomorph of 'one'. But Kanza uses the widespread form, possibly reconstructible as *kišąkka*, which is recorded for several Mississippi Valley and Ohio Valley Siouan, Muskogean and Great Lakes Algonquian languages. Modern Osage has simplified the onset of 'ten', though Haldeman had what would nowadays be represented as /kar-/ (or maybe /gar-/; his depiction of voicing is not always trustworthy). The form in Dhegiha has irregular reflexes elsewhere in Dhegiha: Omaha-Ponca *gthéba*,[1] where <th> is /ð/, has lost the liquid found in the second syllable in other Dhegiha languages and in earlier records of Omaha-Ponca (compare Quapaw *kdébnǫ*; Rankin 1982: 3). The glide which separates the prefix from the form for 'three' in Kanza 'eight' has been apprehended by Haldeman as a front vowel, although the hiatus in the corresponding Osage form has been recognized by Haldeman as such.

Both the Kanza and Osage forms in Haldeman's work include forms of what was originally the enclitic *-xci* 'only' at the end of the form for ONE, and this pan-Dhegiha word is a form which was later borrowed into Caddo as *wists'i'*. Note also the initial [di-] in Kanza 'three', now replaced by /j-/ <y->, and the fact that Haldeman did not notice the nasalization of the vowel in the first syllable of Kanza 'two'. The form 'eight' in Haldeman's Osage is reflected in the modern language, in modern Kanza and (as a loan, namely *kiyátaw*; Rood 1996:608) in modern Wichita. But the earlier form for 'eight' based on 'three' is used in Haldeman's Kanza as a parallel to the form for 'seven' (itself a compound involving 'two'). Primary stress and vowel length and nasalization are well represented in Haldeman's work, especially for Osage.

Of the nouns in Haldeman's record of Kanza, most are similar to their modern counterparts. For the rest, if one allows for a modicum of close phonetic detail (for instance the realization of /a/ as a low rounded vowel in 'nose'), the quality of transcription is rather high. 'Eyebrow' may end in a form of *sábe* 'black' but this is uncertain, while material which is less easy to identify is attached to the end of 'eye' and 'shirt'. The occasional weakness in Haldeman's powers of perception

[1] http://omahalanguage.unl.edu/dictionary/

is seen in the fact that the consonantal sounds in the second part of 'nostril' are represented in Haldeman's work by his symbol for /ʃ/, while the initial consonant of 'tongue' has changed in the 120 or so years between Haldeman's work and Bob Rankin's. 'Warm' seems to include an enclitic, which may be the masculine declarative enclitic *(ey)ao*. Haldeman's remarks about the phonetics of 'nose' and 'tongue' are somewhat surprising, as modern Kanza does not permit [h] in coda position and does not use geminate consonants.

5 Conclusion

Data on Haldeman's recording of Dhegiha languages have been presented and the success or otherwise of Haldeman's system in coping with the segmental phonology of these languages, especially the complex consonantal systems, has been evaluated. Haldeman's ability correctly to hear the phonetic features of a language seems to have varied in competence from one language to another. Although his Kanza and Osage data are the most accurately recorded Dhegiha data of their time (and although very little else was available for Kanza when Haldeman's data appeared), his transcription is still far from adequate. This is possibly the result of his imperception of certain sounds. Nonetheless the transcriptions list some forms which differ in phonological shape or sense from the modern forms of these words, and as such they have some historical significance.

Acknowledgment

I first heard of Robert L. Rankin's Siouan work in the 1980s and grew to know it better in the course of the following decade, thanks to his generosity and that of Dave Costa, which I eagerly acknowledge. Bob's example of someone switching linguistic fields and taking the opportunity to document languages before they passed completely out of existence (which he accomplished at a time before language endangerment was regarded as an important concern) impressed me greatly, as did his use of philological sources. Having presented a paper at the Siouan-Caddoan Conference I was lucky enough to have dinner with him and Giulia Oliverio at the SSILA Meeting in Albuquerque in 1995, where he was kindly, erudite and hilarious company. The study of Native American language is so much the poorer for his passing.

References

Cumberland, Linda A. & Robert L. Rankin. 2012. *Kaá^{n}ze íe wayáje: An annotated dictionary of Kaw (Kanza)*. Kaw City, OK: Kaw Nation.

García, Bartolomé. 1760. *Manual para administrar los santos sacramentos de penitencia, eucharistia, extremaunción, y matrimonio, dar gracias después de comulgar, y ayudar a bien morir a los indios de las naciones pajalates, orejones, pacaos, pacóas, tilijayas, alasapas, pausanes y otras muchas diferentes, que se hallan en las Missiones del Río de San Antonio, y Río Grande, pertenecientes a el Colegio de la Santíssima Cruz de la Ciudad de Querétaro, como son: Los pacuaches, mescales, pampopas, tacames, chayopines, venados, pamaques, y toda la juventud de pihuiques, borrados, sanipaos, y manos de perro*. Mexico City: Imprenta de los herederos de Doña María de Rivera.

Haldeman, Samuel Stehman. 1859. Analytic orthography: An investigation of the sounds of the voice, and their alphabetic notation. *Transactions of the American Philosophical Society* 11. 259–401.

Haldeman, Samuel Stehman. 1860. *Analytic orthography*. Philadelphia: J. B. Lippincott & Company.

Haldeman, Samuel Stehman. 1872. *Pennsylvania Dutch, a dialect of South German with an infusion of English*. London: Nicholas Trübner & Company.

Jones, Daniel. 1909. *The pronunciation of English*. Cambridge: Cambridge University Press.

Kelly, John & John Local. 1989. *Doing phonology: Observing, recording, interpreting*. Mancester: Manchester University Press.

Lepsius, Carl Richard. 1863. *Standard alphabet for reducing unwritten languages and foreign graphic systems to a uniform orthography in European letters*. 2nd edn. London: Williams & Norgate.

Lesley, J. Peter. 1886. Memoir of Samuel Stedman Haldeman, 1812–1880. *Biographical Memoirs of the National Academy of Sciences* 2. 139–172.

Pilling, James Constantine. 1887. *Bibliography of the Siouan languages* (Bureau of American Ethnology Bulletin 5). Washington: Government Printing Office.

Quintero, Carolyn. 2009. *Osage dictionary*. Norman: University of Oklahoma Press.

Rankin, Robert L. 1982. A Quapaw vocabulary. *Kansas Working Papers in Linguistics* 7. 125–152.

Rankin, Robert L. 1994. Notes on the Kansa word list of Maximilian, Prince of Wied. *Kansas Working Papers in Linguistics* 19(2). 201–210.

Riggs, Mary Ann Clark. 1852. *An English and Dakota vocabulary*. New York: R. Craighead.

Rood, David S. 1975. *Wichita grammar*. New York: Garland Publishing.

Rood, David S. 1996. Sketch of Wichita, a Caddoan language. In William C. Sturtevant & Ives Goddard (eds.), *Languages* (Handbook of North American Indians 17), 580–608. Washington: Smithsonian Institution.

zu Wied-Neuwied, Maximilian Alexander Philipp. 1839–1841. *Reise in das innere Nord-America in den Jahren 1832 bis 1834: 2 Textbände und 1 Bildband mit Illustrationen von Karl Bodmer*. Koblenz: Verlag J. Hölscher.

Part II

Applied and descriptive linguistics

Introduction to Part II

Several Siouan languages are without native speakers – "extinct" or "sleeping" – and the rest are endangered. Because of the situation, documentation, retention, revival, reclamation and revitalization are crucial – both for Native communities who want to use their languages, and for linguists seeking to record usage patterns, establish details of historical change or investigate the limits of human grammatical structures. Non-Native linguists, community members and descendants have tended to approach and answer language description and retention questions in different ways, with different interests. The chapters in this part of the volume illustrate a variety of approaches to language documentation, pedagogy, and resource enhancement.

Linda Cumberland ("In his own words: Robert Rankin recalls his work with the Kaw people and their language") interviews Rankin about his long association with the Kaw Language Program, his early fieldwork experiences and the importance of language documentation to ensure that future generations have the ability to reconstruct and revive their language even after intergenerational transmission fails. This chapter highlights Bob Rankin's contributions to applied and descriptive linguistics; for instance Cumberland & Rankin (2012), Rankin (1989). Kaw people's appreciation is evident in the inscription on his tombstone: "Wíblaha^{n} Ká^{n}ze íe shó^{n}sho^{n} ni." ('Thanks to you the Kanza language lives on.')

Jimm Goodtracks, Bryan James Gordon, and Saul Schwartz ("Perspectives on Chiwere revitalization") present a three-pronged, personal account of their individual and collective involvements with the Chiwere (also known as Báxoje Jiwére Ñút^achi, or Ioway and Otoe-Missouria) language. In separate, individually-written sections, each author gives a unique view of past, present, and future prospects of Chiwere, drawn from their various roles as Elder, teacher, dictionary-compiler, language-nest participant, researcher, and language-program director.

Justin T. McBride ("Reconstructing post-verbal negation in Kansa: A pedagogical problem") addresses a common issue in the teaching of a language which no longer has an active speech community: that of gaps in the recorded data. As the numbers of fluent speakers decline, Siouan languages are increasingly "housed" in the inevitably incomplete descriptions of linguists and in non-fluent

tribal and descendant communities. McBride uses a variety of tools, including historical data, syntactic theory, and consultation with speakers of closely related languages to reconstruct an appropriate negative conditional form ('wouldn't') in Kansa (Kaw). The combination of formal syntactic argumentation and pedagogical purpose is unusual but productive.

Jill D. Greer ("Baxoje-Jiwere grammar Sketch") provides an overview of the grammar of Baxoje-Jiwere (Chiwere). Although necessarily very short, it covers the basic facts of the language, from phonology through morphology, syntax, word coinage and variation indexing gender and dialect. As is typical for a grammar of a Siouan language, the verb receives the most attention, with its multitudinous prefixes, suffixes and clitics.

References

Cumberland, Linda A. & Robert L. Rankin. 2012. *Kaá^{n}ze íe wayáje: An annotated dictionary of Kaw (Kanza)*. Kaw City, OK: Kaw Nation.

Rankin, Robert L. 1989. Kansa. In Takashi Kamei, Rokuro Kono & Chino Eiichi (eds.). Trans. by Akira Yamamoto, *Languages of the world* (The Sanseido Encyclopaedia of Linguistics 2), 301–307. Tokyo: Sanseido Press.

Chapter 6

In his own words: Robert Rankin recalls his work with the Kaw people and their language

Linda Cumberland

In this edited transcript of a 2011 interview, Robert Rankin discusses his early training in linguistics, his first contacts with the Kaw people and language, and his subsequent lifelong involvement with the Kaw Language Project.

Robert Rankin was fundamental to the development of the Kaw Language Project, an office maintained by the Kaw Nation at its headquarters in Kaw City, Oklahoma. Under Bob's supervision and with his tireless assistance, the KLP has produced an array of language materials, including teaching materials, a volume of KanzaKaw language texts (*Kaa^{n}ze Weyaje — Kanza Reader*; Kanza Language Project 2010) and a dictionary (*Kaa^{n}ze Ie Wayaje: An Annotated Dictionary of Kaw (Kanza)*; Cumberland & Rankin 2012). In the early 2000s, he and then Language Director Justin McBride worked together to assemble a comprehensive collection of all Kaw language data known to exist and archived it at tribal headquarters. His work with the Kaw people and their native language extended over four decades and produced the only available sound recordings of the language, collected in the 1970s, from the last native speakers of the language. In December 2011, I, as Language Director at that time, sat down with Bob to ask him to recall those early days of his work and the speakers he worked with. This is an edited version of that conversation.

Linda: It's December first, 2011. I'm in the Kaw Nation Language Office with Dr. Robert Rankin to talk about his field experience in the 1970s, recording the last first-language native speakers of the Kaw language. So, Dr. Rankin — Bob —, why don't I just ask you to tell in your own words how you started doing fieldwork down here and about your fieldwork.

Bob: Well, it's hard to know exactly where to start. I was trained in European languages in my college and university work, and I stayed with European

Linda Cumberland. 2016. In his own words: Robert Rankin recalls his work with the Kaw people and their language. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 119–132. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.169 DOI:10.17169/langsci.b94.169

languages right on down through my doctoral dissertation at the University of Chicago. I did two years of work recording Romanian language dialects in Romania — which at that time was a Communist country — under a Fulbright Student Grant in the mid-1960s and had gotten the job at the University of Kansas on the basis of my linguistic training and my knowledge of those European languages. But I had a colleague at the University of Kansas, Dale Nicklas, who had worked most of his life with the Choctaw people in southeastern Oklahoma, and he kept reminding me that it was the duty of every American linguist to try to document at least one Native American language.

I'd always kind of ignored him, but he finally got to me, so I'd begun work in northeastern Oklahoma trying to find people who could still speak Quapaw in the summer of 1972. I was really unable to find more than one or two people who remembered much Quapaw. I familiarized myself with the literature on the Siouan language family, the family of languages related to Dakota Sioux, of which Kaw was one, and I had determined that after Quapaw, the Kaw language was the least well documented in the archival materials available in the Smithsonian and other archives.

And so, about mid-1973, I thought that I would try to find speakers of Kaw, or as I called it at that time, Kansa.[1] Of course, one of the things you have to realize is that it was hard to find people who could talk about a project like this. I had no idea who would know the Kaw language, how many people spoke it, where they were, what their names were, where they were located. This was all sort of a mystery. I remember asking the woman behind the desk at the Osage Museum over in Pawhuska if she knew anybody who could still speak the Kansa language and she gave me the names of three or four families. And so I came over here to see what I could find out.

The gentleman I located lived in the village of Washunga, which at that time was not yet flooded by the Army Corps of Engineers [to create Kaw Dam and Kaw Lake]. I think it was Clyde Monroe. I went to his home. He was bedridden at that time. His family permitted me to talk with him for a few minutes, and I told him the kind of project I was interested in. He said I should talk with Walter Kekahbah or Maude Rowe, Ralph Pepper or Tom Conn. He may have named one or two other individuals, maybe one or more of the Mehojahs. So I did that. He told me that Walter Kekahbah lived in the nursing home up in Newkirk [Oklahoma]. I went up there and I was able to interview him. He claimed that he had forgotten

[1] There are no hard and fast rules for the use of the terms *Kansa* (or *Kanza*) and *Kaw*. Both are used interchangeably in reference to the language, while the people generally refer to themselves as *Kaw*. —LC

most of his Kaw but, really, he hadn't. He made some recordings for me, some brief recordings, and they turned out to match very well with the Kaw that was documented in the late 19th century. So, he remembered a lot more Kaw than he thought he did. But he told me that I really should talk with Maude Rowe, that she had spoken [Kaw] more recently than he had.

I took his advice and met with Maude Rowe's son, Elmer Clark. Elmer invited me over that evening for some watermelon. He said his mom was going to be there, so I got together with her. She agreed to make recordings for me and so I set up a schedule to visit her weekday afternoons at 2:00 where she was living in Pawhuska. So we did that for part of, uh, it would have been the summer of '73, the latter part of the summer of '73.

Well, I don't usually talk about this sort of thing with people, but I was really kind of scared of this whole project. I guess I'm a sort of timid sort of guy and I… it… I don't know. I felt like I was intruding or, uh — I don't know. But anyway, I got over my fear, my reticence, and started making recordings with Maude Rowe. She moved up to Shidler, Oklahoma shortly thereafter. I think she moved the next year. And so I would meet her there.

We would meet for about two hours, from about 2:00 p.m. to 4:00 p.m. every weekday while I was down here. I would come down for about two weeks at a time and then I would go back up to Lawrence [Kansas] and do all the other things that had piled up, waiting for me to attend to them, and then I'd make another trip and we would go over more material. At that time, I acquired a lot of copies of archival materials stored in the Smithsonian Institution, by 19th century amateur linguist James Owen Dorsey. He had worked recording stories and vocabulary in the Kaw language in the late 1880s, early 1890s. So I acquired copies of the various Kaw traditional stories that he had recorded back then and I went over each one of these, word by word, with Mrs. Rowe. She would conjugate the verbs for me and explain portions of the stories that I couldn't understand. So she was a big help all along, and she provided more vocabulary. So we did that on and off in summer vacations. Sometimes I'd come down in the Christmas holidays, the Easter holidays, for a couple days and we'd do more recording.[2]

She suggested that I talk to Ralph Pepper, in Tulsa, so I got in touch with one of Ralph Pepper's daughters, Hattie Lou Pepper, and she introduced me to her dad. Her dad was very hard of hearing when I met him, and to record with Mr.

[2] All of Rankin's recordings with Mrs. Rowe during these sessions were sound recordings, nearly all of which he later transcribed in his field notebooks. Copies of these sounds recordings and notebooks are archived at the Kaw Nation. James Owen Dorsey's recordings, of course, are written only, although Rankin did make sound recordings of Mrs. Rowe repeating or elaborating on some sections of Dorsey's work.

Pepper, I had to basically write out my questions ahead of time so he could read them because he had trouble hearing my voice. He could always understand his daughter very easily but when I would try to talk to him, he had problems hearing, so he would respond to written questions and translate material. Occasionally he would include a little prayer. This would have been already in the mid-1970s, I think. So we continued. I did some recording with Mr. Pepper, although his hearing loss made it difficult. And I did quite a lot of recording with Maude Rowe. Oh, Levi Choteau — he was the other speaker that she mentioned. I had talked to him. I had gone down to Reno, Oklahoma and talked briefly with him, but he was bedridden at the time and wasn't able to do much recording for me.

So, that basically went on. I would come down during vacations, whenever I had some spare time, and do recordings [with Maude Rowe]. This lasted until I started having back problems, sort of in the late 70s. I had to have some surgery then and so that sort of broke things off. Mrs. Rowe passed away in 1978 and Ralph Pepper, in the early 1980s. So at that point, I think, that was pretty much the last of the individuals who could speak fluent Kaw. That's about the story [of work with Kaw speakers].

Linda: What were your impressions of Mrs. Rowe, her personality? What was a typical session like? Just reminisce about Mrs. Rowe a bit.

Bob: Well, that's always a pleasure. She was a delightful woman. Of course at that time I was this 30-year-old and she was a woman in her 70s, I guess, so again I guess I felt a little reticent, a little timid, you know, but she was always a pleasure to work with, to record with. She had a wonderful sense of humor. I would generally spend the mornings at my motel room, going over the material that I was going to ask her about that day. I'd prepare in the mornings, and then I'd go get lunch and to Shidler and record with her a couple of hours, going over the materials I had prepared.

One morning I had gotten lazy and — I don't know what I did, overslept or ate a long breakfast, or something. Anyway, I was not prepared. So I went up with my tape recorder and I opened my notebook of James Owen Dorsey traditional tales that he had tried to write down in the 1800s. So, what I would do is read what he had written down out loud and then I'd get her to tell me what it meant to her, and verify with her that Dorsey had gotten the meaning right and had gotten the pronunciation right. I'd get her to pronounce the sentences, too. And so I was reading along — it was the story of the two raccoons — and I got to one point that I hadn't noticed because I hadn't prepared that day, where underneath the sentences written in the Kaw language, where normally Dorsey would write the

English translation, in this particular instance he had written Latin, not English. So he'd written these Latin words there, and I hadn't bothered to check. So I got to this sentence and read it out loud. And Mrs. Rowe just sat there for a minute. And then she kind of made this funny noise. She kind of went, "Pff, pff, pff," you know, through her lips, and then she just burst out laughing, laughing as hard as she could. I couldn't figure out what the deal was, and then I looked to see the translation and I saw the Latin — and I can read Latin — and it was, uh, shall we say, very questionable language, really ripe. Raccoon had done something really naughty and then had done it again. And I had read this right out loud to Mrs. Rowe! And if she hadn't had a terrific sense of humor, that probably would have been the end of my fieldwork, right that day. I don't think either one of us would have gotten over it. I was so embarrassed! I had used language with her in her native language that I would hesitate to use to my own mom or grandma — maybe even my own wife. It was pretty raw language, but she, heh, just laughed like crazy. We got together with Elmer Clark that evening and we were talking about the day's recording and she said, "Oh, we had a humdinger! It was a humdinger!" [Bob laughs.] And it was.

Linda: Wasn't it [the use of Latin] sort of a product of the times when Dorsey did that? It was my understanding that there were censorship laws that kept them from writing certain things in English, so they wrote it in Latin to avoid putting stuff in print that would have violated [American] censorship laws.

Bob: That's entirely possible. I am not aware of particular laws, but it wouldn't surprise me at all because those were the days when Hollywood kisses couldn't last longer than three seconds and if you were sitting on a bed you had to have at least one foot on the floor. Heh! And that was just movies. So, yeah, I suppose if those stories were going to be published, probably they wouldn't have wanted to write those words in English.

Linda: That was the era of "banned in Boston."

Bob: Yes, absolutely. But you're right. I suspect Dorsey was under some constraints to write the translations of blue vocabulary and used Latin. [3]

[3] A law and subsequent judgment that had been put in place just as Dorsey was doing his fieldwork on Dhegiha languages would have led Dorsey and other ethnographers to use Latin in place of English when translating anything that might have been construed as obscene. First, the federal Comstock Act of 1873, "which grants the post office the power to censor mail containing material that is 'obscene, lewd, and/or lascivious' … The law and its state-level counterparts are initially used primarily to target information on contraception rather than pornography, but soon become the primary vehicles for obscenity-related

Linda: What was Ralph Pepper like?

Bob: Like I said, he was very hard of hearing and at the time that I was able to speak with him, he had also been in an accident. He had been hit by a truck, so he was bedridden. But he was eager to speak the Kaw language. He was happy to record for me. It was just a matter of interacting with him, not being able to speak to him directly and have him understand most of what I said. I really had to write everything out, so it was kind of laborious for him, I think. He would sit there while I scribbled things out and showed him and he would say something. By that time I had done quite a lot of recording with Mrs. Rowe, so I could understand what he said. So that worked out really well. It was very important to get the pronunciation and the sentence grammar that male speakers used because the Siouan languages, unlike the English language, have one grammar for men and a different grammar for women. They actually conjugate verbs a little differently. So it's really important to get both male and female speakers of Kaw, or Quapaw, or Omaha, or Osage any of those languages.

Linda: Was the work with Walter Kekahbah similarly laborious?

Bob: It was difficult for me. I only met with him once or twice. I was just beginning then, so I didn't have much of any notion what the Kaw language was like. He was the first Kaw speaker I met. At the time, when he said that he had forgotten all his Kaw, I took him at his word. But after I had gotten to know something of the language after working with Mrs. Rowe, I went back and listened to the recordings of Walter Kekahbah. They were quite fluent and had a lot of the male speech features that I was looking for in Ralph Pepper's speech, also.

Linda: What was the attitude of these Kaw speakers when they knew what the project was: that you wanted to make recordings and that you were going to document the language? Were they actively supportive or were they reluctant? Sometimes speakers in that generation were reluctant to share their language.

---

prosecutions. The Comstock Act allowed warrantless searches of the mail for 'obscene' materials." [http://civilliberty.about.com/od/freespeech/tp/History-of-Censorship.htm] The judgment came in 1897. "In *Dunlop v. United States*, the Supreme Court rules that print literature — including but not limited to risqué fiction and information on birth control — may be considered obscene under the Comstock law. The *Dunlop* standard is still current law, based on the 2005 prosecution of several Internet sites specializing in written online erotica graphically depicting illegal acts. As Chief Justice Warren Burger wrote in *Kaplan v. California* (1975), "[o]bscene material in book form is not entitled to any First Amendment protection merely because it has no pictorial content." [ibid.]

Bob: I didn't find a reluctance to share. Everybody that I approached was happy to do what they could. They were interested in seeing the language documented, for posterity, you know, for their grandkids and great grandkids. I think in the beginning people didn't know what to expect from me, and for good reason. I didn't know what to expect from *myself* at that time, much less what to expect from them. This was just something that sort of grew, as we worked together. With Mrs. Rowe, I recorded enough with her over the period of several years that she had gotten interested in it, herself, by the time the project began to wind down and she seemed always very eager to help out.

Linda: "…preserving this for kids and grandkids…" How many members of these people's families did you meet?

Bob: I met a lot of Maude Rowe's relatives. I met a couple of Walter Kekahbah's relatives — I think his nephew and his wife. I'd met Ralph Pepper's daughter. I'd met some of the other family, some of his nephews or great nephews and so forth. But Mrs. Rowe's family, goodness! I've met her son, his daughter, *her* son and daughters, and one of *her* daughters has kids now, so I guess I've known four generations or so of that family. Ah — Stormy and Dewey [Storm Brave and Dewey Donelson].

Linda: They keep winning prizes at [the annual Native American Youth] Language Fair. That's a really strong line of language persistence through that family.

Bob: Mm-hm, mm-hm.

Linda: So, once the project wound down with the data collection and you were back at the [KU] campus, what did you do with the data? Obviously, you've stayed engaged with [the Kaw Nation] for decades, but what did you do with it first?

Bob: Well, the first thing you need to do, of course, is transcribe all those tapes. Most of my recordings were on old tape recorders, reel-to-reel tape machines, so that had to be written down, and I got through most of it. I may still have a little bit that still needs to be written down, even after all these years because, you know, you get busy doing other things. I had a teaching position nine months out of the year and it's almost impossible to do a lot of tape transcription during that period. Things just get in the way of writing it down and analyzing the data, you know, and papers for various linguistic conferences…. Then I set about compiling a dictionary. I bought a computer — bought my first computer in 1984 — and began typing all the vocabulary into a dictionary, which is now pretty much complete. And I went through the various stories, trying to figure out

different points of grammar, how the grammar fit together in sentences, and all that sort of thing. It's very time consuming. It's what I enjoy doing, so I can't complain. I do a lot of comparative research, too, comparing material from the Kaw language to material from related, sister languages. So that's another thing that occupies quite a lot of my time.

Linda: I'm curious to know how you linked up again with the Kaw tribe directly. Was that because Justin McBride got in touch with you? How did your very productive collaboration with Justin come about?

Bob: Actually, it began somewhat before that. The Kaw Nation decided at some point to have a language program. I can't remember precisely what year this was. It would have been the 80s or early 90s? I'm not quite sure. Anyway, I got a phone call from Lonnie Burnett. Lonnie was, I think, the first Language Coordinator. He came up to Lawrence and brought a video-cam with him. He wanted me to record the alphabet and some sample vocabulary, some sample words containing the various letters of the Kaw alphabet. We did that, and I think the tape still exists somewhere.

Linda: [pointing] Right in that drawer, right there [in the Language Director's office].

Bob: Well, it was kind of an amateur production on my part, I'm afraid, but it was ok. He did that and then at some point, let's see, Kelly Test was involved — Test or Estes, I don't remember which name she was using. She was quite young at the time. She was involved with the program and then at some point Justin McBride took over the program.

But actually, before Lonnie Burnett started up the language program here, one morning, I was sitting in my office at KU, there suddenly appears in my door this group of people: Little Carol [Clark] and Jim Pepper, Henry and Curtis Kekahbah. Oh, and Johnny McCauley. They just appeared in my office door. I recognized Curtis, and I recognized Little Carol, so, heck, I invited them in — Old Home Week! — and they said, "We know you have these recordings of the Kaw language, and we wonder if you'd like to share them with the Tribe — if you'd make copies of them." That sounded like a fine idea to me and so I said, "Sure." They arranged to have the CDs made from my copies of my tape recordings at a television studio, a professional sound studio in Wichita. So I took all my tapes, all fifty of 'em, or however many there were — 'cause I don't remember how many of them there are — at least fifty *hours*, anyway — and I drove 'em down to Wichita and we met down there, and they copied the tapes and then mastered several

copies — one for me, one for the tribe, one for Elmer Clark's family, maybe one or two others for archival purposes.

Basically, that provided the Tribe with copies of everything that I had done, and that served as the database for Justin's work with the language. He came to the tribe with a BA in Linguistics from OU [University of Oklahoma], and a dedication to seeing the language preserved and documented, and so he went right to work here, and the rest is history.

Linda: Yes. He invented the board game[4] and certainly mapped out the structure for the language CD, which I then completed when I joined the language department. I came in January of 2006. Justin had already been there three years, possibly four.

Bob: I would think so. I would think at least four. I remember he used to be in the tribal headquarters across the street, with the scorpions — heh! I think he and Kelly had been there at least three years. And then they came over here [to the Maude McCauley Clark Rowe Social Services Building].

Linda: Justin hosted the annual Siouan Language Conference here in 2005.

Bob: Ok, well, you can count backwards from there. I think my recordings for Lonnie Burnett had been a year or two before Justin came, so you can count backwards from there and figure out when I had done my initial recordings for him. The delegation to Lawrence that appeared in my office would have been a few months to a year before that, I think. We were just getting the CDs available when he began his work, I think.

Linda: Now, Justin left just a little over a year ago. In fact, he earned a Master's while he was here, while he was working full time, but then he left to begin work on a PhD. I think he had been here 9 years when he left. So he's left quite a legacy.

Bob: Yes.

Linda: And he overlapped with me — I'd been here five years when he left.

Bob: So we probably did the conversion of the tapes to CDs in the mid-90s. I think that's about right. It's amazing now, that the entire stack of tape recordings that I made in the 70s, at slooow tape speed to preserve tape, all fits on one CD now.

[4] The board game, *Wajípha^{n}yi^{n}*, Camp Crier, is a Kaw language learning game that also introduces players to the clan structure of the Kaw tribe. The box version might still be available through the Kaw Nation museum store; an electronic version is included in the Kaw language CD set, available through the Kaw language web site, WebKanza, http://www.kawnation.com/langhome.html.

Linda: You know, it's worth noting that if you hadn't made those recordings in the 70s there would be no recordings of the Kaw language, as far as we know. There might have been some informal recordings, but yours are the only known record of how the language actually sounded. It's extremely important for the history of the Kaw language and the documentation, linguistically, just knowing what the language sounded like from fluent, first-language speakers of it.

Bob: Let me add to that, since you brought it up. Tape recorders became pretty popular in the United States already in the mid-50s and so tons of people had tape recorders of various kinds in the 60s and 70s, certainly, and it may well be that in people's closets or attics or old dusty bookshelves, maybe they made some casual recordings of their grandparents—Grandma, Granddad, uncles — speaking the Kaw language. People, families, may not even remember they have those, but if they do, those recordings, if anyone has any casual recordings, no matter how short, of their elders speaking in the Kaw language, those would be extremely valuable to the language program. If they would be willing to make those available for copies, I would be happy to go through them and make sure that they get translated so that people can follow them.[5]

Linda: Yes. So, the interesting thing, when I hear this whole history, from the moment when you just asked the question about speakers who still spoke the language to the present day, many people have come and gone, but *you* are the constant. You started it; you've been with it every step of the way. You've trained any number of people in this office to work with it; you've seen it through the three essential elements of language documentation: a dictionary, a grammar (your grammatical sketch [Rankin 1989]), and texts. You are the constant throughout all work on Kaw language [in recent times]. So I know that the Tribe is extraordinarily grateful to you and you certainly deserve this sort of aura of being able to walk on water. [Both laugh.]

Bob: I just consider myself lucky to have been able to get paid for doing what I really enjoy doing in life. It's a big thing in life, you know, to have a job that you really enjoy, and this is what I like.

Linda: So now that you're retired, what kind of plans do you have — partly for Kaw language, but what else are you working on?

Bob: Well, I'm still doing a lot of work — comparative vocabulary, that sort of thing. But I think that, especially for the Kaw language, we have a dictionary

[5] Sadly, of course, such recordings would not have the benefit of Bob's knowledge today, but it would still be of great value for such recordings to be added to the archive.

that's nearing completion, we have the Reader, the set of Kaw texts, that represent most of the traditional stories in connected texts that we've inherited from our predecessors. The main thing that's left now is a written grammar, so that's something that I think we all hope we'll be able to work on.

Linda: Did you continue with your work on European languages at all?

Bob: Not really. Actually, when I got into the Kaw language and really enjoying my work with the Kaw people, I never looked back. I left my European studies back in about 1975 and just pretty much came over to what I do now.

Linda: Now, you've had several students who have picked up on certain aspects of Siouan research. Who has done that? Have any of them worked on Kaw?

Bob: Mm, no, not really. There was one young lady whose help I enlisted at KU who analyzed some of the recordings to check the status of long and short vowels, that kind of thing. She did a brief project on that, but she didn't follow up on it. She wound up doing something else, instead. One of my students, Sara Trechter, has taken up the study of Mandan, which is up in North Dakota, and she's done a lot of work with the last living speaker of Mandan. Another of my students, Giulia Oliverio, got interested in a Siouan language that was originally spoken in Virginia in the 1700s — the 18th and 19th centuries, and has been extinct now for quite a while. She collected all the archival material that had ever been done on the Tutelo language and she drew from it as a doctoral dissertation — a lexicon and a dictionary. I've had a couple of students who have followed up. And then Dave Kaufman has been working on collecting the Biloxi material and the Biloxi dictionary.[6]

Linda: So, all these are Siouan languages.

Bob: These are all Siouan languages. Tutelo was originally spoken in Virginia — Virginia and West Virginia. Biloxi was spoken in Louisiana, and Mandan is up in North Dakota, so the speakers of the original Siouan languages got around quite a bit. They migrated over much of North America.

Linda: Just to close out, for the record: Siouan languages — just informally, what makes them so closely related?

[6] David Kaufman served briefly as Kaw Language Director from August 2013 to December 2014. At this writing (early 2015), Justin McBride — now Dr. Justin McBride — has stepped in to re-assemble the Kaw language archives following a physical move to a new location.

Bob: The original, single Siouan language was probably spoken back in Kentucky — or back East, and the speakers broke off in groups and moved westward. And, well, one of the groups became the Mandan, one group became the Crows, one of the groups became the Dakota Sioux; a different group became the Ioway, Otoe, and Winnebago [Ho-Chunk] tribes, and then there was an additional group that split up into five sister tribes: the Kaw, the Osage, the Quapaw, the Omahas, and the Poncas, languages that are still pretty similar to each other after several hundred or a thousand years. And to a certain extent, they can still understand each other somewhat, whereas if they had to meet up with someone who spoke Dakota Sioux, Mandan, Biloxi, or one of those other languages, they wouldn't be able to understand each other.

The Kaws, the people who could speak fluent Kaw, could understand what Osage people said in their own language. Mrs. Rowe used to get her beadwork supplies from Mrs. Robinson, over in Pawhuska, who sold the beads for beadwork. Now, Mrs. Robinson was an Osage speaker and Maude Rowe could go in and speak to her in Kaw and she would speak back in Osage and they would understand each other. So those two languages are quite close. And the speakers of those languages must have diverged from what must have been a single tribe relatively recently. Earlier divisions would have been the Omahas, the Poncas, and the Quapaws. So those five tribes are closely related; their languages are very similar; they have similar vocabularies, whereas the speakers of the other language — the Sioux, the Crows, the Mandans and so forth — they must have split off at a much, much, much earlier date — several centuries — and their languages are very different now.

Linda: A couple of the sources we're working with, George Morehouse and Addison Stubbs, were involved with the [Kaw] tribe pretty closely and did documentation, each in his own amateur way, and each said that Osage and Kaw are the "same language." Maybe you could just put that notion to rest, officially.

Bob: Well, they're similar enough that they could understand what one another was saying. The pronunciation is certainly different. I've heard spoken Osage, I've worked with a group of Osages back around 1980 devising an alphabet at that time for teaching the language (they've developed their own alphabet since then), so I know something about Osage — it's different [from Kaw].

This is where the existing record of the interview ends. It is unclear how the original recording was lost, but in all of the existing copies of the recording, the final minute is missing. Fortunately, as I recall, it consisted only of our concluding expressions of thanks to each other for the interview. At the conclusion of the interview, we went to lunch and then went back to work on the project at hand,

whatever it was that day. Bob was unfailingly generous with his time and knowledge, and refused to let his health issues deter him from continuing his work on the language. When he could no longer make the trip to Kaw City, I made a series of week-long trips to Lawrence over the following year and a half, where he could more easily work around his regular medical appointments, putting in a full eight hours each day. The pleasure he took in his work on Kaw and comparative Siouan was evident and inspiring to all of us who were privileged to work with him.

References

Cumberland, Linda A. & Robert L. Rankin. 2012. *Kaá^{n}ze íe wayáje: An annotated dictionary of Kaw (Kanza)*. Kaw City, OK: Kaw Nation.

Kanza Language Project. 2010. *Kaá^{n}ze wéyaje – Kanza reader: Learning literacy through historical texts*. Justin T. McBride & Linda A. Cumberland (eds.). Kaw City, OK: Kaw Nation.

Rankin, Robert L. 1989. Kansa. In Takashi Kamei, Rokuro Kono & Chino Eiichi (eds.). Trans. by Akira Yamamoto, *Languages of the world* (The Sanseido Encyclopaedia of Linguistics 2), 301–307. Tokyo: Sanseido Press.

Chapter 7

Perspectives on Chiwere revitalization

Jimm Goodtracks

Bryan James Gordon

Saul Schwartz

This chapter examines the Ioway, Otoe-Missouria Language Project's Chiwere revitalization efforts from three perspectives. Jimm Goodtracks provides an account of how he came to be involved in Chiwere language preservation in the 1960s and of the shifting strategies for documentation and revitalization that he has employed in subsequent decades. Bryan James Gordon presents a list of phrases in Chiwere that he prepared while living in Jimm's household language nest in order to communicate with Jimm's grandson. Saul Schwartz contextualizes the Project's activities by describing their role within the Ioway and Otoe-Missouria communities and comparing them to strategies for revitalizing other Siouan and Native American languages. What emerges is a sense of language revitalization as a social process that involves personal dedication, support from family and friends, collaboration between linguists and community members, and a sense of responsibility to previous and future generations.

1 Introduction

This chapter describes the Ioway, Otoe-Missouria Language Project's recent Chiwere language revitalization activities. Chiwere is a Siouan language with three historically attested dialects – Ioway (*Báxoje*), Otoe (*Jíwére*), and Missouria (*Ñút^achi*)[1] – and is a heritage language for three federally recognized tribes: the

[1] *Ioway*, as a term for an American Indian people and their associated language and culture, is also known as *Iowa*, as in the names of the two federally recognized Iowa Tribes. Similarly, *Otoe* is sometimes spelled <Oto>, and *Missouria* can be called *Missouri*. Except where otherwise indicated by angle brackets (<>), all Chiwere words in this chapter are written with the orthography used by the Ioway, Otoe-Missouria Language Project, which is described on the Project's website (Goodtracks n.d.).

Jimm Goodtracks, Bryan James Gordon & Saul Schwartz. 2016. Perspectives on Chiwere revitalization. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 133–165. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.170 DOI:10.17169/langsci.b94.170

Iowa Tribe of Kansas and Nebraska, the Iowa Tribe of Oklahoma, and the Otoe-Missouria Tribe of Indians, also in Oklahoma. Most authoritative scholarly sources place the last fluent Chiwere speakers in the 1990s (Lewis, Simons & Fennig 2013; Parks & Rankin 2001). Though no fluent speakers have been identified since then, a handful of semispeakers remain, and some Ioways and Otoe-Missourias continue to use Chiwere in certain contexts. Nevertheless, most tribal members have few opportunities to hear Chiwere in their daily lives, and written Chiwere constitutes their primary form of access to the language.

The Ioway, Otoe-Missouria Language Project is a community-based Chiwere documentation and revitalization effort headed by Jimm Goodtracks. Though the Project is not formally affiliated with academic or tribal institutions, it has collaborated with both over the years. The Project's documentary work has been funded by the National Science Foundation's Documenting Endangered Languages program since 2007, first through a grant to the Iowa Tribe of Kansas and Nebraska to prepare a dictionary and then through a grant to the Project itself to support ongoing work on an annotated corpus.[2] The Project makes its dictionary, pedagogical materials, and other resources available online through its website (Goodtracks n.d.).

In addition to these public activities, the Project also supports a more private revitalization effort: Jimm's house is a LANGUAGE NEST, in which he raised his grandson, Sage, to speak Chiwere as his first language. During the course of their graduate studies in anthropology and linguistics, Bryan James Gordon and Saul Schwartz had the privilege of living with Jimm and his grandson. They participated in the Project's documentary and revitalization activities as well as the household language nest.

Our chapter begins with Jimm's account of how he came to be involved in Chiwere language preservation in the 1960s and of the shifting strategies for documentation and revitalization that he has pursued in subsequent decades. Bryan's section presents a list of phrases in Chiwere that he prepared while living in the language nest in order to facilitate communication with Jimm's grandson. Saul's section seeks to contextualize the Project's activities by describing their role within the Ioway and Otoe-Missouria communities and comparing them to strategies for revitalizing other Siouan and Native American languages. Our hope is that by combining our perspectives, a richer and more complex picture of Chiwere revitalization will emerge. As can be seen from what we have written,

[2] Work on the dictionary was funded by NSF award number BCS-0553585 *Ioway Otoe-Missouria Dictionary Project*. Work on the corpus is funded by NSF award number BCS-1160665 *Chiwere (ISO 639-3: iow) Audio Archive Project (CAAP)*.

we are each indebted in our own ways to Bob Rankin, to whose memory this volume is dedicated, for his contributions to our work.

2 The Ioway, Otoe-Missouria Language Project (Jimm)

I left Pawnee, Oklahoma, after my graduation from Oklahoma State University in May 1965 to begin my first genuine employment in social welfare and to live as an independent adult. My work in southern Colorado would initiate my course into linguistics and the committed study of our Elders' Báxoje, Jiwére-Ñút^achi ('Ioway, Otoe-Missouria') Language.

I cannot really say when it was that I realized that in southern Colorado I had arrived in an encapsulated leftover corner of early Spanish colonization. The semi-arid desert was irrigated by a system of antiquated hand dug waterways that nourished small family rancheros that surrounded many aldeas ('villages') off the highways and produced lush green fields for small herds of cattle and sheep and individual gardens for the valley residents. The people lived in flat-roofed homes made of adobe, each having its own dug well, and some homes had archaic hornos (clay ovens still in use by the Pueblo Tribes in New Mexico). Only in the highway towns, like Alamosa, Antonito, La Jara, and Manasa, could one enjoy a city water system. The area residents referred to themselves in English as *Spanish*, but when speaking their antiquated Spanish, they referred to themselves as *chicanos* or *la raza*. They were said to be descendants of los conquistadores of some 400 or so years earlier, who in turn married into the Native Peoples (Utes, Apache, and Pueblos) and settled in the mountain valleys, making them their home, while the indigenous people were driven further and deeper into the mountains that are characteristic of the State of Colorado. The area was rich in radio programs in Spanish, Navaho, Taos Pueblo, etc. I was surrounded by non-English speakers everywhere.

During these same years in Oklahoma, the late 1960s, I recall being around many tribal languages usually spoken by the Elders: Pawnee, Otoe, Ioway, Osage, Ponca, Sauk, Kiowa, Creek, Cherokee, and Seminole were just a few of the languages heard frequently. There were a few monolinguals who required the assistance of a translator for their Native Language. It was common to hear Ponca, Ioway-Otoe, and Osage spoken and sung in prayer services, ceremonies, and conversation at the various community dances, handgames, and prayer meetings – both in traditional Native American Church services held in tipis and in the Indian Methodist and Indian Baptist churches. The languages were always present in ceremonials, as in the Iroshka Society, the AsaKipiriru ('Young Dog') Dances,

and Pipe Blessings. It was only natural that any of us young people would have learned and used some of the Native Language of the Elders, and yet it was a knowledge all too often taken for granted. It was overspun by the lure of "fun" and partying, driving youthful passions, and the need to "fit in" for that special interaction that can only be achieved from one's peer group.

Now, as a novice social caseworker, I had arrived back in time in the immense Valle de Sangre de Cristo. The area stirred my interest in the historical, earlier times of our Oklahoma Grandfathers. I became enthralled with the uncommon culture and language of the Chicanos. A younger brother noted how similar these people were to Native People in Oklahoma. Well, not quite, but close enough.

I quickly learned that young Chicanos of my age that lived in the outlying communities could not be relied upon to speak English or know sufficient English to carry on business. Most of their Elders had never traveled beyond the mountain ranges that surrounded the Valley. I learned to speak Spanish and thought of how I came from a small family of relatives who spoke indigenous language(s) of the Land, just as these People. And then I thought to myself: Why did I not pursue developing my limited knowledge from and with the Elders there?

Thus, I was inspired to write down every word and phrase that I knew and remembered of our Grandmothers' Báxoje, Jiwére-Ñút^achi Language. I arranged during my travels home to Oklahoma to spend an increasing amount of time with our old Uncles and Aunties. I asked for the words and terms for plants, trees, and articles of dance clothing. I became more interested in the personal histories of the old People, of their early days in Oklahoma, the changes they have been obliged to conform to, and their Spirituality, which saw them through their transformation from resident possessors and spiritual keepers of the land to a contemporary disparaged minority which seemingly endlessly needed to bend to the will of the dominant controlling Society. On one occasion, at the Iroshka Society ceremonial dances of the Osage in Gray Horse, Oklahoma, Mama introduced me to an Osage woman, saying to her: "My son is an old man made over." She was referring to my passion to know language, old traditions, teachings, and ways, which contrasted with most adolescents, whose interest was minimal at best.

During the years of 1965–1987, I worked to preserve and research the Ioway, Otoe-Missouria Language, oral tradition, history, and customs. I researched all manner of documentation on the Ioway, Otoe-Missouria culture, oral literature, language, lifeways, and ceremonials made by early explorers, fur traders, and missionaries. At the same time, I was mentored by several Elders in the Ioway, Otoe-Missouria communities and one other up in the Northern Plains.

In 1970–1972, I received a National Mental Health Scholarship to attend graduate school at the University of Kansas School of Social Welfare in Lawrence. It was there that I acquainted myself with the KU Department of Linguistics and met Robert Rankin and Ken Miner. I attended the Siouan and Caddoan Languages Conference and met John Koontz from Colorado. I learned that these three men had done extensive research and fieldwork in the Kansa, Quapaw, Omaha, and Winnebago [Ho-Chunk] communities. Through their personal and professional association, they provided me with much insight into the study of academic linguistics and how it could be useful in community or applied linguistics. They provided assistance for my personal work with the Ioway-Otoe Language and Elders of both communities that I had begun about 1965.

Bob and John were particularly helpful in showing how linguistic analysis of words provides insight into the grammar of the language and the construction and root meanings of words. The utility and exactness of linguistic orthography was pointed out to me. When I began serious work on a corpus dictionary for the language, I would share files with them and seek their advice and recommendations for more complete entries. After I worked out an acceptable standard orthography with my Elder informants and advisors, I intended to include a word sounding out such as was the common practice among community members. One of the community Elders wrote a brief dictionary in the 1970s of words and phrases in this folk manner (Murray 1977). For example, he wrote as in the first column of Table 1:

Table 1: Murray's (1977) orthography compared with Ioway, Otoe-Missouria Language Project orthography.

Murray orthography	English translation	Goodtracks orthography
knee	'water'	ñí
the-gray	'footprint'	thígre
wah-yeeng-eh-the	'yellow bird'	wayíŋe dhí
khoo-me-khan-jay	'big stink'	xúmi xánje

Since the communities had become accustomed to view their language written in such a format, I began to include duplicate dictionary word entries in a similar manner. However, Bob Rankin advised that this double format was inappropriate and a disservice to the people and language per se. He remarked how many of the younger community people had already undertaken the study of other

languages, such as Spanish, French, etc., and at the outset, they were obliged to learn the particular alphabet and phonics of the particular language. Thus it followed that language students need to learn the standard language orthography based on a linguistic model. He spoke to the imperfect, changing nature of using English phonics as a model for unwritten indigenous languages, which results in incorrect pronunciations and enunciations of words and spoken conversation. Also, he pointed out that to write indigenous languages in such a "Morse code" manner was to demean the languages as being something less than the highly developed and complex languages that they are, the equal of any of the contemporary dominant languages that presently exist. Observations among the people in the communities gave truth to his admonition, as there were heard and still persist mispronounced words and simply incorrect phonetics.

In the summer of 1971, while still in graduate school at KU, I engaged the services of a professional linguist at Kansas State University, Lila Wistrand-Robinson, who secured a grant for a study of Ioway, Otoe-Missouria Language with the Elders. She produced a lexicon and a set of basic and intermediate level language study books, which were published in the 1970s and combined my personal research with her own (Otoe and Iowa Language Speakers with Lila Wistrand-Robinson 1977a,b). The materials were distributed gratis to the students and families of the three Ioway, Otoe-Missouria communities. Meanwhile, I continued to work on my own dictionary files.

During July 1985, the Otoe-Missouria Johnson O'Malley Program, after being referred to me by several local Elders, employed me to instruct the children in the summer program on the Báxoje-Jiwére Language. A number of the children had been raised in the presence of their grandparents and were familiar with words and phrases from the language. I utilized word games based on such popular games as Bingo and developed songs designed to teach numbers. I used the tune from the old racist nursery song "One little, two little, three little Indians," and replaced the words with *Iyáŋki, núwe, dáñi nampóiñe* 'One, two, three little fingers'. Then we practiced basic oral conversations.

From that time, various individuals have engaged me in several capacities to assist them to learn the songs, language, and the traditional teachings and lifeways of the Elders (traditional stories, sweat lodge, Native American Church, etc.), in which I was mentored over a period of thirty years. On several occasions, the Otoe-Missouria Tribe asked me to return and provide language classes for the community.

Meanwhile, through the years, I compiled a collection of wéka^{n} (traditional stories) and developed comprehensive dictionary files, which combined my own knowledge of the language with that of tribal Elders and manuscript sources go-

ing back to the 1830s (Goodtracks 1992 – present). In 2002, I was asked to provide an introduction to the Báxoje Language in White Cloud, Kansas, in conjunction with the annual Baxoje Fall Encampment held every September by the Iowa Tribe of Kansas and Nebraska, and I was encouraged afterwards to continue the sessions at several locations, including in Lawrence.

In 2003, Bob Rankin and John Koontz encouraged me to write a grant proposal to develop my dictionary files into an online encyclopedic dictionary that would serve as a resource for the Native communities and the work of professional linguists. Bob and John both wrote letters in support of my proposal. The proposal was awarded a grant, and I commenced to edit my files into full entries for both sections of the bilingual dictionary. And during 2004–2005, I partnered with Marge Schweitzer, a retired Oklahoma State University anthropologist, to transfer the aging, deteriorating cassette recordings of the Elders to a digital format, which provided an opportunity to edit and correct various flaws and errors in the original books for basic and intermediate Ioway-Otoe language study (Otoe and Iowa Language Speakers with Lila Wistrand-Robinson 1977a,b). The cassettes for the first book followed the original material closely but not so the remaining cassettes, which contained previously undocumented materials. Thus all additional words, phrases, and the same contained on these recordings were added to newly written bilingual booklets (Goodtracks 2004a,b).

By this time, I had retired from my social welfare career, and in 2005, I unexpectedly became the legal guardian and adoptive parent of my grandson. I resolved to have a language nest for him as, to date, after all the community classes, programs, and highly vocalized encouragements to "Talk Your Native Language," there were no new speakers to be found anywhere (*new speakers* being defined as individuals who daily communicated in the community heritage language).

So I spoke nothing but Báxoje to my grandson, reading to him and telling stories to him in Báxoje, and when he began to speak, he spoke nothing but Báxoje. It is his first language. My family encouraged and supported my efforts, and they would speak to him as well as assist in his care. In the summer of 2008, Bryan Gordon, a graduate student, volunteered to assist with the work on the dictionary edits. In the course of staying with us that summer, he too learned to speak in Báxoje with my grandson and wrote an interesting and comical list of "Phrases in Báxoje Ich^é Indispensible to Living with a Three-Year-Old". (See the list at the end of §3.)

In 2009, the Siouan and Caddoan Languages Conference was held at the University of Nebraska – Lincoln campus. One of the Ho-Chunk community participants from Wisconsin desired to compare the relative closeness between Báxoje

and Ho-Chunk languages. Thus, he proceeded to say Ho-Chunk words to my grandson. My grandson did not comprehend his exercise with the word comparisons. So when the individual said, "súúch" ('red'), Hi^{n}tágwa ('my grandson') replied: "Hiñego ('no'), šúje," correcting his pronunciation to the Báxoje word. Again, the person said "wooná^{n}ži^{n}" ('shirt'), and my grandson responded, "Hiñego, wónayi^{n}," he corrected him. "Warúch" ('eat something'), the man said. "Hiñego, warúje," my grandson continued to correct him and so on. The Ho-Chunk delegate thought it amusing, as did I.

Following the conference, another graduate student, Saul Schwartz from Ohio, who attended graduate school at Princeton University in New Jersey, volunteered to stay the summer. He assisted me with the dictionary edits and also with the home schooling of my grandson. He had prepared himself before arriving by learning basic conversation in Báxoje, and he continued to hone his capacity to speak and understand the language during that summer. For home-schooling classes, he continued with several materials I had developed and composed, then he added other purchased learning aides, such as lettered blocks, dice, picture books, and innovative books and drawings. He became a "big brother" of sorts, and a constant companion for Grandson, taking him on walks, bike rides, and swimming pool visits. He preferred that I continue to sing the teaching songs, such as "Iyáŋki, núwe, dáñi nampóiñe" ('One, two, three little fingers') and the "ABC Uyá^{n}we" ('ABC Song'). Saul returned the following summer to continue his assistance and take on a new mission, namely field research on Native language study and regeneration.

When Grandson was about four years of age, he had noticed that he was the only child speaking Báxoje, and adults outside the immediate family who he was aware could speak some Báxoje-Jiwére would invariably talk English to him rather than speak whatever they knew in the language. Thus, he acquired basic English, and he began to speak it more often than he would his first language. Nevertheless, for some years afterwards, he would often approach me with new English words he heard spoken and ask: "What do they mean when they say 'diarrhea'? Say it in Báxoje!"

So I would respond: "Tahéda rayéthri ra^ú^{n}na."

"Ohh, ok, pí ke," he would respond upon comprehending the new word. Now, it is not so much that he asks, but still every now and then, he will ask. Further, every day I continue to speak in part in Báxoje, his first language. And we continue to use it as a part of family gatherings and prayer ceremonials, such as in the YúgweChí ('Purification Lodge') and Wanáxi Kigóñe ('Spirit Feasts'). The language does not have the prominent place in the home as it did in Grandson's

preschool years, but it remains a permanent presence in the home. He is now ten years old.

Initially, I had pursued this language interest to satisfy my own knowledge. Later, I desired to have knowledge of the Ioway-Otoe Elders for the benefit of my family and my children. And now I desire to record and share the knowledge I keep as a resource in print for the younger generations that may desire to know about their heritage and language. The Elders are all gone on, and my own generation grows older and fewer. I am now seventy-two. So, it will be good when the taped voices of those traditional Elders have been saved in order to instruct the contemporary and unborn generations. And that will be my final project and contribution to the late Elders, the communities, and those individuals who are inspired to speak the language in the manner of their great-grandparents.

I, Jimm G. Goodtracks, on October 29, 2014.

3 Phrases in *Báxoje Ich^é* Indispensible to Living with a Three-Year Old (Bryan)

In the summer of 2008 Hi^{n}táro ('my friend')[3] Jimm Goodtracks hired me to assist his work on his Báxoje, Jiwére-Ñút^achi, Ma^úŋke ('Ioway, Otoe-Missouria, English') dictionary. He was raising Itágwa ('his grandson') Sage Goodtracks, then three years old, in Báxoje Ich^é. Out of respect for the immersion environment, and out of my love for languages stretching back to when I was Hi^{n}tóšge's ('my nephew's') age, I committed to doing my part. I created this list as a practical guide and a learning tool for myself, checking a lot of it with Hi^{n}táro along the way, but not all of it. It needs many corrections yet. Hi^{n}táro and Saul requested the list for inclusion in this chapter. I had originally posted it on a social-media site popular at the time and it was lost, so it was a relief that it could be turned up again.

What you see in the following lists of phrases at first glance reads more authoritarian than a person ought to be with a child. This is because I usually added to this list during the workday, during which Hi^{n}tóšge ('my nephew') provided me with constant company and friendship, plenty of play breaks, walks, shared meals, and many teaching and learning opportunities – but I also wished to stay productive and offer alternative behaviors often. I used the praising phrases more often than the critical ones. I feel bashful at putting phrases invented by a non-fluent learner out in publication, and disclaim that I learned more about Báxoje

[3] See §4 for further explanation and interpretation of the Báxoje kinship terms employed.

Ich^e from Hi^{n}tóšge's and Itúga^{n}'s ('his grandfather's') corrections than from my faulty practice.

To my recollection, the section breaks here reflect those in the original, but they are not consistently thematic, so I have not titled them. I was probably trying to maintain a sort of thematic structure while adding new phrases on the fly. Instead of reorganizing, I've kept the section breaks where they were.

"Hiñégo, míne wá^u^{n}nachi mínahi^{n} bé re hó." ('No, please leave me sitting alone because I am working.') Phrases in Báxoje Ich^é Indispensible to Living with a Three-Year-Old

Wabúhge hi^{n}ñíŋe ke. 'I'm out of bread.'
Ritúga^{n} igwá^{n}xe re. 'Ask your grandfather.'
Inúha^{n} Ritúga^{n} ahósege škúñi ne. 'Don't talk to your grandfather that way again.' (lit. 'Second time your grandfather don't talk saucily to him.')
Wapópoge škúñi ne. 'Don't throw stuff around.'
Wóji^{n} škúñi ne. 'Don't hit.'
^Ú^{n} škúñi ne. 'Don't do that.'

Ra^ú^{n} ramáñišge rakích^e hñašgu. 'If you keep doing that you might get hurt.'
(Chú^{n}ha^{n}we) Tá^{n}grigi gasú^{n} waná^{n}xi ñíŋeñe ke. Gasú^{n} yá^{n} híwe re. 'Now there are no ghosts outside (the window), so go to sleep now.'
Arábechi ma^{n}grída mína hñe ke. 'Since you threw it up there, it's going to stay up there.'
Gasú^{n} srudhéda asrí^{n} jí re. 'Ok, now go and bring that back here.'

Wá^u^{n} hamína ke. 'I'm working.'
Dagú ra^ú^{n} je? 'What are you doing?'
Dagúre raína (je)? 'What are you eating?'
Dagúrehsji ragú^{n}sda (je)? 'What exactly do you want?'
Dagúre uráje je? 'What are you looking for?'
Wagísdóxi škúñišge srúdhe škúñi hñe ke. 'If you don't ask for it you won't get it.'
Wayére (je)? 'Who is that?'
Dagúra (je)? 'What is it?'
Wé? 'What!?'

__ ta^{n}dáre iwéhi^{n}negi je? 'Where did you put my [horizontal] __?'
__ ta^{n}dáre iwéjehi^{n}negi je? 'Where did you put my [vertical] __?'
__ ta^{n}dáre iwénahi^{n}negi je? 'Where did you put my [round] __?'
Ú^{n}nek^u^{n} hagú^{n}ta ke. 'I want you to give it back.'
Rixráñi je? 'Are you hungry?'
Ridáxra^{n} je? 'Is that too hot for you?'
Ñí sráta^{n} ragú^{n}sda je? 'Do you want to drink water?'
Pagrá^{n}da agúje wórata^{n} hñe ke. 'First you have to put on your shoes.'
Dókirašda^{n}. 'Just a little.'

Iyá^{n}kišda^{n}. 'Just one.'
Hókithresda^{n}. 'Just half a piece.'
(Wabúhgehgu / Warók^i^{n} / núxechebáñi / chebáñiwébri) ragú^{n}sdasge Ritúga^{n} iwá^{n}xe re. 'If you want (cookies / pie / ice cream / cheese) go ask your grandfather.'
Yá^{n} híwe ragú^{n}sda je? 'You want to go to bed?'
Išdá^{n} rich^ésge ke. 'You must be tired.'

Urídus^ada^{n}na ke. 'I'm getting fed up with you.'
Wayí^{n} uríxwáñi ke. 'You're crazy.' (lit. 'Your mind fell down.')
Xáp^a re. 'Quiet down.'
Amína ne. 'Sit down.'
Uyéchi wará je. 'Go to the bathroom.'

Dá. 'I don't know.'
Gasú^{n}da áñi^{n} ke. 'I already have it.'
Dagúre šé asrí^{n} je? 'What do you have there?'
Wáji hadúšda^{n} ke. 'I already ate.'
Waráji iríbra^{n} ke. 'You've already eaten enough.'

Áta ihádušda^{n} ke. 'I already saw that.'
Ra^ú^{n}na aríta ke. 'I already saw you do it.'
Gasú^{n}da ihápahuŋe ke. 'I already know that.'

Pi ra^ú^{n} ke! 'You did good / well!'
É^o! 'Watch out!'

Námañí huhe ke. 'There's a car coming.'
Nabráhge: "Šu^{n}kéñi chéxi ke," ériwana ke. 'That sign says there's a dangerous dog.'
Ná^u^{n} asrúchena náwe ú^{n}k^u^{n} ne. 'Give me your hand when you cross the street.'
Xámi amáñi ne. 'Walk on the grass.'
Danáñida chínada hi^{n}wámañi ke; héroda eswéna hi^{n}né ho. 'We walked into town yesterday. Let's go maybe tomorrow.'
Há^{n}we Waxóñida^{n} na^{n}kérida Jíwere Ñút^achi Wóya^{n}we hi^{n}náwi hi^{n}núšda^{n}wi ke. 'We already went to the Otoe-Missouria Encampment a week ago.' (lit. 'We already went to where the Otoe and Missouria sing a Sunday into the past.')
Há^{n}wegi míšda^{n} hamáñi hajé. 'Today I'm going for a walk by myself.'
Rí^e jegí ramína hñe ke. 'You stay here.'
Aréhga je? 'Are you telling the truth?' / 'Are you serious?' / 'Is that so?'
Hi^{n}wáha re. 'Show me.'
Ná^{n}je giwáha re. 'Go show your uncle.' (lit. 'father / father's brother')
Chúgwa jí re. 'Come back inside the house.'

Axéwe šgáje re. 'Go play outside.'
Núwerút^ana amína re. 'Go play on your bicycle.'
Grúya re. 'Clean up.'
Rixúmi ke aréchi kipídha re. 'You smell bad, so go take a bath.'
Šé mi^{n}táwe ke; ritáwe škúñi. 'That's mine, not yours.'
Wagúñetá^{n}i^{n} waná^{n}p^i hi^{n}nágirusdajena rusdá^{n} ne. 'Stop pulling on my rainbow necklace.'

Ga^ída. 'Over there.'
Akína ne. 'Wait.' / 'Watch out.'
Gasú^{n}hsji ke. 'Right now!'
Tórigu^{n}. '[See you / Not until] Later.'
Uxré. 'Soon.'
Náhehi^{n}na gigré re. 'Leave me alone.'
Ruhdá skúñi re. 'Don't touch that.'
Ritúga^{n} nigírixogešge ke. 'Your grandfather might scold you. '
Ritúga^{n} uhágidagešge ke. 'I'll tell your grandfather.'

Hú^{n}che. 'Yes.'
Hiñégo. 'No.'
Ahó. 'Hi there.' / 'Thanks.' / 'Good.' / 'Ok.' / 'I acknowledge that.'
Ahó warigróxi ke. 'Thank you very much.'
Pi ke! 'It's good.' / 'Wow, cool!' / 'Ok.'
Pi je? 'Is that good?'
Ripí je? 'Are you okay?'
Daríhga? 'How are you?'
Pí škúñi ke. 'That's bad.'
Wayí^{n}thwe skúñi ke. 'You're misbehaving.'
Uwáreri škúñi ke. 'I didn't understand you.'
Gasú^{n} ke. 'Enough!'
Rúšda^{n} ne. 'Stop it.'
Dáka^{n}hina uránadhena rúšda^{n} ne. 'Stop turning the light off and on.'
Á^{n}withruje škúñi ne. 'Don't squirt me.'

Jehéhgana šéhe ke. 'I told you so.'
Gasú^{n} srúšda^{n} ke. 'Ok, you're done.'

4 Strategies and challenges in Chiwere revitalization (Saul)

In this section, I would like to provide a broader context for understanding the Project's activities by describing their role within the Ioway and Otoe-Missouria communities and comparing them to strategies for revitalizing other Siouan and Native American languages. Following a long process of domain contraction, Chiwere was used primarily in religious contexts by 1950 (Davidson 1997; Furbee & Stanley 1996; 2002). This resonates with my own observations since I began working with Jimm in 2009. Besides interactions explicitly framed as language learning (e.g. tribal language classes), I have seen and heard Chiwere used for endonyms, salutations, valedictions, alimentation (especially water and common or traditional foods), elimination, kinship terms, personal names, and tribal programs. Some tribal members incorporate Chiwere into their personal or professional activities, especially if they are involved in art, music, and/or activism.[4]

[4] Examples of tribal programs with Chiwere titles include the Otoe-Missouria tribal newsletter, *Wórage: Stories of the People*, and the Iowa Tribe of Oklahoma's eagle sanctuary, Bah Kho-je

Nevertheless, as in many Native American communities (see e.g. Kroskrity 1998), ritual contexts are the most prestigious domains for Chiwere and other indigenous language use for Ioways and Otoe-Missourias, including in Native American Church meetings, ceremonial dances (e.g. Iroška), sweat lodges, and personal prayers. Since access to these activities can be limited or restricted, not all tribal members are exposed to Chiwere in these settings; furthermore, participants themselves do not always know the meaning of the Chiwere words they use or else they memorize their meanings (see e.g. Davidson 1997: 520–521).

The association between Chiwere and religious settings has fostered a sense that the language itself is sacred and that its circulation should be restricted like the distribution of other kinds of ceremonial knowledge (Davidson 1997). As a result, tribal members have few or no opportunities to hear or use Chiwere in their daily lives unless they have family members who make a special effort to use the language on a regular basis or participate in language classes. Currently, two of the three tribes for whom Chiwere is a heritage language have active language programs that offer classes and other educational resources: the Iowa Tribe of Kansas and Nebraska, whose language program is directed by the Tribal Historic Preservation Officer, Lance Foster, and the Otoe-Missouria Tribe of Indians, where Sky Campbell is the tribal language coordinator.

One strategy for reversing language shift in such contexts involves expanding the domains available for heritage language use. Jon Reyhner's (1999: vii) adaptation of Fishman's (1991) Graded Intergenerational Disruption Scale, for example, measures vitality in large part by whether a language is used in communal public spaces like educational and governmental institutions, businesses, and mass media. Jimm and his counterparts in tribal language programs try to raise awareness about language preservation by making Chiwere more visible in public places and encouraging people to use it more often. The last time I visited the Iowa Tribe of Kansas and Nebraska's offices, for example, Lance had put up signs in Chiwere identifying different offices and the restrooms. Sky has made similar signs for the Otoe-Missouria offices, and stop signs in the tribal complex parking lot are also in Chiwere.[5]

The Ioway, Otoe-Missouria Language Project has pursued a similar strategy of increasing Chiwere's public presence. For example, the Project designed and printed a tee shirt that includes a beadwork-inspired floral design, the Chiwere

Xla Chi (*Báxoje Xrá Chí*). For examples of Chiwere use by individual tribal members, see art, essays, and websites by Lance Foster (1989; 1996; 1999; 2009; n.d.[a],[b],[c]), an album by artist and musician Reuben Kent (2004), Jones (2004), and Brett Ramey's (n.d.) website.

[5] I am told that the Iowa Tribe of Oklahoma also has Chiwere signage in their offices.

endonyms for the Iowa, Otoe-Missouria, and closely related Ho-Chunk peoples; an image of an elder and a child wearing ceremonial dance clothes and a sentence in Chiwere that translates, 'The language honors our elders and teaches our children.' The Project also designed mugs that include the Chiwere phrase for 'I love my coffee' with the image of an Oneota-style ceramic vessel superimposed over a medicine wheel.[6]

In both cases, these objects were designed to set up educational interactions through question and answer routines. Since both objects include only Chiwere and no English text, those who do not know Chiwere have to ask someone who does if they want to know what the tee shirt or mug "says." When I have modeled the tee shirt at powwows and been approached by curious tribal members, Jimm has used the opportunity to explain the shared histories of the Iowa, Otoe, Missouria, and Ho-Chunk peoples and to discuss the aesthetic principles and symbolism of floral designs. Similarly, when I have witnessed interactions surrounding the coffee mug, Jimm has explained that the Chiwere word for 'coffee,' *máka^{n}thewe*, literally means 'black medicine,' just like the word for 'tea,' *xámi mákan*, literally means 'herb medicine.' This often leads into a discussion of traditional notions of medicine, healing, substance abuse, etc.

To take a more humorous example, Jimm's license plate reads *DAGWISA*, the Chiwere word for 'What did you say?' When people approach Jimm to ask about his license plate, a typical dialogue goes something like this:

"What does your license plate mean?"

"What did you say?"

"I said, what does your license plate mean?"

"What did you say?" etc.

While in this example the Chiwere language token does not appear in a culturally significant environment, it is still designed to promote an interactional context involving Jimm, who uses the opportunity to impart grammatical and cultural knowledge.

Bringing Chiwere out of religious settings and into the public sphere also carries risks. One of the primary dangers is that the language may lose its indexical associations with the traditional cultural values that motivate its revitalization in the first place. This is a phenomenon that other scholars have described for Tlingit and other Alaskan Native languages (Dauenhauer & Dauenhauer 1998),

[6] *Oneota* refers to an archaeological culture ancestral to a number of historical groups, including the Iowa and Otoe-Missouria. Some tribal members claim elements of Oneota culture as part of their cultural heritage, e.g., in Oneota-style ceramics by Ioway artist Reuben Kent (n.d. Rundle & Rundle 2007). Members of related Siouan groups, such as the Omaha, also see themselves as descendants of Oneota communities (Buffalohead 2004).

Apache (Nevins 2013; Samuels 2006), Elem Pomo and other California Indian languages (Ahlers 2006), Kaska (Meek 2010), and Maliseet (Perley 2011). Some communities, particularly in the Southwest, restrict the circulation of language materials in order to prevent the decontextualization of heritage languages from what are considered to be proper settings and forms of language use (Debenport 2015; Whiteley 2003). Jimm's language revitalization strategies represent another approach to the problem of decontextualization: rather than limiting access to language materials, Jimm promotes the circulation of materials that associate Chiwere with other symbols of traditional Ioway and Otoe-Missouria culture (e.g. the tee shirt and mug described above) and resists the circulation of materials that associate Chiwere with what he believes are dominant society practices and values.

Translation requests are one domain in which Jimm exercises discretion in order to control the cultural associations of Chiwere language tokens. Jimm often receives requests to calque English idioms – for example, 'Go green!' (for a tribal environmental awareness program); 'I ♡ boobies!' (for tribal breast-cancer awareness bracelets); 'Bigg Rigg' (for fans of Otoe-Missouria mixed-martial-arts fighter Johny "Bigg Rigg" Hendricks); and 'They are in a Warthog' (a phrase used in playing the video game *Halo*[7]). These requests are often met with ambivalence since they are seen as having no connection with traditional culture, and Jimm often declines to provide translation services for efforts that would increase Chiwere language use if he believes that such use would undermine traditional values.

Once, for example, a tribal member sent Jimm a list of English terms that she wanted translated into Chiwere. The list focused on terms for genitalia and bodily functions that are considered "bad words" in English. Jimm declined to provide the requested translations since he felt that the corresponding Chiwere terms lacked the negative associations of their English counterparts and worried that English speakers would project those associations onto the Chiwere terms. Many tribal members also believe that there are no "bad words" in Chiwere. Thus, swearing is one domain of language use where there is considerable resistance to using Chiwere. (For a description of an inverse situation, where an indigenous language is used almost exclusively to swear, see Muehlmann 2008.) In a similar vein, Jimm's response to the request to translate 'I ♡ boobies!' for a tribal breast cancer awareness program included a long description of traditional attitudes toward sexuality, which he felt ran counter to the slogan's sexual innuendo. Jimm was also concerned by the request to translate phrases used in playing the

[7] *Warthog* is a type of vehicle in the game.

video game *Halo*, which was intended to enable tribal members to communicate in Chiwere while playing the game. When I explained to Jimm what *Halo* is (a first-person shooter, i.e., a rather violent video game), he expressed reservations that Chiwere be associated with it and replied to the request by saying: "If I cannot contribute to peace and harmony, what the old people called *wapána*, then I cannot contribute at all."

Jimm's responses to these translation requests reflect the attitudes of some tribal members who have strong views about appropriate and inappropriate contexts for Chiwere language use. Occasionally, mere proximity between Chiwere and objectionable content can trigger concerns. Once, Jimm and I were approached by an elder who objected to Chiwere language lessons being posted on YouTube by a tribal member whose account also linked to music videos that contained suggestive and/or violent imagery. The elder was also concerned that the language lessons, which featured the voice of her deceased relative, were now publically available to unrestricted audiences. Of course, younger generations, particularly those who live far from tribal reservations, may appreciate the increased access to their heritage language that online platforms like YouTube provide, and they may have no qualms about viewing a Chiwere language lesson followed by a popular music video.

The challenge of giving an endangered heritage language a wider public presence while maintaining its traditional cultural associations is one faced by many working to revitalize Siouan languages. For example, a request on the Siouan List to translate a line from *Alice in Wonderland* ("curiouser and curiouser, cried Alice") for a polyglot compilation produced multiple responses. While some found the intellectual challenge of translating a Victorian neologism into Siouan languages intriguing, others were less receptive to the request because of its perceived triviality and irrelevance to Native American communities. Bryan wrote: "It's a more distinguished request than pet names[8] and such, but it's not the kind of translation work I would prefer to spend my time on. Why don't people ask us to translate Microsoft Word or a K-12 curriculum or something important?" (Gordon 2014). Jimm concurred: "I have other priorities and am unclear on the need for [a translation of] the particular quote from a story which has nothing

[8] Requests, many from non-Native people, to translate names for pets and children or stock English phrases into Siouan languages are so common that John Koontz (2003a) posted his general responses to such questions on the FAQ section of his website. Once, he was even asked (presumably as a joke) for a Native American name for an RV; he responded in kind with *Hotanke*, an Anglicized spelling of the Dakotan word for the Winnebago [Ho-Chunk] people, from one of whose English names the Winnebago brand of RVs took its name (Koontz 2003b).

in common with Native American Culture…. To spend time on the translation of materials that have no immediate application to the language communities is nonsensical and, for my part, a waste of time" (Goodtracks 2014). Willem de Reuse shared his general guidelines for responding to such requests: "One has to pick and choose. If it is short and culturally appropriate, I generally agree to it…. Then other requests have to be nixed, like the set of 'Spring Break' phrases I once was asked to translate, things like 'I am so drunk,' and 'Where is the bathroom?' " (de Reuse 2014a). As de Reuse (2014b) explained, part of the reason he objected to translating spring-break phrases is because the translations could be circulated in a way that would trivialize indigenous languages. In short, most linguists and activists working on Native American language revitalization face the question of what exactly is a "culturally appropriate" application of a heritage language – and how to prevent indigenous languages from appearing in culturally inappropriate contexts.

One solution to this problem is to embed heritage languages in contexts rich with other traditional cultural symbols. The tee shirt and coffee mug described above are two examples.[9] Another example is a board game produced by the Kaw Nation language department, *Wajípha^{n}yi^{n}*, which Jimm and I adapted for Chiwere and played with his grandson and other relatives. The game encourages players to imagine themselves as camp criers, who move among the traditional moieties of the tribe answering vocabulary questions to accumulate clan counting sticks (Kanza Language Project 2004). Many pedagogical materials produced for Siouan languages also qualify as culturally rich to the extent that their content addresses traditional themes (e.g. Hartmann & Marschke 2010; Kanza Language Project 2010). Chiwere pedagogical materials, for example, feature late nineteenth- and early twentieth-century photographs of Ioways and Otoe-Missourias in traditional dress and emphasize speech genres like prayer and oral histories of elders (Otoe and Iowa Language Speakers with Lila Wistrand-Robinson 1977a,b). Debenport (2015) notes that a "nostalgic" mood also permeates many Pueblo pedagogical materials.

Jimm's dictionary also emphasizes a connection between Chiwere and traditional cultural practices and values by including elaborate encyclopedia-style entries for terms with particular cultural significance. The entry for *mihxóge*, for example, begins by giving a range of more or less word-for-word English trans-

[9] Many other materials produced by the Project also seek to embed language in culturally rich environments. In the past, for example, the Project has published calendars with historical photographs and the names of the months in Chiwere (Goodtracks 1985). The Otoe-Missouria Tribe recently published a similar calendar (Otoe-Missouria Language Department 2014).

lations for the term, including "blessed person; a spiritual person or intermediary; gay, lesbian, homosexual, bisexual; two spirits person; transvestite; transsexual; berdache" (Goodtracks 2008: 6). Following the standard definition field is a long note that begins with a more literal definition of the word as "an individual who has some natural female-like aspect of their character, personality or nature, which is of a mysterious divine origin" based on a morphological analysis of *mihxóge* as derived from *miŋe*, indicating a feminine quality, *xóñita*^{n}, which refers to something sacred, blessed, or mysterious, and *-ge*, a suffix that indicates an innate or natural gift, ability, or state (Goodtracks 2008: 6).

The note then frames this analysis as a reflection of a traditional understanding of homosexuality associated with the elders, who respected *mihxóge* as spiritual leaders. The note includes quotations from elders ("They are *waxóbri*^{n} ['holy'], and they kinda know that and use it," and, "They're not crazy. They just got that born in them. Born in their nature," etc.), which are interpreted for readers in relation to current social conditions (Goodtracks 2008: 7). The note contrasts the elders' traditional views with Judeo-Christian attitudes, in which homosexuality may be seen as a choice or sin. The note also suggests that returning to traditional views would enable *mihxóge* to once again cultivate "their dormant 'medicine powers' and abilities" for the benefit of all (Goodtracks 2008: 7). The note concludes with a final quotation from an elder who gives instructions on how to behave toward *mihxóge*: "Talk to them, be good to them and that's all. But don't hurt them – it'll come back on you. They got medicine." (Goodtracks 2008: 7)

Thus, through a series of metapragmatic framings, the dictionary identifies the semantics of *mihxóge* (as reflected in its morphological composition) with its ancestral and potential future pragmatics. In other words, in defining the word *mihxóge*, the dictionary also teaches readers a set of traditional attitudes and behaviors associated with the elders. Jimm's desire to include the kind of information that readers could use to live the language in this way, which reflects a commitment to meet the needs of community audiences as well as linguists, explains some of the dictionary's unconventional formatting.

Jimm's language nest (discussed in more detail below) provided an opportunity for me and other participants to live the language. Within the language nest, for example, we never addressed or referred to each other as "Jimm," "Sage," or "Saul." Instead, we always put a kinship term before the proper name. This is evident in Bryan's preface to his list of Chiwere phrases, where he refers to Jimm as *Hi*^{n}*táro* 'my friend' and *Itúga*^{n} 'his [Sage's] grandfather' and to Sage as *Itágwa* 'his [Jimm's] grandson.' When I first came to live with Jimm and his grandson, Jimm

explained to me that it is impolite to address someone without a kinship term that expresses a consanguinal, affinal, or "fictive" relation because relatedness is the basis of Ioway and Otoe-Missouria society. Historically, families in Ioway, Otoe-Missouria, and other Native American communities often made strangers, including anthropologists, into relatives through adoption (see e.g. Kan 2001). Jimm locates the origin of such practices in clan origin myths, which describe how clans met each other and formed larger societies by establishing relationships in which they refer to each other as *hi^{n}táro* 'my friend[10].' Like Bryan, I refer to Jimm and his grandson, and they to me, as *hi^{n}táro*. Sometimes I refer to Jimm's grandson as *hi^{n}thúñe* 'my little brother', since after all these years he has become like a little brother to me. By using these kinship terms with each other, we foster and express relationships that have become part of our lives as lived. The examples of *mihxóge* and *hi^{n}táro* reflect how the Project and language nest link Chiwere language with cultural values and social action.

Of course, just as restricting Chiwere to religious settings limits opportunities for language use, associating the language exclusively with a nostalgic conception of "traditional culture" risks alienating those who struggle to see how a heritage language could be relevant to their modern lives. Linguists have noted on the Siouan List that some community members reject games as productive learning activities in language classes because the language is "sacred" or oppose colorful pictures and contemporary vocabulary in pedagogical materials because they are not "traditional" (de Reuse 2014b; Ullrich 2014a,b). Similar attitudes have been described for other Native American languages: Moore (1988: 463–464) mentions an example of how everyday talk in Wasco has become "mythologized" and subject to restrictions once applied only to a specific set of myths. Clearly, an exclusive association between indigenous languages and "sacred" or "traditional" domains may present an obstacle to revitalizing languages as means of everyday communication.

In short, while language revitalization seeks to create new opportunities for heritage language use by adapting the language to current conditions, this quest for relevance is tempered with the recognition that codes can become disasso-

[10] In the clan origin myths, this relationship is established through formal gift exchange and pipe ceremonies. *Itáro* also refers to what are known as *Indian friends* in certain varieties of American Indian English, in which two people are bound to each other by mutual ritual obligations. Between Jimm, his grandson, and me, the semantic sense of the term is closer to the English word *friend* since we are not bound to each other by ritual obligations, but our pragmatic use of the term to address or refer to each other differs from the conventional English usage of *friend* in a way that reflects how Ioways and Otoe-Missourias use kinship terms to express the social and cultural value of relatedness.

ciated from the traditional values that motivate their revitalization in the first place. If what we care about is not only preserving linguistic diversity (in the sense of grammatical structures) but also preserving distinctive cultural worldviews and lifeways, we will have won the battle while losing the war if people are learning and using heritage languages primarily to participate in activities associated with the dominant society. As an Apache bilingual teacher wonders, if children are only learning how to use Apache to order a cheeseburger, what's the point? (Samuels 2006: 551) Revitalizing Chiwere, Apache, and other endangered Native American languages seems to require a balance between enabling language learners to order cheeseburgers or play video games and encouraging them to pursue a deeper engagement with ancestral cultural traditions that they will simultaneously transmit and reinvent for future generations.

A second danger that accompanies attempts to make heritage languages more visible in public spaces is that the languages tend to be used as emblems of indigenous identities without actually producing speakers (Ahlers 2006; Dauenhauer & Dauenhauer 1998: 98; Whiteley 2003: 715). This reflects the fact that as much as some language-revitalization efforts are motivated by a desire to reconnect community members with aspects of traditional culture, they are also informed by a thoroughly modern, nationalist notion of the role of language in political and social life (Kroskrity & Field 2009; Nevins 2013). In other words, people are not necessarily interested in learning the language; rather, they may want to display the code in order to accomplish nonlinguistic symbolic goals. It is often disappointing for linguists to discover, for example, that people may want a dictionary of their heritage language to have, but rarely if ever to take off the shelf to read. Community language classes also have a tendency to serve symbolic and social needs rather than being effective in producing speakers.

A sense of disillusionment with community classes and other forms of language use that affirm identities but rarely produce speakers is common among linguists who work on Siouan languages. In an interview, for example, Bob Rankin told me that accurate documentation is more important than pedagogical materials for producing new speakers:

> So, if we get it right, those of us who are lucky enough to have been able to work with this last group of speakers, then the materials will be available to future scholars, or I always hope that —. You know, every 10,000 kids there's some language genius who's born, there's some little kid who can just pick up languages, and I was one of those, so I know they exist. And there'll be some little Kaw kid or some little Omaha kid who'll pick this up and just attack it hammer and tongs and actually learn it someday. I'm

> not one of those people who optimistically believes that language-retention programs or language-teaching programs are going to resuscitate these languages. People are too busy to learn languages. Language learning is not easy for most people. It's hard work. Everybody wants to know a foreign language, but nobody wants to study a foreign language, and I completely understand that from all those Romance irregular verbs that I had to memorize, but the materials will be there. If the tribe needs them, if scholars need them, they'll at least be there.

Bob presents himself here as something of a messianic realist: he distances himself from "optimists" who believe that classes will produce adults who learn their heritage language as a second language and instead emphasizes the importance of documentation, which will allow those with a gift for picking up languages to teach themselves.

Similarly, as Jimm points out above, the inspiration for his language nest came in part from his disillusionment with community language classes and similar programs that produced a substantial discourse about the importance of language revitalization but failed to produce new speakers who used the language on a daily basis. A disconnect between discourse supporting language preservation and a lack of effective action to reverse language shift is unfortunately a common phenomenon (see e.g. Dauenhauer & Dauenhauer 1998). Like Bob, Jimm has focused on documentation in recent decades. "Even if no one is interested now," he told me, "it will all be there in the dictionary for anyone who comes along and wants to learn."

In addition to putting more energy into documentation, Jimm also established his home as a language nest, in which he raised his grandson to speak Chiwere as his first language. The term language nest originates from and is a translation of a Māori language revitalization program called Te Kōhanga Reo, which focuses on early-childhood language immersion (King 2008). Te Kōhanga Reo inspired similar programs in Hawai'i (Warner 2008; Wilson & Kamanā 2008) and other indigenous communities in North America and beyond. Whereas Te Kōhanga Reo and similar programs are usually communal or corporate in nature, Jimm's language nest is domestic and includes only Jimm, his grandson, and graduate students who have lived with them, such as Bryan and myself.

Jimm's household is the only one I know of where Chiwere is used as a primary language of communication and is one of the few places where Chiwere is spoken on a regular basis at all. White Cloud (or Chína Maxúthga, as we called the small town where we lived) is near the reservation of the Iowa Tribe of Kansas and Nebraska. Chiwere had not been a native language there for many

years. One of the last fluent speakers, Arthur Lightfoot, was born there in 1902, but he moved to Oklahoma in 1935. When he died in 1996, he was one of two or three fluent speakers left and the only one of his dialect. The town itself, named after an Ioway chief, runs from the banks of the Missouri River up into the neighboring bluffs. It was a regional center of commerce and culture during the steamboat era, and the downtown district is listed on the National Register of Historic Places for its "sense of historic time and place as a nineteenth century river town" (Wolfenbarger 1996: 1). For that reason, a few scenes from the film *Paper Moon* (Bogdanovich 1973), set in Great-Depression-era Kansas and Missouri, were filmed there. Today, the area feels like a ghost town. From a population of 1,000 or so in 1868, the 2010 census reported 176 residents, almost 20% Native Americans. Abandoned buildings and houses outnumber those permanently or seasonally occupied. When cars come through, they are usually on their way to the nearby reservation bingo hall and casino. The town is only busy twice a year for semiannual flea markets, but at the last one I went to, in 2012, I heard people complain that numbers were down because vendors and shoppers prefer to go to markets in towns closer to where they lived.

For many years, Jimm only spoke to his grandson in Chiwere, and Jimm told me that Bryan had done the same when he had lived with them for a summer. Since everyone else spoke to his grandson in English, however, his grandson had become bilingual and would use English unless addressed in Chiwere, which did not happen outside the home. Before I went to live with them, I prepared as well as I could to be a productive participant in their language nest. I made and memorized a few hundred flashcards, studied pedagogical materials, memorized Bryan's list of "Phrases in Báxoje Ich^é Indispensible to Living with a Three-Year-Old," and walked around campus listening to Chiwere recordings on my iPod.

The first summer, I spent a lot of time with Jimm's grandson. We were friends since there were few other children around and none that knew or seemed interested in learning Chiwere despite our proximity to the reservation. He was also my primary language teacher. I knew more grammar, but he knew more words, so I would constantly ask him for the names of things: "Dagúra? Jé^e ráye dagúra? Sé^e dagwígana je?" ('What is it? What's the name for this? What do they call that?'). During a walk through the woods, I asked about the Chiwere word for 'leaf', which got us confused as we tried to sort out *náwo* 'path', *náawe* 'leaf', *náwe* 'hand', and *núwe* 'two'. Since we only communicated in Chiwere, we often played in relative silence. When we went to the playground, we would chase each other and jab each other with our fingers. We did not jab each other hard, but the person who was jabbed was supposed to yell, "Ow, pahí^{n}, gích^e

ke!" ('Ow, that's sharp, it hurts!'). Or, we would climb to the top of the slide and ask each other what we saw:

"Díno xáñe arásda je?' ('Do you see a big dinosaur?'), he would say.

"Hú^{n}je, áta ke," I would say. ('Yeah, I see it.')

Our Chiwere repertoires consisted primarily of play routines and household phrases. 'Good morning.' 'Pass the salt.' 'Have you seen my keys?' 'Where's the dog?' 'Good night.' In novel situations, we communicated with varying degrees of success through a combination of Chiwere, English, and body language. My own oral Chiwere skills peaked quickly that summer because Jimm and I usually spoke to each other in English. The topics we discussed were often technical (related, for example, to the formatting of dictionary entries, computer issues, and the like), and neither of us felt comfortable addressing them in Chiwere. From time to time we would talk about talking more in Chiwere with each other, but we never kept it up for long.

After the first summer, I returned and lived with Jimm and his grandson on and off for two years during my fieldwork. I volunteered to help home school his grandson in Chiwere. Chiwere is the first language he learned to read. When we began matching letters to sounds, his favorite activity was going through the consonant and glottal stop combinations: *k^a, k^e, k^i, k^o, k^u.* One problem was that there were only two books in Chiwere for us to read. One, called *Hó Gíthige*, was a story about an uncle and nephew going fishing that Jimm translated from Lakota. The other was *Mischíñe na Náthaje*, a wéka^{n} ('myth') Jimm had hand-illustrated. His grandson was soon bored of reading those two books over and over. On the math side, I was able to teach addition and subtraction in Chiwere but had a hard time trying to explain more advanced topics like multiplication or division.

Pressures to improvise gave our Chiwere novel features, including codeswitching and lexical and grammatical innovation. For example, Jimm's grandson would say "Wanna sgáje?" for 'Do you wanna play?' We were also forced to come up with new words for a number of household objects, some of which found their way into Jimm's dictionary. When Jimm's grandson wanted something to drink, he would say "Dagúra sráhda^{n}?" which means 'What do you want to drink?' He associated this phrase with receiving something to drink because this is what Jimm would ask him before giving him something to drink. Jimm's grandson was unaware that *rahda^{n}* and other *ra-* initial verbs follow an irregular conjugation pattern for first- and second-person forms. Sometimes, I responded by saying what I wanted to drink rather than giving him something to drink to try to help him understand that *sráhda^{n}* is a second-person form – that is, I would

respond to the semantic rather than pragmatic meaning of his utterance. He would say, "Hiñégo, mí^e hasráhda^{n}!" ('No, me, I-you-drink!'), *hasráhda^{n}* being an ungrammatical form that includes both the regular first-person prefix *ha-* and the irregular second-person prefix *s-*. Jimm's grandson seems not to have recognized the irregular pronoun prefixes at all and thus treated the *s-* as part of the word for 'drink'. Moore (1988) describes a similar tendency to lexicalize already inflected forms as stems available for further inflection among younger speakers and semispeakers of Wasco.

One challenge that faced the language nest was a lack of reinforcement beyond the household. All of Jimm's grandson's favorite television shows, movies, and books were in English, and he could rarely if ever use Chiwere to communicate with anyone besides his grandfather, Bryan, and me. Once when we visited another reservation to help Head Start teachers incorporate Chiwere into their classrooms, the tribal language coordinator asked Jimm's grandson to say something in Chiwere. "Won't you say a little something?" he said. "Even a word or two?" Jimm's grandson just stared at him with a shy smile and shook his head. "He doesn't do performances," Jimm said, "he uses the language to communicate." Unfortunately, Jimm's grandson had few opportunities to communicate in Chiwere outside the household, and supporters in the community treated him as something of a spectacle. Others expressed concern that his language acquisition would be delayed or that he would never learn to speak proper English. Over time, Jimm's grandson began speaking to me in English even when I would address him in Chiwere, and I incorporated more English into the home school. We were just as likely to read *Harry Potter* as *Hó Gíthige* and *Mischíñe na Náthaje*, and all our math was in English, which suited me well because my Chiwere abilities found their limit in trying to explain multiplication and division. Chiwere went from being the medium of instruction to a special subject. As Jimm puts it, Chiwere may not have the prominent place that it once did, but it is still a permanent presence.

Reflecting on my experiences, Jimm's language nest has much in common with other home-based attempts to revive languages that are no longer spoken. One of the most famous examples of language revival, often presented as an inspirational model to indigenous communities whose heritage languages are "sleeping" (see Hinton 2008), is Hebrew. According to the popular narrative of Hebrew revival, Eliezer Ben-Yehuda, a late nineteenth-century Zionist, raised his children to be the first native speakers of modern Hebrew. For the previous 2,000 years, Jews had studied Hebrew for religious purposes but used vernacular languages for everyday communication. In the beginning, Ben-Yehuda's own Hebrew skills

were less than fluent, and his lexicon lacked terms for many household objects and activities. Thus, when he wanted his wife to prepare a cup of coffee, "he was at a loss to communicate words such as 'cup,' 'saucer,' 'pour,' 'spoon,' and so on, and would say to his wife, in effect: 'Take such and such, and do like so, and bring me this and this, and I will drink.' " (Fellman 1973: 37–38) Over time, however, Ben-Yehuda came up with words to fill these gaps. Furthermore, family friends urged Ben-Yehuda and his wife to speak other languages to their children because they feared that they would grow up unable to speak at all (Ben-Avi 1984: 50–53). Like Ben-Yehuda's household, Jimm's language nest also involved dealing with gaps between communicative goals and abilities, innovating lexical and grammatical forms, and resisting external pressure to speak dominant languages.[11]

There are also parallels between Jimm's household and a more recent case: the revival of Miami, an Algonquian language, which was labeled "extinct" when the last fluent speaker died in the 1960s. In the 1990s, however, Daryl Baldwin, a tribal member, learned Miami from historical documentation and then taught his wife and four children. Like Jimm and his grandson, the Baldwins also lived in a rural area and home-schooled their children in order to create some separation from outside influences. Nevertheless, all of the Baldwins are bilingual and bicultural. The family's language is characterized by a focus on domestic topics, relatively simple grammar, and changes from classical Miami reflecting some influence from English as well as lexical innovation. The Baldwins agreed to use Miami with each other whenever possible, but in practice around 30% of their conversation time was in Miami, though there was considerable variation depending on the topic (Leonard 2007: 14).

Leonard argues that the Miami case shows the reclamation of sleeping languages as languages of daily communication is possible, but he is careful to note that the measure of success in the Baldwin household is not Miami language fluency. Instead, the goal is to develop language proficiency as a means of enhancing the family's connection to traditional worldviews and their modern Miami identities. Daryl Baldwin refers to the goal as "cultural fluency" (Leonard 2007: 36–37; 2011: 139–140).[12]

[11] While Ben-Yehuda plays a central role in popular accounts of Hebrew revival, the emergence of modern Hebrew was a complex process that went far beyond his efforts (Harshav 1999). Nevertheless, his household still provides an interesting comparative context for examining home-based strategies for language revival.

[12] Another example of a similar language nest can be found in the documentary film *We Still Live Here* (Makepeace 2011), which addresses efforts by Jessie Little Doe to teach her daughter Wampanoag. Accounts of household language nests by Baldwin, Little Doe, and others are included in the recent edited volume *Bringing Our Languages Home* (Hinton 2013).

Similarly, Jimm's goal was never for his grandson to speak only Chiwere forever. Jimm expected his grandson to switch to English as his primary language of communication but hopes that his life will always be enriched by Chiwere language and associated cultural traditions. Like Leonard, my experience living in a household working to revive an indigenous heritage language has led me to believe that such efforts are entirely possible, especially when the heritage language has a strong connection to distinctive cultural practices and values, when there is a collective commitment within the household to use the language for communication as often as possible, and when there is an openness to the new words and grammatical structures that arise when speaking a language that has not been spoken for many years. A supportive broader community or state apparatus may be necessary for language revival on a large scale, but much can be accomplished by an individual "language genius" or family if they dedicate themselves to the task.

5 Conclusion

In this chapter, we have tried to present a multifaceted account of the Ioway, Otoe-Missouria Language Project's Chiwere revitalization activities, including personal accounts of our motivations and histories, a textual artifact from Jimm Goodtracks' Chiwere language nest, and a broader context for understanding how the Project's approach compares to other approaches to language revitalization in Siouan communities and beyond.

There is no way to tell what the future for Chiwere and other sleeping Siouan languages will hold. As our accounts demonstrate, successful language revitalization often depends on a fortuitous and potentially fragile combination of individual dedication, support from family and friends, collaboration with linguists and communities, and a sense of responsibility to previous and future generations. In the words of the Ioway, Otoe-Missouria Language Project's mission statement, the goal of the Project is "to share in part the Elders' desire to continue their language and traditional culture and knowledge," but "ultimately, it is the role of each descendant to continue it further among their own family and relations."

Hó, Náwo Pí ramáñišge tahó! May you walk a good road!

References

Ahlers, Jocelyn C. 2006. Framing discourse: Creating community through native language use. *Journal of Linguistic Anthropology* 16(1). 58–75.

Ben-Avi, Ittamar. 1984. Ittamar Ben Avi, the "guinea pig infant" of modern spoken Hebrew, tells how he uttered his first word in Hebrew (1885). In Levi Soshuk & Azriel Eisenberg (eds.), *Momentous century: Personal and eyewitness accounts of the rise of the Jewish homeland and state, 1875–1978*, 50–53. New York: Herzl Press.

Bogdanovich, Peter. 1973. *Paper moon*. Hollywood: Paramount Pictures.

Buffalohead, Eric. 2004. Dhegihan history: A personal journey. *Plains Anthropologist* 49(192). 327–343.

Dauenhauer, Nora Marks & Richard Dauenhauer. 1998. Technical, emotional and ideological issues in reversing language shift: Examples from Southeast Alaska. In Lenore A. Grenoble & Lindsay J. Whaley (eds.), *Endangered languages: Current issues and future prospects*, 57–98. Cambridge: Cambridge University Press.

Davidson, Jill D. 1997. *Prayer-songs to our elder brother: Native American Church songs of the Otoe-Missouria and Ioway*. Columbia: University of Missouri (Doctoral dissertation).

de Reuse, Willem. 2014a. *Alice translation*. Posted to the Siouan mailing list, 11 April 2014. http://listserv.linguistlist.org/pipermail/siouan/2014-April/009475.html (28 June, 2015).

de Reuse, Willem. 2014b. *Saul and Jan's discussion*. Posted to the Siouan mailing list, 25 April 2014. http://listserv.linguistlist.org/pipermail/siouan/2014-April/009557.html (28 June, 2015).

Debenport, Erin. 2015. *Fixing the books: Secrecy, literacy and perfectibility in Indigenous New Mexico*. Santa Fe: School for Advanced Research Press.

Fellman, Jack. 1973. *The revival of a classical tongue: Eliezer ben Yehuda and the modern Hebrew language*. Berlin: de Gruyter Mouton.

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

Foster, Lance M. 1989. *Mankánye washi*. http://ioway.nativeweb.org/iowaylibrary/mankanye_washi_1.htm (28 June, 2015).

Foster, Lance M. 1996. The Ioway and the landscape of southeast Iowa. *Journal of the Iowa Archeological Society* 43. 1–5.

Foster, Lance M. 1999. Tanji na che: Recovering the landscape of the Ioway. In Robert F. Sayre (ed.), *Recovering the prairie*, 178–190. Madison: University of Wisconsin Press.

Foster, Lance M. 2009. *The Indians of Iowa*. Iowa City: University of Iowa Press.
Foster, Lance M. n.d.(a). *Ioway cultural institute*. Institutional website. http://ioway.nativeweb.org/ (27 June, 2015).
Foster, Lance M. n.d.(b). *Nemaha*. Personal weblog. http://nemaha.blogspot.com/ (27 June, 2015).
Foster, Lance M. n.d.(c). *The sleeping giant*. http://hengruh.livejournal.com/ (27 June, 2015).
Furbee, Nonnie Louanna & Lori A. Stanley. 1996. Language attrition and language planning in accommodation perspective. *Southwest Journal of Linguistics* 15. 45–62.
Furbee, Nonnie Louanna & Lori A. Stanley. 2002. A collaborative model for creating heritage-language curators. *International Journal of the Sociology of Language* 154. 113–128.
Goodtracks, Jimm G. 1985. *Banyi Baxoje-Jiwere Nut^achi: An Indian new year calendar featuring Iowa, Otoe-Missouria people, may 1985 – april 1986*. Perkins, OK: Self-published.
Goodtracks, Jimm G. (ed.). 1992 – present. *Baxoje, Jiwere-Nyut^aji, Ma^unke Ich^e Wawagaxe: Ioway, Otoe-Missouria language to English dictionary*. Boulder, CO: Center for the Study of Native Languages of the Plains & Southwest. http://iowayotoelang.nativeweb.org.
Goodtracks, Jimm G. 2004a. *Text for Baxoje~Jiwere Ich^e – Ioway ~Otoe Language – CD I*. Lawrence, KS: Baxoje Jiwere Ich^e – Ioway Otoe-Missouria Language Project.
Goodtracks, Jimm G. 2004b. *Text for Baxoje~Jiwere Ich^e – Ioway~Otoe language – CD II*. Lawrence, KS: Baxoje Jiwere Ich^e – Ioway Otoe-Missouria Language Project.
Goodtracks, Jimm G. 2008. Mihxóge. In Jimm G. Goodtracks (ed.), *Baxoje, jiwere-nyutâji, maûnke ichê wawagaxe: Ioway, Otoe-Missouria language to English dictionary*. Lawrence, KS: Baxoje, Jiwere Ich^e – Ioway, Otoe-Missouria Language Project. http://iowayotoelang.nativeweb.org/pdf/m_btoe_feb2008.pdf (28 June, 2015).
Goodtracks, Jimm G. 2014. *Here is a little curious inquiry for all of us to ponder*. Posted to the Siouan mailing list, 11 April 2014. http://listserv.linguistlist.org/pipermail/siouan/2014-April/009473.html (18 June, 2015).
Goodtracks, Jimm G. n.d. *Welcome to the Ioway, Otoe-Missouria language website*. Topeka, KS: Baxoje Jiwere Ich^e – Ioway Otoe-Missouria Language Project.

Gordon, Bryan James. 2014. *Here is a little curious inquiry for all of us to ponder.* Posted to the Siouan mailing list, 11 April 2014. http://listserv.linguistlist.org/pipermail/siouan/2014-April/009468.html (18 June, 2015).

Harshav, Benjamin. 1999. *Language in time of revolution.* Stanford: Stanford University Press.

Hartmann, Iren & Christian Marschke (eds.). 2010. *Texts with analysis and translation and an audio-cd of original Hocąk texts* (Hocąk Teaching Materials 2). Albany: State University of New York – Albany Press.

Hinton, Leanne. 2008. Sleeping languages: Can they be awakened? In Leanne Hinton & Kenneth Locke Hale (eds.), *The green book of language revitalization in practice*, 3rd edn., 413–418. Bingley, West Yorkshire: Emerald Group Publishing.

Hinton, Leanne (ed.). 2013. *Bringing our languages home: Language revitalization for families.* Berkeley, CA: Heyday Books.

Jones, Matthew L. 2004. Wahtohtana héda Ñyút^achi Mahín Xánje Akípa: The year the Otoe and Missouria meet the Americans. *Wíčazo Ša Review* 19(1). 35–46.

Kan, Sergei (ed.). 2001. *Strangers to relatives: The adoption and naming of anthropologists in Native North America.* Lincoln: University of Nebraska Press.

Kanza Language Project. 2004. *Wajípha^{n}yi^{n}: A Kanza language game.* Board game. Kaw City, OK: Kaw Nation.

Kanza Language Project. 2010. *Kaá^{n}ze wéyaje – Kanza reader: Learning literacy through historical texts.* Justin T. McBride & Linda A. Cumberland (eds.). Kaw City, OK: Kaw Nation.

Kent, Reuben. 2004. *Bidanwe.* Compact disc.

Kent, Reuben. n.d. *Turtle Island storyteller Reubin Kent: These native ways.* http://wisdomoftheelders.org/turtle-island-storyteller-reubin-kent/ (4 June, 2015).

King, Jeanette. 2008. Te kōhanga reo: Māori language revitalization. In Leanne Hinton & Kenneth Locke Hale (eds.), *The green book of language revitalization in practice*, 3rd edn., 119–128. Bingley, West Yorkshire: Emerald Group Publishing.

Koontz, John E. 2003a. *FAQ.* http://spot.colorado.edu/~koontz/faq.htm (28 June, 2015).

Koontz, John E. 2003b. *Names.* Personal website. http://spot.colorado.edu/~koontz/faq/names.htm (28 June, 2015).

Kroskrity, Paul V. & Margaret C. Field (eds.). 2009. *Native American language ideologies: Beliefs, practices and struggles in Indian country.* Tucson: University of Arizona Press.

Leonard, Wesley Y. 2007. *Miami language reclamation in the home: A case study*. Berkeley: University of California – Berkeley (Doctoral dissertation).

Leonard, Wesley Y. 2011. Challenging "extinction" through modern Miami language practices. *American Indian Culture and Research Journal* 35(2). 135–160.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2013. *Iowa-Oto*. 17th edn. Dallas: SIL International. http://www.ethnologue.com/language/iow (28 June, 2015).

Makepeace, Anne. 2011. *We still live here*. Lakeville, CT: Makepeace Productions.

Meek, Barbra A. 2010. *We are our language: An ethnography of language revitalization in a Northern Athabaskan community*. Tucson: University of Arizona Press.

Moore, Robert E. 1988. Lexicalization versus lexical loss in Wasco-Wishram language obsolescence. *International Journal of American Linguistics* 54(4). 453–468.

Muehlmann, Shaylih. 2008. 'spread your ass cheeks,' and other things that should not be said in indigenous languages. *American Ethnologist* 35(1). 34–48.

Murray, Franklin. 1977. Iowa Tribe of Oklahoma language. In Iowa Tribe of Oklahoma (ed.), *Bah Kho-Je: People of the grey snow*. http://bahkhoje.com/wp-content/uploads/2014/11/Franklin-M.-Murray-An-Original-{Ioway}-Speaker-1896-1983-First-{Ioway}-Dictionary-created-by-him.pdf (22 April, 2015).

Nevins, Marybeth Eleanor. 2013. *Lessons from fort Apache: Beyond language endangerment and maintenance*. Hoboken, NJ: Wiley-Blackwell.

Otoe and Iowa Language Speakers with Lila Wistrand-Robinson. 1977a. *Jiwele-Baxoje Wan'shige Ukenye ich'e – Otoe-Iowa Indian language book I: Alphabet, conversational phrases and drills*. Park Hill, OK: Christian Children's Fund American Indian Project.

Otoe and Iowa Language Speakers with Lila Wistrand-Robinson. 1977b. *Jiwele-Baxoje Wan'shik'okenye ich'e – Otoe-Iowa Indian language book II: Simple, compound and complex sentences; songs and stories for practice*. Park Hill, OK: Christian Children's Fund American Indian Project.

Otoe-Missouria Language Department. 2014. *Jiwere-Nut'achi banyi koge nuwe agrį grebrą agrį thatą: 2015 Otoe-Missouria calendar*. Red Rock, OK: Otoe-Missouria Language Department.

Parks, Douglas R. & Robert L. Rankin. 2001. Siouan languages. In Raymond J. DeMallie & William C. Sturtevant (eds.), *Plains*, vol. 13 (Handbook of North American Indians), 94–114. Washington: Smithsonian Institution.

Perley, Bernard C. 2011. *Defying Maliseet language death: Emergent vitalities of language, culture, and identity in eastern Canada*. Lincoln: University of Nebraska Press.

Ramey, Brett. n.d. *Hóragewi*. Personal website. http://www.horagewi.com/.

Reyhner, Jon. 1999. Introduction: Some basics of language revitalization. In Jon Reyhner, Gina Cantoni, Robert N. St. Clair & Evangeline Parsons Yazzie (eds.), *Revitalizing indigenous languages*, v–xx. Flagstaff: Northern Arizona University.

Rundle, Kelly & Tammy Rundle. 2007. *Lost nation: The Ioway 1*. Moline, IL: Fourth Wall Films.

Samuels, David W. 2006. Bible translation and medicine-man talk: Missionaries, indexicality, and the "language expert" on the San Carlos Apache reservation. *Language in Society* 35(4). 529–557.

Kroskrity, Paul V. 1998. Arizona Tewa Kiva speech as a manifestation of a dominant language ideology. In Bambi B. Schieffelin, Kathryn A. Woolard & Paul V. Kroskrity (eds.), *Language ideologies: Practice and theory*, 103–122. Oxford: Oxford University Press.

Ullrich, Jan F. 2014a. *Citation / quotation conventions for list?* Posted to the Siouan mailing list, 25 April 2014. http://listserv.linguistlist.org/pipermail/siouan/2014-April/009556.html (28 June, 2015).

Ullrich, Jan F. 2014b. *Saul and Jan's discussion*. Posted to the Siouan mailing list, 25 April 2014. http://listserv.linguistlist.org/pipermail/siouan/2014-April/009559.html (28 June, 2015).

Warner, Sam L. Noʻeau. 2008. The movement to revitalize Hawaiian language and culture. In Leanne Hinton & Kenneth Locke Hale (eds.), *The green book of language revitalization in practice*, 3rd edn., 133–144. Bingley, West Yorkshire: Emerald Group Publishing.

Whiteley, Peter. 2003. Do "language rights" serve indigenous interests? Some Hopi and other queries. *American Anthropologist* 105(4). 712–722.

Wilson, William H. & Kauanoe Kamanā. 2008. Mai loko mai o ka ʻiʻini – Proceeding from a dream: The ʻAha Pūnana Leo connection in Hawaiian language revitalization. In Leanne Hinton & Kenneth Locke Hale (eds.), *The green book of language revitalization in practice*, 3rd edn., 147–176. Bingley, West Yorkshire: Emerald Group Publishing.

Wolfenbarger, Deon. 1996. *National register of historical places registration form for the White Cloud historic district*. Submitted to the National Park Service, United States Department of the Interior. http://pdfhost.focus.nps.gov/docs/NRHP/Text/96000701.pdf (28 June, 2015).

Chapter 8

Reconstructing post-verbal negation in Kansa: A pedagogical problem

Justin T. McBride

Despite the fact that there are no L1 speakers of Kansa, and the handful of learners are mostly novice-range speakers, the Kaw Nation has been actively engaged in revitalization efforts for many years. The absence of speaker knowledge poses a major problem for curriculum developers insofar as the quality of Kansa pedagogical materials is often limited to what can be uncovered from analysis of documentary materials – mostly those of Dorsey and Rankin. These sources, though essential, are far from complete. For instance, they lack many constructions that potential language learners would want to know, including how to express what in English is captured by the word *wouldn't*. In such cases, syntactic analysis can be used to reconstruct certain areas of Kansa grammar. Kansa is a left-branching, head-marking language with canonical (S)OV word order. Several features of its syntax seem to complicate an X-Bar treatment of Kansa, but the placement of negation (NEG) in the post-verbal complex seems to violate a number of principles all at once. This gives rise to contradictory expectations for its location in different contexts. In this paper, I discuss one way of reconstructing Kansa NEG to fill a pedagogical need. While not arriving at any definite theoretical conclusion, I do arrive at a possible one, and conclude with a set of recommendations for curriculum developers dealing with this and other such problems.

1 Introduction

Among the many problems plaguing the revitalization efforts of languages without L1 speakers is that a large number of the useful, conversational things that learners might want to say are simply unknown. These things may include greetings and pleasantries, common expressions for introducing self and others, stating likes and dislikes, making and fulfilling requests for additional information, telling time, and so on – all of which people use with great frequency in their own L1s and expect to be able to say in an L2. In fact, language teachers usually want to teach these sorts of conversational forms early on in classes as stock constructions that can build both competence and confidence in their learners. However,

Justin T. McBride. 2016. Reconstructing post-verbal negation in Kansa: A pedagogical problem. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 165–182. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.171
DOI:10.17169/langsci.b94.171

with no speakers around to ask, there may only be the products of linguistic research available as the next best thing. Perhaps there is a dictionary, a text series, or simply a set of field notes. Yet, even the most diligent field worker may not think to elicit very practical expressions such as 'hello, my name is [blank],' 'I did not understand what you said; please repeat it,' or 'how do you say [blank] in the [blank] language?'

The case of the Dhegiha Siouan language Kansa (also known as Kanza or Kaw) is precisely as described above. Dorsey's 1880s-era field work yielded a rather large set of slip files, two dozen texts collected from nine separate consultants, and hundreds of pages of ethnographic notes, all of which Rankin used in his own extensive work with the last Kansa speakers in the 1970s and early 1980s. Following the deaths of his consultants, Rankin continued working on Kansa for the rest of his life. Neither Dorsey nor Rankin intended their work to be used as-is for revitalization and curriculum development purposes, but this is what happened: Such efforts must begin somewhere, and their material was the logical starting point. Fortunately, Rankin was willing to contribute to this enterprise, and he often worked on Siouan language pedagogy side-by-side with other linguists (these included myself and several other contributors to this volume), both in the classroom and behind the scenes. Yet, even with Rankin — himself a lifelong educator — and a team of Siouanists at the head of Kansa language classes, the learner outcomes were often far less than could be expected of other beginning language classes; the source material was simply incomplete. As a consequence, many basic things remain unknown for Kansa and, accordingly, unused among language learners.

Consider a common English expression such as 'She *wouldn't* go, 'Jonathan *wouldn't* do that,' 'they *wouldn't* give it to me,' or the like. To my knowledge, there is no recorded translation of this expression in the available Kansa materials. I honestly cannot recall the exact circumstances of how this lacuna was discovered, but I remember that it came up in the Kaw Nation's Thursday night community language class in Kaw City, Oklahoma, in the mid-2000s. Perhaps some thoughtful student simply asked, "How do you say, 'she wouldn't go,' in the language?" Surely I knew that the answer to the question would involve post-verbal negation and some use of both the potential and non-continuative enclitics, but I was flummoxed as to how to order these elements. Whatever the circumstances may have been, once it became apparent that I could not immediately provide an answer based on my working knowledge of Kansa syntax, I probably explained that I would have to do more research and return with a defi-

nite solution later. Little did I know then that I *wouldn't* have a satisfying answer the next week, month, or even year!

Part of the problem lies in just how one would go about trying to find the answer. It would ideally involve reviewing the available texts and field elicitations with an eye toward finding how Dorsey, Rankin, or someone else may have recorded it. Those working on Kansa have, of course, done a great deal of secondary research like this; the lacunae are numerous, the learners are curious, and the available scholarly analysis is of high quality. Nevertheless, the construction does not appear in the materials. Failing that, the next step would involve reconstructing the form from a set of near-matches combined with knowledge of the language's syntax. Yet, syntax is one area where Kansa and the other Dhegiha languages are not always described in the greatest detail. Both Rankin's brief grammatical sketch of Kansa 1989 and his later sketch of Quapaw 2005 discuss a variety of syntax topics, as does Quintero's (2004) book-length grammar of Osage. But all of these works are overviews of Dhegiha grammar, and are ultimately too general to offer fine-grained perspective on such a specific question.

In this chapter, I will attempt a basic generative syntactic analysis of Kansa post-verbal negation. Bear in mind that I am ultimately looking for a pedagogical solution, not a theoretical one. As such, I do not advocate any particular theory of formal syntax and feel fairly free to borrow liberally from several eras of transformational grammar all at once. I am fully aware that this juxtaposition of concepts may make my analysis problematic for strict syntacticians, and perhaps also for dedicated pedagogues who may find any such analysis tedious to begin with. I do this not to alienate potential readers or to break any new theoretical ground, but simply to predict an unattested enclitic order using the formal means within my disposal. I also hope that my analysis and the discussion that follows will help to shed some light on a few philosophical principles that I consider very important to anyone working in Siouan languages:

- Gaps in the available documentation of languages are not necessarily insurmountable challenges;
- Grammar must occasionally be reconstructed in order for it to be taught;
- Formal analysis is not, by mere virtue of its formality, better than other means of acquiring grammatical knowledge; yet
- Formal analysis of some manner or another can serve practical pedagogical purposes.

1.1 X-Bar considerations

Kansa, like other Mississippi Valley Siouan languages, particularly those of the Dhegiha branch, is head-marking with a canonical (S)OV word order (see, for example, Quintero (2004: 421) for the Dhegiha language Osage; Rankin (2005: 488–490) for Quapaw, also Dhegiha; and Cumberland (2005: 369) for Assiniboine, a Dakotan language). Moreover, it appears to follow the same sort of left-branching syntactic pattern that Boyle (2007) described for Hidatsa (Missouri Valley Siouan). Although this paper is concerned with the syntax of the Kansa post-verbal complex, it is important to point out some grammatical features that complicate an X-Bar analysis of Kansa, including those as follows.

(a) Left-branching: Tree structures for Kansa and the movement of elements within them appear to run counter to the right-branching patterns typical of X-Bar theory.

(b) Radical pro-drop status: For the most part, only nominal subjects and objects appear independently in the sentence, the former presumably in the [SPEC, TP]; all else is handled by verbal inflection.

(c) Concept of word: Just how much of enumeration and derivation is left up to morphology versus syntax is essentially still up for grabs; as a consequence, so, too, are the classifications of enclitics, auxiliaries, affixes.

(d) Abstract tense: The TP in Kansa is at best misnamed given the language's general absence of tense marking, and the projection below the topmost Kansa CP is probably little more than an agreement-checking level.

These features are crucial to any full description of Kansa syntax, and they have very interesting implications for syntactic theories as a whole. Nonetheless, while I take these points as fundamental assumptions for the analysis that follows, they are actually not altogether relevant for me to discuss in greater length given the narrow focus of this chapter.

1.2 Aspect

Tense may be absent in Kansa, but verbal aspect is quite developed. Figure 1 shows the general division of aspect in Kansa. The primary division is between what I have termed simple and augmented aspect. Simple aspect is obligatory in all clauses while augmented is not. Simple aspect is further divided into continuative/imperfect (CONT) and non-continuative/perfect (NCONT) aspects, which are

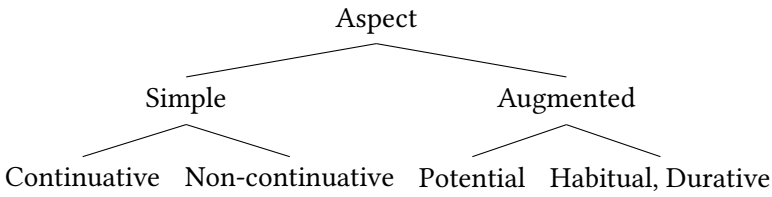


Figure 1: Kansa aspect

in complementary distribution. CONT is marked on verbs through a complex series of post-verbal enclitics (Rankin (2005: 485) argues that these enclitics "are actually conjugated as fully-fledged auxiliary verbs" in the closely related Dhegiha language Quapaw; see also Rankin (2004), for a much more detailed discussion) that also carry with them a sense of the subject's physical orientation in space. These include such categories as CONT-LIE, CONT-SIT, CONT-STAND, CONT-MOVE, etc. Moreover, these auxiliaries agree with the phi features of the verbal subjects. Non-continuative aspect (NCONT), on the other hand, is marked in two ways: A null form (-∅) is used with 1SG, 1DU, and 2SG subjects; the verbal enclitic *-(a)be*[1][2] is used with 1PL, 2PL, and 3CN subjects. This suggests a person and number configuration as shown in Figure 2. The augmented division includes potential (POT)[3] on the one hand and several habitual (HAB) and durative (DUR) aspects on the other. The features HAB and DUR function syntactically in the same way as POT. I classify these as augmented due to the fact that they can be combined as needed with either simple aspect to generate compound aspects such as POT CONT, POT NCONT, HAB NCONT, etc. The feature POT consists of the underlying enclitic *ce*, which only surfaces as such when no other post-verbal elements – aspect or mood – follow it; this is very rare, but it does occur. It most often takes the

[1] I have written the initial vowel in parentheses to avoid a digression into what is occasionally known as ablaut in Siouan. Suffice it to say, this initial vowel surfaces only when the final vowel of the element to which it attaches ends in *-e*, presumably due to a V1+V2=V2 rule involving Kansa /e/ and /a/. For a more detailed treatment of this phenomenon throughout Siouan, see Rankin (1995).

[2] Both to save space and to preserve consistency with source material where appropriate, all Kansa words in this chapter are written only in the practical orthography. This system is phonological in nature, but uses fewer special characters, allows digraphs and trigraphs, and makes use of English-based capitalization and punctuation standards that potential language learners may regard as normal.

[3] Note that POT is occasionally regarded in the Siouan literature as an irrealis marker (see Quintero 2004; 2009).

shape of *ta* through a sequence of regular phonological changes.[4] Though the phonology of this variation is understood, the mechanism behind it is not, a fact that raises some interesting questions about its enclitic status. Note that Figure 2 is concerned with the post-verbal arrangement of person and number considerations, and there are pronominal prefixes that are shared between numbers for the same person, including, for instance, for 1DU and 1PL and for 2SG and 2PL.

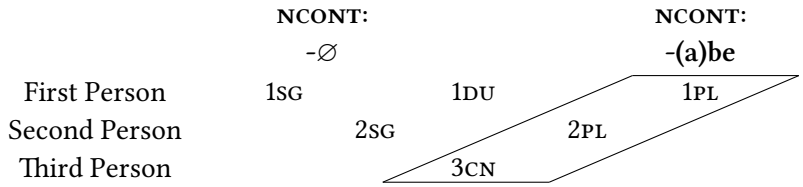


Figure 2: Person and number categories in Kansa with respect to NCONT marking

1.2.1 POT enclitic status

POT, unlike other post-verbal enclitics, is syntactically dependent on what comes before it (it is enclitic to the main verb, presumably as the head of a PotP) but phonologically dependent on what comes after it (its shape is determined by its proximity to the end of the clause). Furthermore, owing perhaps to its consonantal rather than vocalic onset,[5] it does not interact phonologically with the main verb. As such, the POT enclitic is somewhat different from that of, say, NCONT.[6]

1.2.2 Aspect order

The clauses in example (1) and their templatic arrangement in Table 1 illustrate some representative combinations of the major aspects and the order in which

[4] The so-called ablaut considerations mentioned in footnote 1 are presumably responsible for two allomorphs of the POT enclitic surfacing in different post-verbal phonetic environments. These forms include *ce* and *ta*, the former of which further exhibits routine spirantization of the initial stop before a front vowel.

[5] The other augmented aspect enclitics, HAB and DUR, also feature consonantal onsets, a fact that may strengthen the notion of augmented aspect as a natural class in Kansa.

[6] The fact is represented in the Kansa practical orthography by a space between the main verb and POT where no such space is left between the main verb and NCONT.

they typically occur post-verbally.[7,8,9]

(1) a. Wipághe tá mi^{n}khe.
∅-wi-p-(g)aghe *ta* *mi^{n}khe*
3CN-2CN.PAT-1SG.AGT-1SG.CONT.SIT POT 1SG.AGT.make
'I will make them for you' (KR, p. 192)

b. Yuzé ta akhá.
∅-∅-yuze *ta* *akha*
3CN-3CN-take POT 3CN.CONT.REST
'S/he was about to take it.' (KR, p. 200)

c. Hne tábe.
hn-(y)e *ta* *-(a)be*
2PL.AGT-go POT -NCONT
'You (PL) will have gone.' (KR, p. 192)

d. Ozhú tábe.
o-∅-∅-zhu *ta* *-(a)be*
in-3CN-3CN-pour POT -NCONT
'S/he would plant it.' (KR, p. 111)

Sentences (1-4) suggest the following canonical order of post-verbal aspect elements: V, POT, CONT/NCONT. This can be represented in tree form as shown in Figure 3.

2 The problem, in formal terms: NEG and aspect

Kansa syntax involves the use of post-verbal negative (NEG) enclitics, particularly as used in different aspect combinations. Kansa NEG has two separate forms:

[7] In this paper, I mark pronouns using AGT for agent and PAT for patient, without regard to the various inflectional realizations found throughout Dhegiha; the use of null pronouns in third person makes the classification as agent or patient irrelevant. I mark number using SG for singular, DU for inclusive dual, and PL for plural. I also use CN, after Kelly's (1992) Hebrew gender convention, to represent so-called common number in third person where singular and plural have collapsed in Kansa.

[8] All clausal examples in this paper come from sentences in McBride & Cumberland (2009), *Compiled Kanza texts*, or McBride & Cumberland (2010), *Kanza reader*, abbreviated CKT and KR, respectively. Corresponding page numbers appear after the English glosses.

[9] The analysis of pronominals here differs from that presented either in Quintero (2004) or Rankin (2005), where all 1SG.AGT and 2SG.AGT pronominals are represented by archiphonemic WA- and YA-, respectively, and phonological rules are needed to explain their phonetic realization. To simplify things, I have simply shown final realizations in the analysis.

Table 1: Order of post-verbal aspect elements

example	V	POT	CONT/NCONT
(1a)	∅-wi-p-(g)aghe	ta	mi^{n}khe
(1b)	∅-∅-yuze	ta	akha
(1c)	hn-(y)e	ta	-(a)be
(1d)	o-∅-∅-zhu	ta	-(a)be

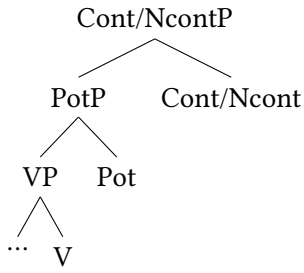


Figure 3: Order of post-verbal aspect elements

It appears either as *-(a)zhi* or *-mazhi*, the latter of which is only used with 1s subjects.

2.1 NEG with POT and CONT

When NEG is used in either CONT or POT CONT aspects, it appears consistently before both, as shown in examples (2) and Table 2.

(2) a. Gó^{n}yazhi akhá. *∅-∅-go^{n}ya-(a)zhi-akha*
3CN-3CN-want-NEG-3CN.CONT.REST
'S/he does not want it.' (CKT, p. 211)

b. Ashká^{n}mazhi tá mi^{n}khe. *a-shka^{n}-mazhi ta mi^{n}khe*
1SG.AGT-move-1SG.NEG POT 1SG.CONT.SIT
'I will not be stirring around.' (CKT, p. 40)

These examples suggest a canonical order of V, NEG, POT, CONT, as seen in Figure 4.

Table 2: Order of post-verbal NEG, CONT, and POT elements

example	V	NEG	POT	CONT/NCONT
Gó^{n}yazhi akhá.	∅-∅-go^{n}ya	(a)zhi		akha
Ashká^{n}mazhi tá mi^{n}khe.	1SG.AGT-move	-1SG.NEG	POT	1SG.CONT.SIT

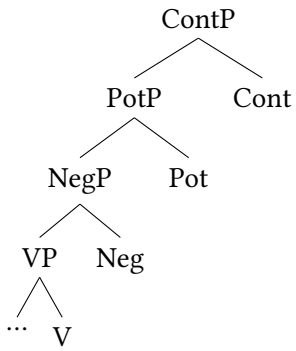


Figure 4: Order of post-verbal NEG, CONT, and POT elements

2.2 NEG with NCONT

However, when NEG appears with the phonetically realized NCONT *-(a)be*, it seems to fall after NCONT, as seen in the clauses of (3) and Table 3.[10]

(3) a. Píbazhi.
∅-pi-(a)be-(a)zhi
3CN-be.good-NCONT-NEG
'S/he was bad.' (CWK, p. 208)

b. Shká^{n}bazhi.
∅-shká^{n}-(a)be-(a)zhi
3CN-move.around-NCONT-NEG
'S/he did not stir.' (KR, p. 180)

[10] The verb in (9) undergoes a complex phonological process that turns the pronominal a^{n}(g)- + the instrumental i- into a^{n}ya^{n}-.

c. A^{n}yá^{n}kikiyabazhi.
a^{n}(g)-i-∅-kiki-ye-(a)be-(a)zhi
1AGT-to-3CN-RECIP-see-NCONT-NEG
'We did not see each other.' (KR, p. 263)

Table 3: Order of post-verbal NEG and NCONT elements

example	V	NCONT	NEG
Píbazhi.	∅-*pi*	(a)be	-(a)zhi
Shká^{n}bazhi.	∅-shká^{n}	-(a)be	-(a)zhi
A^{n}yá^{n}kikiyabazhi.	a^{n}(g)-i-∅-kiki-ye	-(a)be	-(a)zhi

Here, the order appears to be V NCONT NEG. This contradicts the canonical orders seen above, as demonstrated in Figure 5.

(1-2)	V		POT	CONT		
(5)	V	NEG		CONT		⇐
(6)	V	NEG	POT	CONT		⇐
(3-4)	V		POT	NCONT		
(7-9)	V			NCONT	NEG	⇐

Figure 5: Contradictions of Kansa NEG placement

In short, the data suggest that NEG appears both before the slots reserved for POT and simple aspect and after the slot reserved for simple aspect. Note that there do not appear to be clearly identifiable examples of NEG with POT NCONT, the case that would best clarify the ambiguity of Kansa NEG placement and help me to answer the question I was posed about the Kansa equivalent of *wouldn't*. With no attested form in the corpus, it is difficult to say whether it is an ungrammatical form or simply a gap in what was recorded. How would the combination of NEG, POT, and NCONT look with a 3CN subject where *-(a)be* would most certainly surface? Would it appear as *tabázhi, -(a)zhi tábe,-(a)bazhi ce*, or something else entirely? What would such a form tell us of the syntax of Kansa negation? It seems that the Kansa equivalent of the English sentence 's/he would go,' *ayé tábe*, would provide insight into how the equivalent of 's/he would not go,' might look. Yet, the data do not steer us toward any clear solution.

2.3 NEG with person and number

One final consideration must be mentioned before commencing a proper examination of the problem set. Recall that the phonetic realization of NCONT is restricted to only 1PL, 2PL, and 3CN subjects. Thus, the remainder of forms, namely those with 1SG, 1DU, and 2SG subjects, will not clarify these issues. This can be seen in (4) and Table 4.

(4) a. Kó^{n}blamazhi.
∅-k-(g)o^{n}-bl-(y)a-∅?-mazhi-∅?
3CN-1SG.AGT-want$_1$-1SG.AGT-want$_2$
'I do not wish it.' (KR, p. 188)

b. Phímàzhi
ph-(h)i-∅?-mazhi-∅?
1SG.AGT-arrive.there-NCONT-1SG.NEG-NCONT
'I did not reach there.' (KR, p. 92)

Table 4: Ambiguity involving NEG with null NCONT

example	V	NCONT	NEG	NCONT
Kó^{n}blamazhi.	∅-k-(g)o^{n}-bl-(y)a	-∅?	-mazhi	-∅?
Phímàzhi	ph-(h)i	-∅?	-mazhi	-∅?

3 Analysis

3.1 Enclitic placement

NCONT and NEG resemble one another more than they resemble POT, both syntactically and phonologically. This fact at least suggests they are members of a common grammatical class. For one, as neither independent words nor simple suffixes, NCONT and NEG seem to be subject to more restrictive placement considerations than the CONT auxiliaries in the post-verbal environment. This distinction seems to be reinforced by the fact that NCONT and NEG are phonologically dependent on preceding material. Secondly, their placement appears to be more restricted than that of POT.

Logically speaking, there are three environments in which NCONT or NEG may occur: 1) after POT; 2) after one another (e.g., NEG after NCONT); or 3) after the main verb. There have already been examples of the first two, but let us review all three for the purpose of classifying these environments.[11]

(5) a. *ahíbe* ← Environment 1: NCONT after V
3CN.AGT.MOVE.arrive.NCONT
's/he arrived there' (KR, p. 92)

b. *shká^{n}bazhi* ← Environment 2: NEG after NCONT
3CN.AGT.MOVE.NCONT.NEG
's/he did not stir' (KR, p. 180)

c. *ozhú tábe* ← Environment 3: NCONT after POT
in.3CN.3CN.pour POT.NCONT
's/he would plant it' (KR, p. 111)

At this point, it is necessary to distinguish between the distributions of NCONT versus NEG. In the data above (as elsewhere in Kansa), at no time does NCONT appear after NEG. Also, recall the complementary distribution of NCONT and CONT, a distribution that is unlike that of NEG and CONT. On the other hand, (7-9) above demonstrate that NEG can appear after the NCONT enclitic. If one further stipulates that NEG follows the null realization of NCONT in (10-11), it is possible to claim that NEG in Environment 2 is required to attach to NCONT whenever possible. Furthermore, while NEG can appear with either POT (7-9) or CONT (5-6), it appears unable to come after either of these. Thus, the distribution of NCONT and NEG is as follows:

(15) Distribution of NCONT: Environments 1 and 3

(16) Distribution of NEG: Environments 1 and 2

NEG presumably arrives in these environments by means of head-to-head incorporation and/or excorporation as described by Roberts (1991).[12] It can either

[11] Concerning (12) *ahíbe*, Kansa has a class of motion verbs including 'go', 'come', etc. that include a motion prefix *a-* on certain forms. I have termed this MOVE in the gloss. Note that the semantics of these verbs does not preclude use of them in either continuative/imperfect or non-continuative/perfect aspect; both are completely grammatical.

[12] Roberts defines excorporation as "successive cyclic head-to-head movement where one head simply 'passes through' another, first incorporating and then moving on" (1991: 211).

arrive at the verb (Env. 1) in continuative/imperfect aspect or at the the main verb plus NCONT (Env. 2) in non-continuative/perfect aspect. Such enclitic lowering derives a new verb. Thus, NEG appears to attach to the lowest verb in the TP as seen in Figures 6 and 7.

(6) Ashká^{n}mazhi tá mi^{n}khe.

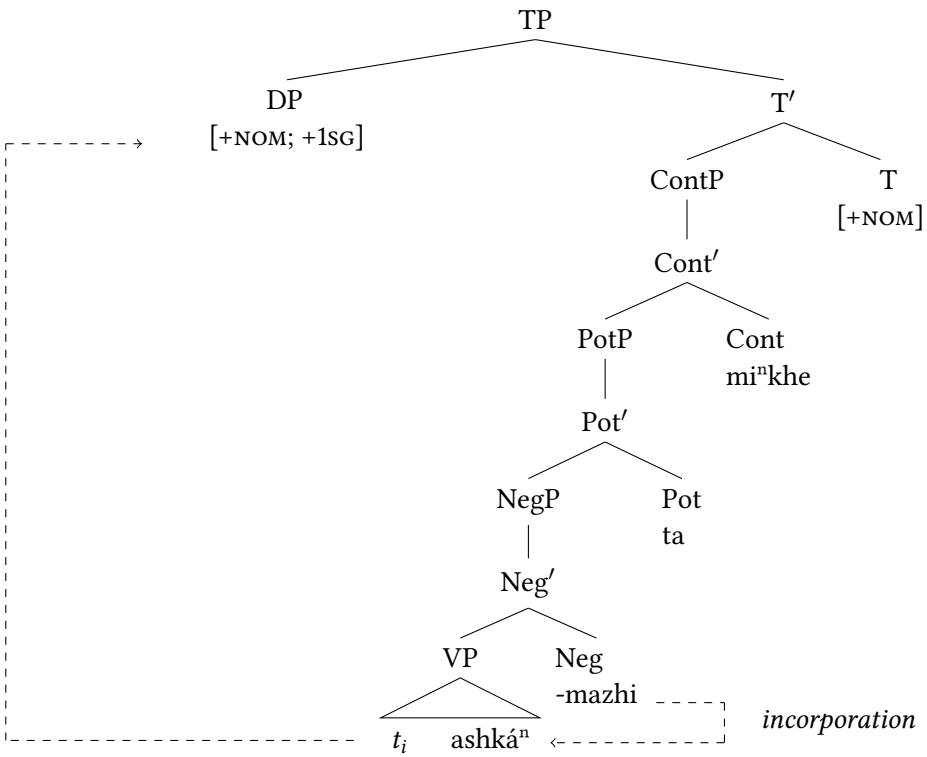


Figure 6: Tree of (6) NEG with CONT aspect

3.2 Feature expansion and prediction

This solution is not particularly satisfying for several reasons. The first is that feature checking does not appear to motivate the enclitic lowering. It is possible, however, to adjust for this simply by adding features that may or may not be checked through movement. We may assume, however, that if an enclitic of any type can move to check a nearby feature, it will do so. Such a process would

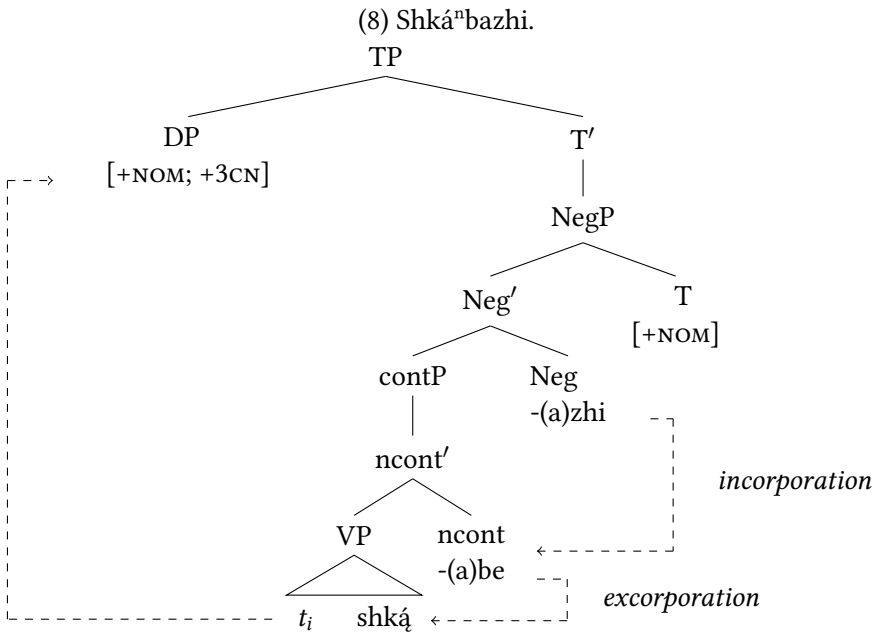


Figure 7: Tree of (8) NEG with NCONT aspect

account for all enclitic lowering. The second drawback is that if NEG lowers before NCONT, the order of enclitics will be incorrect. Therefore, NCONT must somehow lower first. The third is that the status of the enclitics within the tree structures is not as clear as one would like. Are they really V heads, or are they just NEG, NCONT, POT, CONT heads? If they head their own projections, it would seem that their classifications together or separately would require a great deal of justification. On the other hand, classifying them all as V heads would require perhaps even more justification.

Nevertheless, these are exclusively theoretical concerns, and there are mechanisms within formal syntax that can be used to address them. My goal here is not to grind a theoretical axe, but merely to find a pedagogical answer to a student's question. Does my model do this? Yes: The predicted order of post-verbal elements in a Kansa sentence equivalent to English 's/he would not go' is as follows: V POT NCONT NEG, or *ayé tabázhi*. This consists of an inflected main verb, *ayé* (3CN.AGT.MOVE.go), followed by a compound enclitic *tabázi*, consisting of *ta* (POT), *-(a)be* (NCONT), and *-(a)zhi* (NEG).

I was happy with this possible solution, but — given the aforementioned theoretical concerns — not entirely so. Thus, when I presented an earlier version of this paper at the 2011 Siouan and Caddoan Languages Conference, I put the question to several Omaha and Ponca Elders in attendance. While they seemed to indicate that such a construction would not be at all common in their respective languages, they agreed that the cognate form of Kansa *tabázhi* would be the preferred option. This does not confirm the Kansa prediction, of course, but it does seem to suggest that the analysis leading to my prediction was at least on the right track.

4 Conclusion

In this chapter, I have shown how syntactic analysis of textual data relating to a question put forth by an eager learner can be used to extend our knowledge of Kansa and fill in gaps in the source material. But numerous big conceptual questions remain, even beyond the theoretical ones mentioned above. For instance, how useful is this particular analysis and application if, as has been apocryphally suggested for Omaha-Ponca, the English expression may occur at a far higher frequency than the equivalent Kansa expression? With no L1 speech community around to offer guidance, perhaps there is no way to answer this question. On the one hand, the deployment of a form that would not have been used in earlier times is the very nature of language. On the other hand, if it pragmatically separates the L2 speakers of Kansa from L1 and L2 speakers of very closely related languages, its use may work against larger speech community goals privileging the taking of cues from still vital Siouan languages rather than English. On a different level, is the prediction of an order of post-verbal elements, even one seemingly matching cross-linguistic evidence, a sufficient stopping place for analysis? Perhaps the predicted result offers a false confidence in the approach taken. Put in a slightly more philosophical way, is extensive analysis done on a dormant language of any value on its own terms, or does it derive its true worth from practical application in revitalization efforts? Certainly from the perspective of potential learners, the language benefits when it can be put to greater use, regardless of what theoretical or applied linguists may say. As I mentioned earlier, there are many, many problems that plague such situations!

In spite of these challenges, work like this can be useful both to linguistic theory and for practical purposes. For starters, it can be used to show that even deep holes in the available documentation can be filled with a little theoretical elbow grease. This is comforting to know, and I hope that my analysis can show

one way that it can be done. There are, of course, others. My speaking with the tribal Elders at the conference was what ultimately gave me confidence in my solution. I was lacking this confidence after just looking at the problem from a theoretical point of view. Nevertheless, in order to frame the question properly so that it could even be asked (and later taught), I did require some preliminary reconstructive work. The mere formality of the theory underpinning that reconstructive work did not make my solution somehow correct, but neither did it make it unattainable. At the risk of closing this chapter perched atop a linguist's soap box, I would add that language teachers should not fear formal syntax; it is just one more arrow in their quiver, and I hope I have shown here that it can be put to service in solving practical pedagogical problems.

Abbreviations

1, 2, 3	first, second, third person	NOM	nominative
AGT	agent	NCONT	non-continuative/ perfect aspect
CN	common number	PL	plural
CONT	continuative/imperfect aspect	PAT	patient
DU	dual	POT	potential
DUR	durative	RECIP	reciprocal
HAB	habitual	REFL	reflexive
MOVE	motion-verb prefix	SG	singular

References

Boyle, John P. 2007. *Hidatsa morphosyntax and clause structure*. Chicago: University of Chicago (Doctoral dissertation).

Cumberland, Linda A. 2005. *A grammar of Assiniboine: A Siouan language of the northern Plains*. Bloomington: Indiana University (Doctoral dissertation).

Kelly, Page H. 1992. *Biblical Hebrew: An introductory grammar*. Grand Rapids, MI: William B. Eerdmans Publishing Company.

McBride, Justin T. & Linda A. Cumberland (eds.). 2009. *Compiled Kanza texts*. Kaw City, OK: Kaw Nation.

McBride, Justin T. & Linda A. Cumberland (eds.). 2010. *Kaá^{n}ze wéyaje – Kanza reader: Learning literacy through historical texts*. Kaw City, OK: Kaw Nation.

Quintero, Carolyn. 2004. *Osage grammar* (Studies in the Anthropology of North American Indians). Lincoln: University of Nebraska Press.

Quintero, Carolyn. 2009. *Osage dictionary*. Norman: University of Oklahoma Press.

Rankin, Robert L. 1989. Kansa. In Takashi Kamei, Rokuro Kono & Chino Eiichi (eds.). Trans. by Akira Yamamoto, *Languages of the world* (The Sanseido Encyclopaedia of Linguistics 2), 301–307. Tokyo: Sanseido Press.

Rankin, Robert L. 1995. *On Quapaw (and Siouan) ablaut*. Paper presented at the Siouan and Caddoan Languages Conference, Albuquerque, NM, July 1995.

Rankin, Robert L. 2004. The history and development of Siouan positionals. *Sprachtypologie und Universalienforschung* 57. 202–227.

Rankin, Robert L. 2005. Quapaw. In Heather Hardy & Janine Scancarelli (eds.), *The native languages of the southeastern United States*, 454–498. Lincoln: University of Nebraska Press.

Roberts, Ian. 1991. Excorporation and minimality. *Linguistic Inquiry* 22(1). 209–218.

Chapter 9

Baxoje-Jiwere grammar sketch

Jill D. Greer

This synchronic grammar follows a descriptive approach to Baxoje-Jiwere of the Mississippi Valley branch of the Siouan language family. It expands upon prior published and unpublished documentation, based upon fieldwork conducted from 1987–96. Baxoje-Jiwere-Nyut'achi is a "sleeping language," with no fully fluent speakers at present, but with revitalization efforts underway in each of the three native communities of Red Rock and Perkins, Oklahoma, and White Cloud, Kansas. The sketch begins with phonology, morphology, then syntax, with special attention to the complex system of verbal affixes; the interesting phenomenon of noun incorporation within certain verbs; verb classes (regular stems, irregular stems in *r/l*, *w*, and *d*, and the causative construction); positional verbs, which may serve as auxiliary verbs; and SOV word order, with clause-final and utterance-final enclitics marking relation to the following clause, source of evidence, sentence type, and gender of speaker. The topic of language variation concludes the sketch, with gender differences documented for greetings and interjections; brief tables illustrate phonological and lexical distinctions associated with both tribal dialects.

1 Introduction

Baxoje-Jiwere belongs to the Mississippi Valley branch of the Siouan language family, and is the native language of the Plains/Prairie tribes known today as the Otoe-Missouria and Ioway (Goddard 1996: 3,8). While their original homelands were in northern Missouri, southeastern Nebraska, and the state of Iowa, during the late 19th century the two tribes relocated to a north-central portion of Indian Territory in an attempt to avoid Euro-Americans' increasing encroachment on their reservations and the assimilation policies of the BIA. One segment of the Ioway chose to stay on a portion of their original reservation near the Missouri River in northeastern Kansas (Wedel 2001; Schweitzer 2001).

The following sketch is based upon fieldwork in central Oklahoma which I conducted mainly between 1987 and 1996 while a graduate student and research assistant within a larger team, led by Dr. Louanna Furbee, and including Lori A. Stan-

Jill D. Greer. 2016. Baxoje-Jiwere grammar sketch. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 183–229. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.172 DOI:10.17169/langsci.b94.172

ley.[1] The research was conducted with the permission of the 1987 Otoe-Missouria Tribal Council, among members of both the Otoe-Missouria and the Ioway Tribes living in an approximately 100 mile radius of Red Rock, OK. It was funded initially by a University of Missouri Faculty Development Grant, then generously supported by the National Science Foundation Documenting Endangered Languages Program and the American Philosophical Society's Phillips Fund.

Báxoje is the Ioway tribe's name for their people and language. *Jíwere* is the native Otoe term for themselves (and the language), while *Nyút'achi* refers to the Missouria people/language.[2] The native language spoken by these two tribes has frequently been called *Chiwere* in the existing literature (Whitman 1947; Marsh 1936; Wedel 2001; Schweitzer 2001). However, because this spelling makes it more likely for English speakers to mispronounce the first sound of the Otoes' self-name, I prefer to use <J> instead, because the voiced allophone is far less likely to be aspirated by language learners with English as their first language. Goodtracks also follows this orthographic shift.

In addition to the two contemporary communities centered in Red Rock and Perkins, Oklahoma, respectively, there is also a Northern Ioway Nation located on their original reserve in White Cloud, Kansas. Sadly, there are no L1 speakers of Baxoje-Jiwere, but a few individuals may be semi-speakers. Language renewal efforts are underway in each of the small communities, so there is hope that while yet sleeping, the ancestral tongue may still be awakened.

Many factors led to this particular effort to document Baxoje-Jiwere, but the original impetus was the collegial friendship between two University of Chicago linguists (both students of Eric Hamp), the late Robert L. Rankin and N. Louanna Furbee. These two scholars both landed jobs in the Midwest, the former at KU in Lawrence, Kansas, and the latter just a few hours away at Mizzou. They remained in touch throughout the 1970s and 80s. As Bob adopted Siouan languages as his primary research focus, he saw the urgent need for more linguists. He would tease Louanna that since she was employed by the primary research university in the state that was named for one dialect of this highly endangered Siouan language, it was her duty to start doing research on it.

His good-natured urging came to fruition in 1987, when a critical mass of graduate students interested in language study surrounded Louanna and she offered

[1] Stanley's Ph.D. dissertation (Stanley (1993)) includes a life history of Dr. Truman W. Dailey, one of the primary speakers and contacts within the Otoe-Missouria Tribe, available at the University of Missouri-Columbia Library.

[2] Because the Missouria language was not recorded, I omit the name *Nyút'achi* when referencing the language in general, although the Missouria people and history are remembered in Otoe tribal heritage in the conjoined name today.

a special seminar on Siouan languages. About nine eager students enrolled in the course, myself included. Bob came to Mizzou to give a beginning workshop to Louanna's class, with stacks of handouts full of concrete suggestions such as questions to ask and topics to cover in the field.

His help did not end there, but continued throughout the years, giving feedback on papers, guiding our elicitation of forms for the Comparative Siouan Dictionary, reading much earlier versions of this sketch, and countless other generous acts on his part. Thus, without the initial friendship between Robert L. Rankin and N. Louanna Furbee,[3] there would have been no Missouri Chiwere Language Project grammar. This work is dedicated to them both. All errors are of course my own.

2 Sound system (phonology)

2.1 Consonants

2.1.1 Stops

There are three sets of stops distinguished by these features:

a. Aspiration /p^h, t^h , k^h/

b. Glottalization /p', t', k'/

c. Plain (neither aspirated nor glottalized) /b, d, g/

The "plain" sounds can be either voiced or voiceless, but the two allophones would have been heard by native speakers as the "same." Different scholars of Baxoje-Jiwere have used either or both [p/b, k/g, t/d] for the plain (lenis) series. Variation may have existed between closely related forms within the three historic speech communities, within some families, or even with particular speakers. Notes by earlier researchers suggest that individuals' speech did display such tendencies, but the data are too limited to address such topics at present (Whitman 1947). In addition, the glottal stop /ʔ/ does appear in word-initial, medial, and word-final positions, but in the first two instances, it serves primarily to prevent amalgamation and preserve semantic content before certain vowel-initial morphemes such as verb stems. In those settings, its function is morphological,

[3] The essential role of Louanna Furbee as major professor, grant writer, P.I., fieldworker, editor, friend, and all around pillar of strength cannot be overemphasized. The MCLP (Missouri Chiwere Language Project) original materials are archived at Luther College, Decorah, Iowa.

rather than phonemic per se. Likewise, it tends to appear in word-final position only for a limited set of morphemes, namely interjections and sentence-final particles/enclitics. In those instances, its phonetic abruptness carries an iconic meaning of emphasis, doubt, or even impatience (cf. Tables 18 and 19).

2.1.2 Affricates

As with the stop series, there are three contrasts: plain affricates /č/-/ǰ/, aspirated /čʰ/, and glottalized /č'/.

2.1.3 Fricatives

The plain series has a larger set of sounds than the glottalized versions.

a. Plain: /θ ð s š x h/

b. Glottalized: /θ' s' x'/

2.1.4 Nasals

The four nasal consonants are /m n ñ ŋ/. The latter two phonemes /ñ/ and /ŋ/ were significant as indices of tribal identity. Baxoje speakers favored *ñ* in words where Jiwere speakers typically said *ŋ*, such as 'horse': *šu**ñ**e* in Ioway vs. *su**ŋ**e* in Jiwere.[4] However, there are clear cases of /ñ/ in both dialects, such as the shared indefinite plural *-ñe*. Word-initial /n/ often palatalized to [ñ] before front high vowels /i, į/.

The /ŋ/ cannot occur word-initially, and probably is historically derived from phonological environments where a velar stop followed a nasal vowel. Note that there is a very strong tendency to pronounce an epenthetic homorganic nasal consonant when nasal vowels precede stops, probably for economy of effort, or making the word "smoother," as some elders liked to put it, as in the /m/ in *nąm**m**pʰo* 'finger'.[5]

[4] That example also illustrates another common pronunciation difference between the distinct versions of this language, namely the plain /s/ at the beginning of words for Otoe, where Ioway produces /š/ instead.

[5] Amelia Susman's (1943) work on Ho-Chunk (Winnebago) mentioned the same tendency in that very closely related Siouan language.

2.1.5 Liquids

There has been some difficulty defining and representing the liquid sound in Baxoje-Jiwere. Phonetically, it has been described as resembling an unreleased "flap" [d] like the medial sound in *latter*, the plain [r] found in Spanish, and a variation upon the [l] sound (Whitman 1947: 235); Rankin also included [ð, n, y] as possible phonetic reflexes (Wedel 2001: 346; Schweitzer 2001: 447). For orthographic consistency, the symbol /r/ will be used.

2.1.6 Glides

Glides include /w/ and /y/.

2.2 Vowels

2.2.1 Oral vs. nasal

There are both oral and nasal vowels in Baxoje-Jiwere. They include /a i o u e/ and /ą į ų/. Frequently /ą/ would be realized as a nasalized schwa.

2.2.2 Vowel allophones as gender indexicals

Phonetic vowel quality sometimes differs significantly in particular words used by female speakers especially; in those contexts, there is also an [ɛ] and sometimes an [æ]. These variations are limited to a particular small domain of the overall vocabulary of the language, and serve a social-indexical function. (Cf. section 5.2. on sentence-final particles and interjections.)

2.2.3 Vowel length

Robert Rankin transcribed long vowels from a recording of a key word list by a OtoeJiwere speaker, but I have been unable to perceive length on the same recording. No minimal pairs clearly establish phonemic significance of vowel length between etymologically unrelated words.[6] Thus, at present there is scant evidence to support the idea of *phonemic* vowel length, although the revised*Plains* volume of the *Handbook of North American Indians* presents a list of long and short vowels, based on Rankin's analysis (Wedel 2001: 432; Schweitzer 2001: 447).

[6] John Boyle's student presented a brief paper on this topic based on spectrographic analysis of MCLP recordings, but that paper has not been published.

However, there are very prolonged vowels that occur when morphological boundaries have been "blurred" during amalgamation. The greatly extended length preserves the mora from the contracted morpheme, and sometimes affects the stress pattern as well. It seems to be primarily a morphological rather than phonological process.

2.3 Stress/accent

Stress is both volume and pitch-based, with phonemic value in Baxoje-Jiwere, as in *ráwe* 'beaver' and *rawé* 'to count' (Goodtracks n.d.), or *gísa* 'to laugh at another (v.)' vs. *gisá* 'a knot (n.)' (Dorsey n.d.) When a root word with two syllables has additional affixes attached to it, the basic stress (and pitch) pattern can change, typically with primary stress shifting to the left in the case of prefixation, and addition of a secondary stress in the case of infixes or suffixes. An adequate prediction of stress patterns is beyond the scope of this grammar.[7]

2.4 Syllable structure

There is a strong tendency to end all syllables with a vowel,[8] thus (V) and (CV) are very frequent syllable shapes. Initial consonant clusters are allowed (CCV), but examples of CCCV have not been discovered, nor have (VCC). The consonant clusters shown in Table 1 may begin a syllable.

2.5 Longer sound patterns/prosody

For length constraint, phrase-level prosody is included under §4, Syntax.

2.6 Phonological processes

2.6.1 Elision

One of the most common changes, elision is characteristic of rapid speech, such as the final vowels mentioned in Footnote 8 which frequently are deleted.

[7] Cf. discussions of Dorsey's Law in Miner (1979) and Hale & White Eagle (1980).

[8] The few exceptions to the preference for vowel-final syllables would be represented as a CVC structure. However, such instances only appear in informal speech and seem to be elision. During quick speech, the final unstressed vowel disappears, yet speakers give the full "precise" pronunciation with final vowel if asked to repeat or clarify what they said. This seems to have been a major aspect of the historical sound changes separating Ho-Chunk from Jiwere.

Table 1: Syllable-initial consonant clusters

a. stop + liquid:	br-	*bra-*‘separated, spread in layers, sliced, flat’
	gr-	*grų* ‘to curse’
b. stop + glide:	pʰy-	*pʰyúbrą* ‘mint, medicine tea, Indian perfume herb’
	gw-	*gwák’ų* ‘to wipe off, scrape off, dry oneself (one’s body)’
c. fricative + stop:	sd-	*sdą* ‘to stop, cease, leave off’
	sg-	*sga* ‘to be white, shiny’
	šg-	*šgų́ñį* ‘no; not; (does) not’
	θg-	*θga* ‘to be white’ (old form)
	hg-	*hga* ‘to be white’ (Ioway)
d. fricative + liquid:	sr-	*sroge* ‘to remove object from inside a hole’
	θr-	*θríje* ‘easily, softly, slowly’
	xr-	*xra* ‘eagle’
e. fricative + glide:	sw-	*swąhi* ‘to soften’ (flesh, leather, stale bread)
	šw-	*šwąra* ‘soft (buckskin, flesh, cloth)’
f. fricative + nasal:	sn-	*sni* ‘cold’ (Wistrand-Robinson 1972)
	θn-	*θni* ‘cold’ (possibly archaic; Dorsey in Goodtracks (n.d.))

2.6.2 Vowel harmony and nasal spread

The nasal quality of a nasal vowel may “spread” regressively (from right to left) to nearby vowels. (Ho-Chunk scholars have documented such nasality spread not just to directly adjacent vowels, but also across the consonants /h/ and /w/ to the closest non-adjacent vowel (Helmbrecht & Lehmann 2010: 7–8).

2.6.3 Vowel ablaut

This well-known phenomenon among Siouan languages involves /a/ and /e/, which may alternate in a variety of settings, especially before particular verbs or certain suffixes, suggesting it is morphologically conditioned. Motion verbs are one set of verbs that trigger ablaut. Some verbs ending in *-e* such as *ugwe* ‘to enter’ and *re* ‘to go’ will also ablaut to final *-a* before *-wi* ‘definite plural’ as does the indefinite plural *ñe* > *na* before the definite *-wi* also. Conversely, verbal prefixes with final /a/ will ablaut to /e/ before the possessive *gra-* and the verb *udwáñį* ‘to fail to reach, fail to come up to’ (Whitman 1947: 239-40), as well as *doye* ‘to break’. The instrumental prefix *gi-* ‘by hitting’ (with an ax, hammer, or other object in the hand) will trigger ablaut from /a/ to /e/ in the pronominal prefixes which attach directly to it. The dative *gi-*, however, will not trigger the same

vowel change, despite the identical phonetic shape, supporting the idea that it is not a purely phonological process.

Examples:

(1) *Čʰúgwá-wi re.*
house.enter-DEF.PL IMP(male speaker)
[from čʰi 'house'+ ugwe̲ 'to enter' + -wi̲ 'DEF.PL']
'Come in the house, you-all.' (Marsh 1936: "Giants" Bk2 LN49)

(2) *Iwálà-wi ho.*
yonder.go-DEF.PL HORT(male speaker)
[from i-'there'+ wa-'directional' + re̲ 'to go'+ -wi̲ 'DEF.PL']
'Let's go over there!' (Marsh 1936: "The Twins" LN65)

(3) *He-grahi kʰi.*
1P.AGT-love DECL(female speaker)
[from ha̲-1P.AGT + grahi 'love']
'I love him.'

An alternative analysis accounts for the vowel change before *-gra* 'POSS' and *-gi* 'DAT/BEN' as two vowels coalescing. Since the key morphemes in question are consonant-initial, there would have to be an underlying vowel, either /e/ or /i/. The /e/ matches the target vowel, and parallels the 3PL form found in the independent pronoun *eʔe*, and the possessive *etháwe* 'his/hers/its' and *ethéwi* 'theirs'. But there is precedent within Baxoje-Jiwere for /a/ + /i/ to become /e/, which Whitman (1947: 239) called 'amalgamation.' The volume reviewer likewise suggested that possibility, *igra-*. That shape/meaning resembles the 3PL inalienable prefix on kin terms (Table 2), which parallels cognate Lakota forms and matches the reconstructed Proto-Siouan **i-* Possessive (on non-verbs) (Rood 1979). This analysis also reserves the term 'ablaut' to stem-final vowels, as has been the norm within Siouan scholarship. [9]

3 Words/morphology

3.1 Nouns

Many nouns can function fully as verbs, complete with the extensive system of prefixes and suffixes described later in the verbal template. Siouan languages are

[9] Unfortunately, *udwáñį* 'to fail to reach, come up to' and other verbs with separable prefixes preceding PRON prefixes also show the shift from final /a/ to /e/, [*uhédwañį* 'I fail to reach'] and these cases do not fit neatly into the proposed explanation (Whitman 1947: 240).

classified as strongly verb oriented, with very few prefixes or suffixes limited only to nouns.[10] Certain verbal prefixes transform that state/action into something more noun-like, as in the following example, wherein the verb 'to eat' becomes 'something to eat upon': *wá:ruǰe* 'table' < *wa-* 'indefinite object' + *a-* 'upon' + *ruǰe* 'eat'. Without the locative *a-* 'upon', the first vowel is not lengthened, and the stress remains on the second syllable: *warúǰe* 'something to eat, food.' Because there is a ∅ third person pronominal prefix, 'food' sounds identical to the third person singular sentence 'He ate (something).'

3.1.1 Possessing: inalienable vs. alienable

Native American languages often distinguish people and things extremely close to a person's identity and self (INALIENABLE) versus other entities that separate more easily (ALIENABLE). The former category includes kinship terms and in Baxoje-Jiwere, the formal social ties of friendship and parenthood.[11] The prefixes meaning inalienable possession are bound morphemes similar in shape to first and second singular person patient pronouns, but they differ in having an expressed third person form (which is sometimes dropped in fast speech), as shown in Table 2. (See Table 8 for personal pronominal prefixes.)

3.1.2 Address form *-o* 'speaking to this one'

While *hį́tʰára* 'my friend' is the unmarked referential form, a person would switch to *hį́tʰáro* 'my friend (address form)' while speaking directly to the special friend (formally established as cultural role).[13] Kin terms also take the same address morpheme when speaking directly *to* that person. The identical substitution of /o/ for final vowel affects line-final words in songs as well (Davidson 1997). There is no vowel variation by gender for this morpheme.

[10] Helmbrecht (2002) gives an extended discussion of ways to distinguish between nouns and verbs in Ho-Chunk (winnebago).

[11] While body parts may be inalienably possessed in other languages, it is not the case in Baxoje-Jiwere. Frozen remnants of such a system are evidenced if one interprets the initial *i-* in the following body parts as representing the third person *i-* inalienable prefix found in kin terms and other life-long social relationships like formal friendship and parenthood ('(his/her) child' *ičičį́ŋe* (Otoe), *ičičį́ñe* (Ioway)): *ihdóge* 'elbow', *iréje* 'shoulder', *isdą* 'eye'.

[12] See Goodtracks (1992 – present) dictionary for complete inventory.

[13] The friendship would have been initiated by parents of two children of the same sex, formalized with a ceremonial feast, and thereafter a lifelong bond of reciprocity and obligation existed between the two, to be recognized by this word *-tʰara* 'friend'. The ultimate duty came at the death of one friend, when the other would sit with the deceased's body for the duration of the wake, traditionally 4 days before burial would take place (Whitman 1936, Davidson 1997).

Table 2: Inalienable possession

Kin term[12]	Inalienable possessive prefix		
	1st person sg.	2nd person sg.	3rd person sg.
Father	*hį́-ka* 'my father / father's brother'		
	hį-daje (old) 'my father'	*naje <ri-qje* 'your father'	*aje < i-aje* 'his/her father'
Mother	*hį́-na* 'my mother / mother's sister'		
	hį́-hų archaic 'my mother'	đi-*hų* /*ri-hų* 'your mother'	*i-hų* 'his/her mother'
Man's elder brother	*hį-yį́na* 'my elder brother'	*ri-yį́na* 'your elder brother'	*i-yį́na* 'his elder brother'
Woman's brother	*hį-čįdo* 'my(FEM) brother'	*ri-čįdo* 'your(FEM) brother'	*i-čįdo* 'her brother'
Grandfather	*hį-t^{h}úga* 'my grandfather'	*ri-t^{h}úga* 'your grandfather'	*i-t^{h}úga* 'his/her grandfather'

3.1.3 Names

A proper name uniquely identifies someone, for both address and referential purposes. It also may encode key identity features (gender, clan membership, personal attributes/characteristics, or significant events relating to that person).[14] Both dogs and horses were named also (cf. Whitman 1936 for traditional Otoe-Missouria dog names).

Gender. Some names were identical for both genders within the same clan, but often a woman's form differed by the addition of *-mį* 'feminine' suffix. A nickname could be coined to tease someone, as when one elder told another they should call me *Toské-mį* 'Quick/Speedy-Woman,' because I had done something so quickly that it surprised them. While names for men were not specially marked, there was a masculine morpheme *-do* that occurs in words for male noun referents such as 'boy,' 'buck,' and 'bull'; see Table 3.

[14] The Reverend James Owen Dorsey collected names, their meanings, and clan identification during his brief fieldwork in the late 19th century. The Smithsonian Institution has his field notes, truly a rich resource for individuals interested in discovering more about names, now available in their digital archive.

Table 3: Gender affixes: *-do* 'MASC'; *-mį* 'FEM'

a) *ič^{h}i**dó**įñe* 'boy-child'	< *i-* 'at/around'+ *č^{h}i* 'house'+ *-do* 'MASC'+ -įñe 'small/DIM' [Ioway]
b) *ič^{h}i**mį**ŋ(e)* 'girl-child'	< *i-* 'at/around'+ *č^{h}i* 'house' + *-mį* FEM + -įŋe 'small/DIM' [Otoe]
c) *t^{h}a**do*** 'buck, male deer'	< *t^{h}a* 'deer'[a] + *-do* 'MASC'
d) *č^{h}é**do*** 'bull buffalo'	< *č^{h}é* 'buffalo, bison' + *-do* 'MASC'

[a] With white-tailed deer, a buck is clearly the "marked form" if the visible feature of antlers was the primary basis for assigning group membership.

Diminutive suffix *-įŋe, -šį́ŋe* [O-M]; *įñe, šį́ñe* 'small /DIM [Ioway]'. There are also cases in Ioway tales where the protagonist's name is created from a verb + diminutive suffix: [V + DIM > Name].

(4) a. *Bé-ñe-ĺŋe*
throw.out-INDEF.PL-DIM
'The Outcast' < 'Little One(They)Threw Away' (Marsh 1936: "The Outcast" LN141)

b. *Hįnú-šį́ŋe číla*
my.first.son-DIM dear
'My dear Little-Son' (Marsh 1936: "The Wanderer" LN200)

3.1.4 Number

Nouns do not inflect for plural or case; numerals may follow the noun to give an exact number, or verbal suffixes reveal plural information instead. Numbers may act as stative verbs, with patient inflection, as also happens in other Siouan languages such as Quapaw (Rankin 2005: 481) and Lakota (Ullrich 2008: 708).

Numerals. One through ten are the basics from which other numbers are expressed. Eleven through nineteen are formed using the formula 'X over ten' lit. 'ten-over-one': *grebrą agri (i)yąk^{h}i*, 'ten over two', etc. Multiples of ten become 'two tens' (lit. 'ten (be) two') *grebrą núwe* 'two tens over one', up to ninety-nine. An interesting example of word coinage is the large quantity 'one thousand'; it is expressed by the word *kóge* 'box or trunk', because shipments of money (presumably annuity payments from Washington, D.C.) arrived in packing boxes, each of which held one thousand dollars.

Ordinal numbers. Baxoje-Jiwere may use either a prefix *i-* or a suffix *-yą*.

i- 'ordinal marker': (Marsh 1936: "The Giant" Bk2)

(5) *walúxawe i-θátʰą dahá?-e*
bundle ORD-five it.is.standing-that.one
'that fifth upright bundle' (LN 25)

(6) *walúxawe i-šágwe dahá?-e*
bundle ORD-six it.is.standing-that.one
'that sixth upright bundle' (LN 30)

(7) *walúxawe i-šáhmą dahá?-e*
bundle ORD-seven it.is.standing-that.one
'that seventh upright bundle' (LN 34)

(*i-* 'ordinal marker') + *-yą* 'indefinite article' (Marsh 1936: "The Wanderer")

(8) *Dáñ-ɟ́=yą (utʰą?ɟwagi aškų).*
three-ORD=INDEF
'A third time (he makes them appear to him, it seems).' (LN 34)

(9) *Hetále idóyą dahági síge alé gų́?waškų.*
hetále i-dowe=yą dahági síge alé gų́?waškų
then.it.is ORD-four=INDEF time.it.is again it.is.this he.do.it.it.seems.'
'And then, he does it again for the 4th time, it seems' (LN 35)

3.1.5 Compound nouns

Baxoje-Jiwere compound nouns,[15] shown in Table 4 often have the modifying word precede the base noun, while other times the modifier(s) follow it. These words can also include names, i.e. *mąkʰá ruǰè* 'medicine eaters' denoting those who participate in the religious traditions surrounding the sacred sacrament peyote.

3.1.6 Culture contact and word coinage

There was strong resistance to borrowing from European languages throughout Plains tribes in general,[16] so it is not surprising that Baxoje-Jiwere speakers also chose to coin new words, or extend the meaning of existing words. For instance,

[15] In other Siouan languages, e.g. Lakota and Crow, there can be a greater degree of noun incorporation. See Ullrich (2008: 738); de Reuse (1994); Graczyk (1991).

[16] Cf. Brown (1999), also Larson (2004)

Table 4: Compound nouns

a. *čʰína*	'village' < čʰi 'house' + -na 'horizontal?'
b. *čʰína wanàxi*	'cemetery' < čʰina 'village' + wanaxi 'spirit, ghost'
c. *walúšge čʰìna*	'giant(s) village' (Marsh 1936: "The Wanderer" LN100)
d. *hį́dúŋe-nąwų*	'mouse + paths' (Marsh 1936: "The Wanderer" LN67)
e. *wanáxi waxòñitʰą*	'spirit/ghost + be holy/sacred' (Davidson 1997)
f. *máyą uhąwe*	'heaven' < 'land + full.of.light' (Davidson 1997; Goodtracks n.d.)
g. *máyą wàtahe*	'Wanderer' < máyą 'land' +wa- 'directional' + dahe 'be standing'
h. *wą́ŋegíhi*	'Chief/Headman' <wąŋe 'man' + gi-'BEN/DAT' + -hi 'CAUS'
i. *wą?kwás'ose*	'warrior/veteran/soldier'< wąŋe 'man' + was'ose[a] 'brave'
j. *wą?šige*	'person' < wąŋe 'man' + šige 'again'+/or -ge 'NOM'
k. *wą?ší k'uč'e*	'man-hunter' < wą?šige 'person' + k'uč'e 'to kill'
l. *tʰà waθlų*	'roasted deer' < deer + to roast (Marsh 1936: "The Wanderer" LN75)
m. *istą čʰi*	'(menstrual) period' (literally 'be alone-house'

[a] Whitman (1947) noted glottal stop marking morpheme juncture. It seems especially prevalent when the deleted sounds/syllable involves /ŋ/.

the Ioways chose the part of a bird that powers its motion to name that revolutionary object, the wheel: *ahu* 'wing' > wheel (wagon/car).[17]

(10) a. wagon = *námąñį* < na 'wood' + mąñį 'moving/walking'
b. train = *námąñį dàk'o* < námąñį 'wagon' + dák'o 'thunder/fire'
c. photographs/pictures = *įje wagaxe* < įje 'face' + wagáxe 'writing'
d. Saturday = *hąwe ukʰìθre* 'day-half' < hąwe 'day' + ukʰíθre 'half, be split into two' [because the Tribal Agency was open from morning to noon on Saturdays]
e. piano = *nayąwe* 'wood sings' < na 'wood' + yąwe 'to sing'

The existing word for 'metal' *mąðe* originally referred to copper, available from the Great Lakes region in particular, and found throughout late Woodland through Mississippian periods in the Mississippi River valley and tributaries. European silver and gold coins were called 'white/light' or 'shiny' metal, *mąðé θka*. The different types of coins led to this unique descriptor for 'penny' < 'coin (white/shiny-metal)+ red' *mąðé θka šùje*. This new unit formed a single compound noun, as shown by the phrase *mąðé θka šùje iyą* 'a penny, one penny.'

[17] Keith Basso described the Western Apache (Athabaskan) words for automobiles in similar ways, but in that case it was a hand/arm = front wheel and foot = rear wheel set extension (1990: 17).

3.1.7 Degrees of noun incorporation

Table 5 demonstrates various ways that the words now functioning as compound verbs are conjugated. The left-most column represents the least degree of noun incorporating into the verb, because the personal pronominal prefixes still attach directly to the verb: [Noun [Pronominal prefix + Verb]]. Or a speaker might prefer to add an auxiliary verb to carry person/number inflections, rather than inflect the main verb; see center column. Finally, a fully fused/incorporated noun-verb lexeme accepts the2 pronominal prefixes attaching directly to the left-most edge of the word, as represented in the far right-hand column [Pronoun + [Noun + Verb]]. The table shows some variation, and speaker preference seems to have been involved. Forms with *ho* 'voice' (11-13) appear to be more fully fused than other nouns were.

There is an intriguing case from another Marsh text in which the noun seems strongly associated with a certain verb but it was in the third person with ∅ affix, so the conjugation pattern is unknown: *tʰá čʼèhi mąñà* 'he went deer-hunting' (Marsh 1936: "The Wanderer" LN47)

3.1.8 Nominalizing prefixes

Certain prefixes commonly attach to verb stems to form a nominal. To illustrate, the three prefixes in (11) all incorporate the basic *wa-* 'indefinite object' (sometimes contracted with a locative prefix also) to action word(s).

(11) a. *wa-*
wagáxe 'paper' < wa- 'INDEF.OBJ' + gaxe 'to scratch, write'
warúwaha 'bundle'< wa- + ruwaha 'to show with hands'

b. *wi-*
wíːų 'tool' < *wa-* 'INDEF.OBJ' + *i-* 'at, to' + *ʔų* 'to do, make, create'
wíːroːha 'kettle' < *wa- + i- + róhą* 'plenty, lots, much, many'
wíːkʰąhį 'bridle' < *wa- + i- + kʰąhį* 'blood-vessel, sinew, cord' *(Marsh 1936:* "The Outsider" *LN65)*[18]

c. *wo-*
wóːčʰexi 'difficult times, trials' < *wa-* 'INDEF.OBJ' + *u-* 'in' + *čʰexi* 'be.cruel/stingy'
wóyąwe 'festivity' < *wa + u- + yąwe* 'sing'*?* [19]

Table 5: Conjugating different verbs with nouns attached

Jiwere gloss	[N+[PRON-V]]	[N+V] PRON-AUX	PRON-[N+V]
1) *hóθige* 'to fish' ['fish + split']	*ho-**he**-θige* '**I** am fishing'	-	-
2) *nąsje pʰiskųñį* 'to be unkind' ['heart be good-not']	*nąsje-**hį**-pʰiskųñį* '**I** am unkind'	-	-
3) *nąsje pʰi* 'to be kind' ['heart be-good']	*nąsje **ri**-pʰi* '**you** are kind'	-	-
4) *nątʼúdą* 'to pity'	*nątʼų-**he**-dą* '**I** pity him'	-	-
5) *irodaxra* 'to have a fever' ['body-burn/be hot']	*iro-**hį**-daxra* '**I** have a fever' iro-**ri**-daxra '**you** have a fever'	*irodaxra **hį**ñį**wi*** '**we** (PL) have a fever' [añį 'have']	-
6) *iroruθʼa* 'to be shaken up, excited' ['body be-pushed?']	***wawa**-roruθʼa**wi*** '**we**'re shook up pl.' (first response) *roruθʼani* '**I** am shook up'	*roruθʼa **hį**ñi-**wi*** '**we**'re shook up' (second response)	-
7) *iroθetʰą* 'to abuse' ['body + ?']	-	-	*i**ri**roθetʰą* '**you** were abused' (1PSG & PL also)
8) *irokʰupʰi* 'to be handsome' ['body +?look+good']	-	*irokʰupʰi **hį**ñiwi* '**we** look good' [<*añi* 'to have']	*i-**ri**-roskʰupʰi* '**you** are handsome' (1PSG also)
9) *rosje* 'to sweat' [<'body+?']	-	*rosje-**ri**-ñe* 'they made **you** sweat' [CAUS]	***wawa**-rosje**wi*** '**we**'re sweating' (1PSG also)
10) *dąwe* 'to awaken, open eyes' [<*isdą* 'eye(s) + move']	-	-	***ha**-dąwe* '**I** awakened'
11) *hohga* 'to belch' [<*ho* 'voice' + sound symbolic *hga*]	-	-	***ra**-hohga* '**you** belched' (1PSG & PL also)
12) *hoxga* 'to hiccup' [<*ho* 'voice' + sound symbolic *xga*]	-	-	***ha**-hoxga mąñi* '**I** am hiccupping'
13) *hoxu* 'to cough'[a] [<*ho* 'voice' + sound symbolic *xu*]	-	-	***ha**-hoxu* 'I coughed.'

[a] Note 12 is lexicalized, as is its Lakota cognate, relative to Biloxi, which treated 'cough' still as separable, inflecting after *ho* 'voice' (Rankin et al. 2003: 186)

3.2 The verb and its many parts: The verb template

In Siouan languages, the most complex morphology involves the verb, which may include basic verb stem, plus up to ten "slots" or positions for a number of possible prefixes, as well as at least four positions for potential suffixes. Figure 1 (at end of chapter) is the representation of all fourteen potential affix positions and which prefix/suffixes can appear in each of those places.

Described in more detail, beginning at the front or left-most position of an inflected verb, the prefixes may occur as follows (Whitman 1947: 246, Marsh 1936, also Hopkins & Furbee 1991). Negative numbers represent positions preceding the verb root; positive numbers follow the root.

Position [-10] 1st-person patient pronouns.

(12) a. *hį =singular* 'me'
b. *wa$_{1a}$- = dual* 'us two' (first half of separable morpheme)

Position [-9] The second *wa*- set.

(13) a. *wa$_{2a}$-* 'them, something'; INDEFINITELY EXTENDED OBJECT (also detransitivizes the verb)
b. *wa$_{2b}$- 'toward,* DIRECTIONAL*' [precedes all person prefixes except hį- 1*SG*. patient* 'me']

Two examples of the first meaning, *wa$_{2a}$-*, give an idea of its flexibility as both derivational and inflectional morpheme:

(14) a. ***wa**naxi* 'spirit, ghost'
< wa$_{2a}$- 'INDEFINITELY EXTENDED OBJECT' + naxi 'breath, life'.
b. *Hinage* ***wa**-t^{h}a* *naha* *waye:re* *na?*
woman **PL.PAT**-1SG.see those.ones who.are3PL Q
'Who are the women that I saw?'

[19] Length is presumed here from the overall language pattern. Marsh rarely marked vowel length in the narratives, except on interjections within dialogue, when they were greatly lengthened for emphasis.

[19] This form and derivation is from Jimm Goodtracks. Marsh (1936: "The Outcast" LN160) gives *wóyawe* with non-nasal /a/, perhaps from to *yawe* 'stab' (which might refer to the preparation of meat for feasting or the the piercing that took place during mourning a chief).

Whitman considered directional *wa-* to parallel both *gra-* and *gi-* of template positions -3 and -4 in some functions.[20] The next case illustrates directional *wa-* frequently found in prayer songs.

(15) *Hįyįno* ***wa**-hį-na-wi.*
our.elder.brother DIR-1PL.AGT-go-DEF.PL
'We're going toward Our Elder Brother (Jesus).' (Davidson 1997)

Position [-8] Locatives.

(16) a. *a-* 'on, upon, over',
b. *u-* 'in, within, into',
c. *i-* 'at, to, by' (Whitman 1947: 241)

The locatives combine with the prefix *wa*$_{2a}$- 'INDEFINITELY EXTENDED OBJECT' to make a "heavy" syllable with a longer vowel, which usually attracts stress (cf. nominal prefixes.)

(17) a. *wa:* < *wa*$_{2a}$- + *a-* 'on'
b. *wo:* < *wa*$_{2a}$- + *u-* 'in'
c. *wi:* < *wa*$_{2a}$- + *i-* 'at, to, by'

Position [-7] Object/patient pronouns.

(18) a. *wa*$_{1b}$- 'us (1PL.PAT; speaker & another, usually listener)'
b. *ri*$_{1}$- 'thee (2SG.PAT)'
c. *mį-* 'me (1SG.PAT)'

Position [-6] Agent pronouns (first and second person).

(19) a. *ha-, he-* 'I/1SG.AGT'
b. *ra*$_{1}$*-, re-* 'thou/2SG.AGT'
c. *a-, e- '3PL.AGT' with motion verbs only*[21]

[20] Cf. Boyle (2009) for a discussion of the *wa-* prefixes across the Siouan languages, quoting the late Carolyn Quintero on Osage *wa-*, which was especially interesting. Based on these analyses it may be more elegant to conclude that in Baxoje-Jiwere there is only one *wa-* which does a wide variety of things to the verb, including the various functions within the different glosses given above. At present, it does not seem crucial to determine whether they are best described as two distinct morphemes *wa-*, or as a single *wa-* quite flexible in meaning. In the future, as more work on comparative Siouan *wa-* emerges, perhaps the issue can be resolved.

[21] Whitman did not list the *e-/a-* prefixes within the ordering of preverbal elements, probably

Position [-5] Reflexive *k^{h}i-* '(to) oneself'. This prefix relates the event/state described by the verb back to the agent, usually translated as 'oneself.' If *k^{h}i* reduplicates, giving *k^{h}ik^{h}i*, it adds the sense of reciprocal action 'to/with each other'.

Position [-4] Possessive *gra-* 'one's own'. The possessive prefix gives additional information about social relations between persons and things mentioned in the verb complex.

(20) Excerpt from the Otoe-Missouria Flag Song:
E-gra-ña-gri-ñe.
*e-**gra**-añi+a–gri-ñe*
3OBJ[ablaut]-POSS-have+3PL-come.back.(home)-PL.INDEF
'They brought it (the flag) back home.' (Greer 2008)

Position [-3] Benefactive/dative *gi$_1$-* 'for, to'.

Position [-2] Instrumentals (describing how an action was completed).

(21)
a. *ba-* 'by cutting'
b. *bo-* 'with a blow'
c. *da-* 'by heat or cold'
d. *gi$_2$-* 'with object away from the body, by pushing or striking with an object'
e. *nq-* 'with foot/feet'
f. *ra$_2$-* 'by mouth, teeth'
g. *ri$_2$-* 'with held object, toward the body, pulling with an object/tool'
h. *ru-* 'with hand, toward oneself, by pulling with the hand'
i. *wa$_3$-* 'with hand away, by pushing with the hand'

According to Whitman (1947: 246), these nine prefixes transform a passive verb into an active one, or a stative verb into a transitive one (Rankin 2005: 483). They make very specific distinctions in the world of human activity. 'Long horizontal object being cut in two' *-gruje* is an interesting yet abstract verbal root;

because they are limited to motion verbs. However, since motion verbs do occur frequently, it seems preferable to include them as possibly archaic forms. The two also occur in 3rd p. possessive pronouns *et^{h}awe* 'his (singular)', *et^{h}ewi* 'theirs (definite pl.)', and *aré* 'it is' (independent pronoun that primarily serves as demonstrative now, loosely 'that').

someone or something must do the cutting, and the various ways that action is accomplished can be encoded very precisely (and concisely) with these prefixes, as in *wa*$_3$- 'with hand away (from agent's body)' *-gruje* > *wagruje* 'to saw'. Siouan scholars have sometimes distinguished between "inner" and "outer" instrumentals, with the latter a smaller set consisting of 'by extreme temperature/heat', 'by cutting with a knife,' and 'by shooting/blowing' (Rankin 2005: 483-485); however, I have not found data pertaining to that distinction in Baxoje-Jiwere thus far.

Position [-1] 2nd person *s*-. Archaic form that stands for 'you' (second person) on a small number of specific verb stems. Siouan scholars have found related forms in the Mississippi Valley subgroup (e.g. Quapaw allomorphs *š-/ž-*), even extending into Proto-Siouan, suggesting it is of ancient origin (Rankin 2005: 479-480). Over time, it was probably replaced in less common verbs by the regular second person forms *ra-, ri-*, but remained in very frequent verbs, which are more resistant to change.

(22) *Arastawi* *k^{h}e*
a-ra-s-da-wi *k^{h}e*
on-you.AGT-archaic.2-see-PL.DEF MASC.DECL
'You (all) see it.' (Final line, Otoe-Missouria Flag Song, Greer 2008)

Position [0] Verb root/stem.

Position [+1] Post-positioned person affixes + causative suffix *-hi* 'to make something happen, to cause something'. One way to form an active verb from a stative one is by adding the causative suffix *-hi*; so *č'e* 'to die' becomes *č'ehi* 'to kill' (literally 'to cause to die'). Since the causative *-hi* occurs after the verb stem, personal pronoun affixes also come after the verb, but immediately before the *-hi*, rather than their usual pre-verbal positions. Sometimes the *-hi* itself is omitted (as in the following example), but the pronominals' marked position after the verb, plus the meaning 'to cause (something)' are still present.[22] The word *nayįhi*

[22] One possible origin of this unusual case of pronominal prefixes shifting to the end is that *hi* was once truly an independent verb, and over time, the forms were re-analyzed by speakers as single unified words. Then the initial verb of the compound was no longer conjugated. In that light, it is interesting to note that there is another *hi*, the motion verb meaning 'arrive here' (Taylor 1976; Hopkins 1988). That would parallel English idioms such as 'to come to pass' for 'happen, take place,' or 'go and X' as in 'Sam went and punched the man'.

'to heal, cure' literally means 'to cause one to stand up, to stand X up.' The chorus of a NAC song by Edward Small (Ioway) exemplifies an instance where *-hi* does not overtly appear. Still, the translation and the location of the PRO prefixes after the verb stem *nayį* 'to stand' give evidence of the causative *-hi* having an underlying presence.

(23) *Hįyį́no | Wak^{h}ąda- yįŋe | mąya čegi wahire nayį-**wa-ra** na*
our.elder.brother | God- son) | land this sick stand-**3PL.PAT-2.AGT**
and
'Elder Brother, Son of God, you heal the sick in this land.' (Davidson 1997)

Likewise, it occurs in this sentence from missionary scholars Hamilton & Irvin (1848: 43: #53):

(24) *Č'e-wa-[∅]-hi k^{h}e.*
kill-3.PL/INDEF.EXT.OBJ-3-CAUS DECL(male.speaker)
'He killed them'.

Position [+2] Negation *-skųñį* 'not'.

Position [+3] General Plural suffix *-ñe* 'they/them. Usually limited to third persons, Whitman (1947) called it an indefinite form; perhaps the term 'general plural' is more appropriate.

Position [+4] Definite Plural *-wi* 'DEF.PL'. Usually 'we' or 'you-all', it may occur with any grammatical person

(25) *wa-wa…-wi, hį-…-wi, ra-…-wi, ri-…-wi, ∅…-wi).*
(1PL-PAT…-PL, 12AGT…-PL, 2AGT…-PL, 2PAT…-PL, 3…PL)

Both suffixes can pluralize any personal pronoun, no matter if that pronoun is in the role of an actor, patient or object (direct or indirect). They only index number, and definiteness vs. indefiniteness. Specifically, it says there are more than one for second and third person forms, and three or more for the first person dual form. The two potential plurals above differ by whether the people or things being referenced represent given or new information.[23]

[23] Think of the parallel indefinite article being used in the formula which begins many English fairy tales, 'Once upon **a** time, there was **a** princess…'

Thus, they are not interchangeable. They reference the speaker's knowledge about the group, how specific group membership is, whether persons' identities are known, if they have already been mentioned in a story before this point or not, and so on. It makes sense for the definite plural to appear with the first person plural for pragmatic reasons. It is difficult to imagine a situation in which 'we' might mean a group with unknown or uncertain membership. Second person plurals also usually take the definite plural, for the same reason, although some rare exception might occur. However, it is very possible to imagine situations involving third persons to be either definite ('the gourd dancers from Red Rock, Oklahoma') or indefinite in nature ('everyone on Earth who knew my uncle'). Just as one might expect, zero third person-inflected verbs occur with either plural suffix, depending on the meaning intended.

(26) *wówak'ųñawi*
wa-∅-u-wa-k'ų-**ña-wi**
12PAT-3PAT-LOC-12PAT-gave-**INDEF.PL**('they')-**DEF.PL**('us')
[vowel ablaut to *ña* from INDEF.PL *-ñe* when before *-wi*]
'They gave it to us.' (Whitman 1947: 240)

Position [+5] Mood/Aspect *-hñe, -hna* 'will, shall'. The modal suffix seems similar to a future tense, but probably is more accurately expressed as 'an action that is not yet completed' according to Rankin (2009).[24] The *e-* itself ablauts to *a-* with verbs of motion.[25]

Position [+6] Evidential and gender indexical particle. It is not clear that these enclitics are actually part of the verbal complex, rather than serving as an audible coda indexing the gender identity of the speaker of an utterance and the degree of certainty of the speaker for the information given. The enclitics are not tied absolutely to the speaker's gender, but may also reflect the gender of a character during dialogue in a narrative, or original speaker's gender in reported speech/quotatives). They do not seem to function as truly "free" morphemes, as they carry only secondary stress, and there is basically no pause between the verbal complex and the sentence-final particle, which tend to form a single prosodic contour. Because it is such a rich and complex set, with meanings that are not

[24] While Rankin (2009) included auxiliary verbs, adverbial intensifiers, positionals, and more within his comparative Siouan post-verbal template, this analysis will not follow his template for those morphological elements at this time.

[25] Comparatively speaking, there is not yet an elegant historical explanation of ablaut across the various members of the Siouan language family (Rankin 2005: 466-468).

easy to gloss, these particles are listed in Table 18, rather than being included in the verbal template per se.

3.3 Auxiliary verbs

Auxiliaries may appear alone, inflected with the full variety of verbal prefixes. When they are not the main verb, they will follow it (and any verbal suffixes attached to it). In third person and inanimate subjects, the auxiliary verbs may not be inflected, but otherwise they would still bear first and second person prefixes, which strongly tend to be animate for practical contextual reasons. The same pattern is found in most SOV languages (Rankin 2005: 490).

Positionals/modals. After the main verb, there is often a second verb expressing the action or position of the agent, or a distinct clause describing the activity/position of the speaker. If one witnessed an event, a proper Baxoje-Jiwere description would include whether someone was sitting, lying, standing, or moving around while it occurred.[26] Beyond clarifying the bodily orientation of the person or thing being described, there are also various aspectual meanings that may be conveyed. One such aspect is continuative as if the action takes place over an extended time frame, rather than occurring at a single moment or limited duration.[27] They include:

(27) a. *mą́ñį* 'going around, moving (in the characteristic way for that creature)'
b. *mį́na* 'sitting /dwelling'
c. *nąŋe* 'be in a sitting position'
d. *háŋe* 'be in a lying or reclining position'
e. *dáhe* 'be in a standing /upright position'
f. *nąyį* 'to stand something/someone up'

[26] Davidson 1997 outlined the key role these auxiliary verbs played in creating vivid images in Native American Church songs composed by Otoe-Missouria and Ioway speakers.

[27] In addition to the continuative aspect, Rankin (2005: 484–485) also distinguished a habitual ('always') aspect (Quapaw *nq*), an imperfective 'used to X' derived from Proto-Siouan /*ʔõ/ 'do', a potential 'will/would X' (Quapaw *tte*). Negation, imperative, and narrative forms were grouped with the auxiliary aspects, too. More complex moods could be created with combinations of these forms, such as potential + continuative, or negative potential continuative 'to not go on X-ing'. However, I have grouped the imperative and narrative particles with the general sentence-final enclitics, in Table 18.

3.4 Pronominals

Baxoje-Jiwere has overt prefixes for first and second persons, while third person is represented by ∅ morpheme. There are also three numbers expressed: one (singular), we two (dual inclusive), or more than two (distinguished by the plural suffixes discussed in §3.2). Each person's role is identified relative to the action of the verb, as agent/actor or patient/object. There is potential confusion caused by homophony between one allomorph of first person singular patient *hį-* 'me' and the first person dual agent *hį-* 'we two'. The other allomorph for 1PAT.SG, *mį-*, mirrors the form in the independent first person pronoun, as well as the independent possessive first-person pronoun. The first person plural can only be expressed by addition of the definite plural suffix *-wi* (see (+4) above), denoting the speaker, hearer, and one or more additional people as either agents *hį-…wi*, or patients *wa-wa-…wi*.

'You' is composed of second person singular agent ra_2- and patient ri_2-, and also second person plural agent and patient forms. See Table 6 for further illustration.

Table 6: Personal pronominal prefixes

	1SG 'I/me'	1DUAL 'we/us two'	1PL 'we/us all'	2SG 'thou/thee'	2pl 'you'
Agent	ha- (he-)	hį-	hį-[+ -wi]	ra- (re-)	ra-[+ -wi] (re-) [+ -wi]
Patient	mį- hį-	wa_{1a}- wa_{1b}-	wa_{1a}-[+ -wi] wa_{1b}-[+ -wi]	ri-	ri-[+ -wi]

The parenthetical forms with final /e/ show the vowel change that takes place when the prefix is followed by certain derivational morphemes such as *gra-* 'one's own' (possessive), represented in the verb *gra-hi* 'to love, have pity on someone'. The agentive forms *ha-* 'I', *ra-* 'thou' will become *he-, re-* in other complex verbs such as *nąt'udą* 'to pity (someone/something)'.[28] A potential origin for this word is *nąhje* 'heart' plus *u-gi-dą* 'be depressed toward' (Whitman 1947: 243). If that analysis is correct, the benefactive prefix *gi-* 'for' would be the conditioning morpheme for that particular case. Another example is *gi-t'ą* '(it) flies', despite the

[28] Whitman (1947) has the plain [u] here while I heard it as a nasal [ų], perhaps just spreading from the surrounding environment (Davidson 1997).

fact that the *gi-* prefix itself only is fully apparent in the plain ∅ third person form (Whitman 1947: 242).

Third person singular is typically marked by a zero morpheme, although an *e-* prefix may rarely occur, especially with the possessor prefix 'one's own', and with independent possessive third-person *et^{h}awe* 'his/hers (singular)' or *et^{h}ewi* 'theirs'. The demonstrative form *-ʔe* combines with many prefixes, including third person *e-*, resulting in *eʔe* 'it is that one.' Motion verbs provide an exception to that rule, with an *a-* prefix in plural contexts.[29] Once again, we see an /a/-/e/ alternation.

Independent pronouns, shown in Table 7, appear for emphasis or clarity, but are not required grammatically to complete a sentence, provided that the verb is properly inflected.

Table 7: Personal pronouns (Hamilton & Irvin 1848; Marsh 1936).

Person	Independent	Possessive
1 Singular	míre	mɨt^{h}áwe
1 Dual Inclusive	híre	hɨt^{h}áwe
1 Plural Inclusive	-	hɨt^{h}éwi
2 Singular	ríre	rit^{h}áwe
2 Plural	-	rit^{h}éwi
3 Singular	éʔe	et^{h}áwe
3 Plural	aré	et^{h}éwi

3.5 Conjugating verbs

3.5.1 Regular verbs

A verb stem is considered REGULAR if it follows the verbal template of prefixes in its ordering, and the stem itself does not change in form, regardless of any shift in person or number. Verbs are grouped according to whether they are ACTIVE or STATIVE, with the agentive pronominal prefixes inflecting the active verbs, and the patient pronominal prefixes forming the subject of stative verbs, as well as the objects of transitive verbs; see Table 8.

[29] Marsh (1936); Taylor (1976).

Table 8: Regular verb paradigm

Person	Active verb	Stative verb	Transitive verb
1SG	*ha-mąñį*	*hį-yą , mį-yą*	*ha-k'e*
	'I walk/move'	'I sleep'	'I dig (it)/ I dug (it)'
1DU.INCL	*hį-mąñį*	*wawa-yą*	*hį-k'e*
	'We 2 walk'	'We 2 sleep'	'We 2 dig (it)'
1PL.DEF	*hį-mąñį-wi*	*wawa-yą-wi*	*hį-k'e-wi*
	'We-all walk (>2)	'We-all sleep'	'We-all dig (it)'
2SG	*ra-mąñį*	*ri-yą*	*ra-k'e*
	'You (sg) walk'	'You (sg) sleep'	'You (sg) dig (it)'
2PL.DEF	*ra-mąñį-wi*	*ri-yą-wi*	*ra-k'e-wi*
	'You-all walk'	'You-all sleep'	'You-all dig (it)'
3SG	*∅-mąñį*	*∅-yą*	*∅-k'e*
	'He/she/it walks'	'He/she/it sleeps'	'He/she/it digs (it)'
3PL.DEF	*∅-mąñį-wi*	*∅-yą-wi*	*∅-k'e-wi*
	'They walk'	'They sleep'	'They dig (it)'
	(known)	(known)	(known)
1SG.INDEF	*∅-mąñį-ñe*	*∅-yą-ñe*	*∅-k'e-ñe*
	'They walk'	'They sleep'	'They dig (it)'
	(unknown)	(unknown)	(unknown)

3.5.2 Irregular verb stems in *d-, r-, w-*.

All irregular verb stems begin with *d-, r-*, or *w-* sounds (Whitman 1947: 243). Note that the stem-initial consonant defines the class, and determines which conjugation will be irregular; however, there may also be prefixes attached to that stem. When any of those prefixes come before the personal pronoun, they do not influence each other (no amalgamation). These irregular verbs share another anomaly; in second person agent forms, in addition to the expected *ra-*, the archaic Siouan second person *s-* also appears (Slot -1 on verbal template). Examples of irregular compound stems include:[30]

[30] Twentieth century elicitations seem to exhibit a tendency toward including the regular pronominal prefixes, in addition to the verb stem changes. However, Dorsey's slip file only has one speaker who doubles the inflection on these forms; this tendency to move toward the regular pattern may reflect the decline in everyday language use, leading to a preference for the most familiar inflections to be added onto the irregular verb stem changes (Dorsey n.d.).

D- stems. The stem's initial /d/ becomes /t/ to indicate 1AGT, instead of having the regular first person agent pronominal *ha-*. The stem change does not occur in any other person; even first person patient constructions take regular 1AGT *hį-*. Second person agentive form is doubly inflected; both 2AGT *ra-* and archaic 1AGT *s-* attach to the stem-initial consonant; see Table 9.

Table 9: D- stem

a-dá 'to see'	á̲ta	'I see (it/him/her)'
	aráṣda	'You(SG) see (it…)'
	há̲da	'We two (1SG & 2SG) see (it,…)'
	há̲dawi	'We (PL) see (it,…)'[a]
	adá	'he sees her/it'
	a̲rí̲ta	'I see you'
	ą́raṣda	'You(SG) -Archaic 2P see me'
	wáwa̲dáwi	'(he) sees us (PL)'

[a] Stress shifted left to reflect a "heavy" syllable resulting from two vowels coalesced together, *hį-* 1PAT plus *a-* 'on' LOC.

R- stems. There are two irregular verb classes beginning with /r/. In the first paradigm, shown in Table 10, the liquid /r/ is followed by back vowels /a, u/, giving *ra-* or *ru-* as the stem's first syllable. First person agent is marked by /r/ becoming /d/. The second person form inflects twice, with regular 2AGT /ra-/ and archaic 2AGT /s, š/.

Table 10: R-stem 1

rumi 'to buy'	ha̲dumi	'I bought (it)'
	ra̲s̲tumi	'You (SG) bought (it)'
	hárumi	'We two bought (it)'
	rumi	'He/she bought (it)'

The second subclass of irregular verb stems, shown in Table 11, begin with /r/ paired with front vowels /i, e/. The /ri-, re-/ verb stems demonstrate a shift from /r/ to /j/ to mark 1AGT forms, while the archaic 2AGT /s, š/ morpheme inflects the unchanged stem alone.

Table 11: R- stem 2

ré ‘to go’	haj̲é	‘I go’
	sre	‘You go’
	hį̨re	‘We two go’
	hį̨rewi	‘We go (pl.)’
	ré	‘He/she/it goes’

W- stems. These verbs have an initial voiceless bilabial glide /w/ which becomes a voiceless aspirated bilabial stop /p/ in the 1AGT form. The regular 2AGT *ra-* may be present with some verbs, but is absent in others, while all W-stems take the archaic 1AGT /s, š/ inflection; see Table 12.

Table 12: W- stem

awáđo ‘to point at, point to’	ápa̲đo	‘I point at (it)’
	aš̲wáđo	‘You point at (it)’
	háwađowi	‘We (PL) point at (it)’
	awáđo	‘He/she/it points at (it)’

Additional verbs may conjugate regularly in all other persons, but preserve the archaic 2AGT /s, š-/ inflection. These mixed verbs include common words: *e* ‘to say, *hijé* ‘reach a standing position’, *áñį* ‘to have’, *hiwé* ‘reach a lying position’, and *dahé* ‘be standing’ (Whitman 1947: 243).

3.5.3 Other special conjugation patterns: motion verbs

Like all Siouan languages, the Baxoje-Jiwere system of motion verbs has a rich set of distinctions. One intriguing dimension is the vertitive, which allows a concise and powerful way of expressing the notion of leaving home or predicting a safe homecoming.[31] Otoe-Missouria patriotic songs often have this powerful motion verb, poetically highlighting the fear involved when soldiers leave home, and joy when they return safely to their families.[32] Motion verbs are also distinguished

[31] While English lacks the motion verb equivalent to the vertitive, the compound noun ‘homecoming’ is perhaps the closest in meaning and emotional power.

[32] Scholars of related Siouan languages such as Assiniboine have also analyzed these verbs in terms of how they appear in traditional narratives, where the notion of ‘belonging’/ home

by a third person plural prefix *a-* which changes to *e-* in the same conditioning environments in which first and second person prefix vowels also shift from /a/ to /e/, namely before the benefactive prefix *gi$_1$-* and the possessive *gra-*. (See Table 13).

Table 13: Jiwere motion verb stems (Taylor 1976: 293)

Destination:	Arriving Motion		Motion Prior to Arrival	
	non-vertitive /	vertitive	non-vertitive /	vertitive
here . . .	jí	grí	hú	gú
there . . .	hí	-	rá	grá

Note the initial consonant cluster echoes the possessive prefix [*gra-* 'one's own']; the shared phonological shape plus semantic congruity between vertitive and possessive is surely no coincidence.

3.6 Adverbials

There are basic adverbial morphemes that may combine to express a wide range of meanings, with parallels to the personal pronouns (both independent and bound) in recognizing not only distinct first and second persons ('I' vs. 'you'), but also 'we two (you and I),' dual inclusive.

3.6.1 Spatial elements

Baxoje-Jiwere identifies five distinct places relating to the discourse context:[33] 1) location of speaker 'my spot here' *ǰe-*, 2) location of listener 'your spot' *se-*,[34] 3) shared area of persons conversing together 'our here' *i-* (location of both you and me), 4) 'there' *ga-*, beyond the immediate discourse zone, e.g. a distant but visible location, and 5) 'place beyond their sight (usually far away) *hari-* (similar to archaic English *yonder*). These spatial elements combine with morphemes that distinguish between a fixed spot close at hand (*-gi*), a stationary spot slightly

location also can be used to mean the place where a person or animal was located at the beginning of the story (by the river/point A), versus where they ended up later on (inside a cave/point B) (Cumberland 2005).

[33] My M.A. thesis details the system of deixis in Baxoje-Jiwere (Hopkins 1988).

[34] This form *se-* with initial /s/ representing second person is very likely related to the archaic 2AGT /s/ found in conjugations of some conservative (irregular) verbs also (Rankin 2005: 480).

further off (*-da* 'at there'), and motion toward a location (*-gu* 'to'). The directional sense of the prefix *wa$_{2b}$-* 'motion toward' may follow first or second person forms to complete the variety of distinctions recognized.

3.6.2 Negatives

Two basic forms can negate the main verb, *skųñį* 'not' and *ñįŋe* '(be/have) nothing'. Thus, while the stative verb *p^{h}i* 'be good' expresses a positive attribute, the opposite meaning results from adding *skųñį* 'not', giving *p^{h}i-skųñį* literally 'good-not'; 'no good, bad, ornery'. At the clausal level there can be additional ways to make it clear that something is false. (Cf. §4 on syntax, especially the evidential enclitics in sentence-final position.)

3.6.3 Time elements

While some Baxoje-Jiwere words for space do apply metaphorically to time, there are also specific temporal adverbs. They tend to occur at the beginning of the sentence, as in this verse from NAC prayer-song composed by the late George Washington Dailey (Otoe-Missouria):

(28) *Go:č^{h}i Hįyįno* *hį-ha-wi-yiyi*
now our.elder.brother(male.spkr) 1PL.AGT-say-PL.DEF-chant
'Oh, My Lord, we're calling upon Your name, now.' (Davidson 1997)

3.7 Other morphological processes

3.7.1 Sound symbolism

In Baxoje-Jiwere, there are two characteristics of such mimetic words that attempt to recreate certain sounds or material aspects of the world:

a. Often they use fricatives, which sometimes form sets of related words which vary only in the fricatives' place of articulation.

b. Many also are stative verbs, especially ones related to topics of color shade, intensity of hue, or other changes in sense perception, as in volume of noise, or roughness of texture.

This phenomenon is common in most Siouan languages, and can create interesting semantic sets differing by a single consonant sound (Rankin 2005: 468-469). The "lighter/less intense" word is usually associated with a front and/or

upper place of articulation, while the greatest intensity of meaning is found with the "deepest" back sounds. It has been documented for Ho-Chunk and Dakota in particular. Baxoje-Jiwere sound symbolic vocabulary sets include those in Table 14.

Table 14: Sound symbolism

šá-kh'e	1) 'swishing sound made in water' 2) 'sound made by hitting or dragging of a chain'
thá-kh'e [probably th = θ]	1) 'rattling of a rattlesnake'; 2) 'rattling of corn in granary or in pile outside'
khó-kh'e	'ripping of calico, roar of falling water, sawing or scraping sound of tool on wood, whizzing of a whirled stick (a bullroarer)' (Dorsey 1892: 3)
to-tó-khe	'repeated sharp sounds, such as the crackling or snapping of twigs and small branches, or frequent gunshots'
tópě	'pattering sound'; *nątótopě* (no gloss given: I posit 'the sound of dancing feet'.)
ʔé-ghe	'the sounds of filing, grating, gnawing, or scratching on metal, bone, hard wood, etc.' (Dorsey 1892: 4–6)[35]
kh'é-ghe,kh'á-ghe,ká-ghe	'crow (bird, n.)' (initial syllable imitating crow's call) (Dorsey 1892: 8)

Note also the terms for upper body noises with variation in the medial fricative: *hohga* 'to belch' [*ho* 'voice' plus sound symbolic *hga*]; *hoxga* 'to hiccup' [*ho* 'voice' plus sound symbolic *xga*].

Although this is not an exhaustive list, let me add my personal favorite, *hé ʔši* 'sneeze', which beautifully imitates the sound of sneezing, and takes an active/agentive conjugation.[36]

3.7.2 Reduplication

Adult/standard reduplication. Another kind of sound symbolism is reduplication, copying part (or all) of a particular word. If a stative verb such as a color is

[35] Dorsey's orthography for consonants retained here.

[36] Dorsey gave Dhegiha *hé-tchį* 'sneeze' (Kwapa *hě-shį*), and 'snore' *zhą-khdhú-de* (1892: 8).

reduplicated, it means the color is scattered here and there (as in patches, spots, stripes), rather than in a continuous or "solid" distribution. For an active verb, it gives an iterative meaning, whereby *gis'é* 'drip' becomes *gis'és'e* 'drip several drops'. For less concrete activity, the reduplication can convey that the verb's action is somehow partial or incomplete. For example, the form *up^{h}a'p^{h}arehi* 'understanding only bits and pieces, imperfectly comprehending' comes from *up^{h}arehi* 'to understand, notice, investigate'.[37] In Baxoje-Jiwere, reduplication seems to have been a very productive process.

Reduplication in baby talk. In addition to adult reduplication, there is also "baby talk" or caretaker speech, a simplified version of ordinary phonological forms. Based on the limited sample available, it appears to have involved producing an exact copy of a monosyllabic morpheme, such as CV-CV. If the word is polysyllabic, then everything after the first syllable would be deleted. Some of the morphemes have been so simplified that it is not always clear from which word the simplified "baby" form originated. However, the primary difference between adult reduplication and "baby talk" is semantic. The latter had no notion of something being repeated or scattered. Caretaker speech must have made it easier for little ones to learn to speak. Perhaps it originated as an adult imitation of the adorable way young children pronounce things themselves. Examples are in Table 15. Other items elicited include the repeated form + the normal diminutive suffix, *-iñe* 'little one': *mamá**į̣ñe*** 'baby' (Ioway), *hahá**į̣ñe*** 'baby colt, horsey' (Davidson 1997).

Table 15: Baby talk reduplication

dáda	'something to eat'
ǰíǰi	'hot (to touch)'
nąną	'something forbidden because of potential danger or pain'
bobo	'penis' abbreviated from *buje* 'acorn cap, penis'

[37] The latter example came from the late Rev. Arthur Lightfoot and Dr. Truman W. Dailey conversing about white missionaries' partial understanding of Indian beliefs at the Missouri Chiwere Language Project in July 1992.

4 Word order/syntax

Baxoje-Jiwere is classified as an SOV language. However, a verb (for third person forms, a "plain" (uninflected) verb) may function as a grammatical sentence,[38] since the independent pronouns are optional, and there is a ∅ third person pronominal prefix corresponding to 'he, she, it.'

4.1 Noun phrases

4.1.1 Adjectival forms

The head noun should come first in the noun phrase, followed by everything that describes it in any way, including stative verbs showing shape, color or size ('large, round, yellow'), which may also inflect as a main verb in other contexts, demonstrating they are not true adjectives.

4.1.2 Determiners, demonstratives, articles and more

Determiners identify which person or thing is being discussed, if it is a specific individual(s) or a generic one, how many there are, and so forth. They include quantifiers, demonstratives, and at least one definite article and an indefinite article, which all follow their "head". So 'a white horse' when spoken in proper Baxoje-Jiwere order would be 'horse white a' *šųñe ska iyą* Ioway / *suŋe θka iyą* Otoe-Missouria. Quantifiers would begin with specific numerals, as well as other words relating to quantity of a group for countable objects and for animate beings ('few, many, all, most, …') or for quantities of mass nouns such as flour, soup, water and so forth ('some, much, little, …').

4.1.3 Article(s)

Indefinite article *-yą, -iyą* 'a, one' is derived from the word for 'one' *iyąkʰi*.[39] Definite article is *-ge*.[40] *Gilbert-ge daniŋe.* '(That) Gilbert was drunk (again)!'

[38] There also needs to be a final particle that tells the gender of the speaker, as well as how certain the speaker is of the information being given, and the way the listener should respond (by listening and talking, by obeying what was said, by joining in with the speaker). These S-final particles are discussed in a later part of the grammar.

[39] Lakota also utilizes the 'one' morpheme as an indefinite article (Ullrich 2008: 755-756).

[40] Until very recently I followed Marsh's (1936) analysis of Baxoje-Jiwere, which included no definite article. I would like to thank Johannes Helmbrecht (2015 p.c.), and Iren Hartmann (2008 p.c.), whose wonderful work on Ho-Chunk and excellent questions about possible cognates in Jiwere have forced me to reconsider the function of *-ge*. I cannot explain how it was overlooked,

While earlier researchers did not identify a definite article for Baxoje-Jiwere, it seems likely that this is an oversight, due to the relatively small amount of data collected, and its lack of frequency compared to the English definite article. There certainly needs to be further examination in this area, considering its complexity in other Siouan languages (Rankin 1977; 2005; Rood & Taylor 1996: 455).

4.1.4 Interrogatives

Those words that are used to ask questions about quantity or number fall into this category.

(29) a. *tahéna* 'how many, how much?' (*tana* in Hamilton & Irvin 1848)
b. *taheda* 'how far?'
c. *danáha, danáhaje* 'which?'[41]

(30) *Bi-rawe tahena ra-gusta ja?*
'moon-count how.many 2P.AGT-want(irreg.verb.2/s-/) Q.FEM
'How many calendars do you want?'

4.1.5 Indefinite quantifiers

Such words give information as to scope, for instance which members of a collective group are included (or excluded) in the utterance. For example:

(31) a. *dáhi, áhi* 'each, every'
b. *bróge* 'all'

Table 16 presents the demonstrative pronouns paired with the corresponding deictic directional prefixes. Note the latter's strong parallels to and semantic association with discourse participants/persons in the context of the speech event.

Aré 'it is' "points" back at something previously mentioned, and appears with great frequency in the texts collected by Gordon Marsh (1936), according to Hopkins & Furbee (1989). It can be paired with the emphatic bound morpheme *-sų* 'indeed' (*aré?sų* 'indeed!' (emphatic)), and even 'stacked' with the first person deictic prefix *je-* 'this (here)' to give *járe* 'this one-it is', and other additional complex compounds.

except that its representation in the data collected was too infrequent to attract notice. More review of the existing data is needed to confirm the current interpretation.

[41] Cf. the similarity of sound shape in the cognate set found in Lakota (Rood & Taylor 1996: 455-457).

Table 16: Comparison of demonstrative pronouns to deictic directional prefixes

Demonstrative Pronouns		Deictic Directional Prefixes (Hopkins 1988)	
jeʔe	'this one'	je-	1 LOC 'near me', 'this here'
seʔe	'this one [near you]'	se-	2 LOC 'near you' [also še- Ioway]
		i-	inclusive 12 LOC 'here'
eʔe	'it is he/ that one'	e-	3 LOC 'near her/him/it'
are	'it is'	a-	*unattested; possible ablaut form of /e/ with *re* 'to go'
gaʔe	'that one'	ga-	'there'

4.2 Subordinate clauses

Main clauses normally occur sentence-finally, while subordinators(s) transform the first clause(s) into a supporting or modifying syntactic role, signaling duration, exact sequence of events, if events were actual or potential, etc. These subordinating particles include *-sge* 'if', *-da* 'when', *-sji* 'but, although', *nuʔa* 'but'. The temporal particle fills that function as follows:

(32) *Hįyįno|* *wo-waxoñitą* *rit^{h}awe urak^{h}i-ñe*
'Our.Elder.Brother| ceremony-sacred your they.tell.about-INDEF.PL
***da**|* *waʔų* *warup^{h}i* *Rire [∅] a-ñe* *(h)na*
when| the.work wonderful(it.does) you 3- say-INDEF.PL Imperfect.'
'Elder Brother, when they tell about Your ceremony and the wonderful work it does, they say it's You.'

This complex sentence begins with a kin tem (addressed to Jesus), a subordinate clause indicated by subordinator *-da* 'when', then finally the main clause (Davidson 1997: Song #16).[42]

[42] Edward Small (Ioway) composed this song after being healed during a NAC worship service.

4.3 Relative clauses

The Baxoje-Jiwere language tends to place the head noun first within the relative clause. An optional special marker *-naha*[43] 'the one(s) that X' immediately follows the clause it acts upon, as in *hinage at^{h}a naha* 'the woman that I saw' (lit. 'woman I saw (her) that one').

(33) a. Relative clause as the object of the sentence: *John hinage at^{h}a*
John woman I.saw.(her)
naha uk^{h}ič'e k^{h}e.
that.one (he)spoke.with.(her) MASC.DECL
'John spoke with the woman that I saw.'

b. Relative clause as the subject of the sentence: *Hinage at^{h}a naha*
Woman I.saw that.one
John uk^{h}ič'e k^{h}e.
John (she).spoke.with.(him) MASC.DECL
'The woman that I saw spoke with John.'

c. Relative clause as the direct object of the verb phrase: *Sam wawagaxe*
Sam book
hapagaxe naha araje k^{h}e.
I.wrote.it that-one (he).read.it MASC.DECL
'Sam read the book that I wrote.'

Because the relative-clause marker is optional, and the 3rd person pronoun is zero, it can be difficult to translate some sentences, even though the general meaning is clear.

4.4 Conjoined clauses

The conjunction *heda* 'and' may occur at the beginning of the second sentence. Within more rapid speech sequences, it is common to instead have the particle *-na* 'and' occur at the end of the first main clause, separating it from the one to come.

[43] Dorsey (n.d.) gave *daha* as another potential relative clause marker, in an example sentence referring to an object rolling under a tent flap that was not fastened down: *t^{h}ą gri were* ***daha,*** *ruθewi re* '**That which** has gone outside, get ye' (spelling and punctuation adapted to modern conventions). Further study on Baxoje-Jiwere demonstratives' potential relationship to positional verbs in a classificatory system is very much needed (Cf. Rankin 2005: 3).

4.5 Beyond statements: Other kinds of sentences

4.5.1 Directives/requests/commands

These ways to "boss" others are linguistically interesting because many languages omit both first person and verb stating 'I am telling you' to do something. Sometimes second person form is also omitted. The "pragmatic skewing" occurs because overt first and second person forms may be considered too direct, and thus rude (Heath 1998). This politeness pattern holds true with Baxoje-Jiwere directives. One speaks to children in a more direct manner than adults, since few question the authority of parents/elders to tell kids what to do. If speaking to an adult, it would be more polite to use a different form, *ne/nɛ*. However, songs demonstrate expressing a plea with the stronger command particle, *re*:

(34) *Hįyįno* *wa-a-wa-da-wi* *re*
Our.Elder.Brother 1PAT-look.at-DEF.PL [command (male speaker)]
'Elder Brother, look at us!' (Davidson 1997)

Finally, one may make a very polite request by using the dual/first person inclusive plural form with hortative enclitic *t^{h}o*. 'Let us all call on the Creator's name', or 'Let's go to the handgame!'

4.5.2 Questions

There are three ways to correctly form questions:

a. Declarative sentence + sentence-final question particle. Word order does not vary; it is an evidential ending particle that signals an answer is expected, because the speaker is asking, not telling something. As with many of these ending particles, the exact form varies by the speaker's gender: *ǰe* 'Q (male speaker)' / *ǰa* 'Q (female speaker)'.

b. By using interrogatives such as *wayé:re* 'who (is it)?' or *dagú:re* 'what (is it)?' The interrogative word receives the special question-sentence melodic contour, which includes lengthening the stressed vowel greatly and making its pitch higher, plus pronouncing the final syllable's pitch lower than usual.

c. Finally, one can create a question by simply omitting all S-final particles, and using the interrogative intonation pattern. See (25b) below. In Ioway/Otoe-Missouria speech, the question pattern is made with a much longer (and

slightly higher pitched) vowel in the penultimate syllable of the sentence, and a drop to a lower pitch in the last syllable.

(35) a. *Wabúθga ra-gústa jà?*
bread 2SG-want(it) Q.female.speaker
'Do you want any bread?'

b. *Wabúθga ra-gú:sta?*
bread 2SG-want(it)
'You want some bread?'

c. *Ra-gústa dagúre?*
2SG-want(it) what(is.it)
'What (do) you want?' or 'You want what?'

5 Variation in speech by social group

5.1 Tribal identity and language use

The Otoe-Missouria and Ioway people spoke mutually intelligible dialects of one language. After a devastating enemy attack in the late 18th century, most surviving Missouria fled from their village in Missouri to the Otoe village in southeast Nebraska (Schweitzer 2001). Geographic separation between these two tribes ended about forty years before any language records exist. Although recognition of a leader, "Missouri Chief," is documented in Indian Territory ca. 1885,[44] there is no data on unique Missouria dialect features.[45] At the phonological level, general tendencies have been noted (Table 17). It is not as simple as always substituting one sound for another, yet listeners certainly noticed the distinctions.[46]

[44] Cf. the diary of Miss Emma DeKnight, who taught at the Otoe tribal boarding school at that time (DeKnight ms., University of Oklahoma Archives, Norman, OK).

[45] J. O. Dorsey identified a tiny bit of data as specifically Missouri, but it related to only a single speaker, so I prefer to avoid any discussion of the Missouri dialect at present.

[46] There has been intermarriage for a long time, so 100% dialect consistency for a speaker would be very unlikely, regardless of tribal membership. Dialects may be a matter of tendencies, rather than always/never. Family members might use different speech within a household, such as Mr. and Mrs. Small, Ioway and Otoe respectively. The couple understood each other but didn't speak exactly the same (Marsh 1936).

Table 17: Dialect differences

I. Phonological variation	Baxoje	Jiwere
A. Difference in fricatives:		
1. Word-initially	[š]	[s]
	šúñe	*súŋe*
	'horse'	'horse'
2. In consonant clusters		
a) before [g]	[sg/hg]	[θg]
	*wa**h**ge*	*wa**θ**ge*
	'dish/plate'	'dish/plate'
	***h**ga*	***θ**ga*
	'(be) white'	'(be) white'
b) before [ǰ]	[ʔǰ/hǰ]	[sǰ]
	*nąʔ**ǰ**e, nąh**ǰ**e*	*nąs**ǰ**e*
	'heart'	'heart'
B. Difference in nasal consonants:		
1. Medial position, esp. before final -e	[ñ]	[ŋ]
	*č^{h}idói**ñ**e*	*č^{h}idói**ŋ**e*
	'little boy'	'little boy'
II. Select lexical differences		
A. Nouns:	*mamáiñe*	*šúwe*
	'little baby'	'little baby'
B. Interjections:	*sik'*	*dąrah*
	'incredible!'	*tan-rah* (Marsh 1936)
		'incredible!'

5.2 Gender-marked speech

Three distinct lexical sets signal speaker's gender.

1. Kinship. The first set is kin terms as outlined in Goodtracks' dictionary. Gender is distinguished not only of referent (mother vs. father, etc.) but certain terms vary by sex of speaker as well, especially siblings' words for each other and words for one's in-laws. Birth order establishes seniority and thereby determines respect relationships, and is reflected in words denoting sons, daughters, and siblings, which served as familial address terms.

2. **Sentence-final particles**. The second set of gender-indexical terms distinguish between declarative statements, requests,[47] commands, dubitatives, quotatives, and more.[48] These important enclitics audibly punctuate the sentence, informing the listener how to interpret the speech segment. (See Table 18. The list may not be exhaustive.) These enclitics occur in combination with each other, especially when expressing emphasis: *kʰe hųʔ* 'Indeed!'('This I declare! male speaker').[49]

Note that a narrator will use the character's gender marker during dialogue, rather than indexing his or her own identity. Based on the songs I collected, while mixed gender singing does occur (females may join in during various worship and powwow songs), it is the men who traditionally begin the songs, and texts reflect that with male forms.

3. **Interjections**. The final morpheme set indexes a speaker's gender, and is usually sentence-initial (See Table 19). It is sometimes only a subtle difference, such as a final vowel shift, while other forms show little apparent derivational relationship between the two forms at all.

Acknowledgments

Various sources of funding have supported this work over the years, from the initial grants which made my graduate work possible as research assistant to Dr. N. Louanna Furbee [NSF "Documenting the Chiwere Language" BNS 88-18398 and 902-1337], APS "Chiwere Oral Traditions" [Jill D. Greer, P.I.], as well as linguistic consulting work with Jimm Goodtracks, P.I. for his dictionary project [NSF "Ioway-Otoe-Missouria Dictionary Project" BNS 0553585].

[47] Earlier scholars have often called the "inclusive request" form in Table 18 the HORTATIVE. marker, related to the rather old-fashioned word to "exhort" someone to do something.

[48] Trechter (1995) presents a thorough analysis of gender enclitics, including the circumstances where a speaker's gender was not the determining factor, for various pragmatic and contextual reasons, including quoted speech.

[49] Dorsey's manuscripts (n.d.) gave the male declaratives as distinguished by tribe, with *kei* as the Otoe form and *ke* as Ioway, while he listed *kʰi* 'Ioway female declarative', but *hą* for Otoe women's equivalent.

Table 18: Sentence-final particles showing mood, evidentiality and gender

S-final particle type	Male speaker	Female speaker
Declarative 1	kʰe	kʰi
Declarative 2	k'a	hą
Completed Action 'not continuing into present' (Dorsey n.d.)		
Inference (2nd-hand source) 'I think'	no	na (?)
Command	re	rɛ, ræ
Polite Command	ne	nɛ
Inclusive request 'Let us …' /'Would that'	tʰo , dáhò, hdaʔo	tʰa
Question marker (optional)	ǰe	ǰa
Tag question	ʔa	kʔa
Narrative marker 'It seems'	asgų	asgų
Quotative	ʔe	ʔɛ
Emphatic	hųʔ	æ, ʔa, ʔ
Surprise/excitement 'Exclamation point!' (Dorsey n.d.)	t'o	t'ų:

Table 19: Interjections showing mood and gender

Interjection gloss	Male speaker	Female speaker
'Oh, my!' (Pity, love, sympathy, compassion)	hé:hą	inà:, hina:
'Say! Hey!' (Change subject)	kàró	unknown
Joy, Happiness (while singing or talking)	íyà	íyà[a]
Greeting/Acknowledgement, Thank you! Approval/Sanction	ahó, hó	ahá, há
'Hmph! Aw, Heck!' (critical/doubtful; prior speaker isn't telling it right)	dɛʔ[b]	hɛʔ
'Well! (GT) Whew!' (Almost!; something nearly happened, but didn't, either good or bad)	gwí, kwí	hí
'Well, well [Whitman]; Oh, my!' (negative response, as in niece/nephew teasing uncle/aunt too harshly; surprised in a bad way)	hé:hą[c]	háraʔ
'My goodness! Surely not! No way!' (Negative response; surprise, shock)	báʔ, hubaʔ, húʔ húbaʔ (L-R in order of increasing emphasis)	dóʔ, dóʔò (greater emphasis) ga: (Rankin 2009)
'Yes' (Affirmative)	húǰé	húǰɛ̀
'No' (Negative)	hiñégo	hiñéga

[a] Not traditionally female but some use it now.
[b] Both male and female forms = short vowel [ɛ].
[c] Also glossed as doubting truth

Abbreviations

1, 2, 3	first, second, third person
12	first+second person (first person dual)
AGT	agent
AUX	auxiliary
BEN	benefactive
CAUS	causative
DAT	dative
DECL	declarative
DEF	definite
DIM	diminutive
DIR	directional
DU	dual
EMPH	emphatic
EXT	extended
FaBr	father's brother
FEM	feminine
HORT	hortative
IMP	imperative
INDEF	indefinite
INCL	inclusive
LOC	locative
MASC	masculine
MoZ	mother's sister
NOM	nominalizer
OBJ	object
ORD	ordinal
PAT	patient
PL	plural
POSS	possessive
PRON	pronoun
Q	question
SG	singular

References

Basso, Keith H. 1990. Semantic aspects of linguistic acculturation. In Keith H. Basso (ed.), *Western Apache language and culture: Essays in linguistic anthropology*, 15–24. Tucson: University of Arizona Press.

Boyle, John P. 2009. *The indefinite prefix [*wa-] in the Siouan languages*. Paper presented at the Siouan and Caddoan Languages Conference, University of Nebraska – Lincoln, June 2009.

Brown, Cecil H. 1999. *Lexical acculturation in native American languages* (Oxford Studies in Anthropological Linguistics 20). Oxford: Oxford University Press.

Cumberland, Linda A. 2005. *A grammar of Assiniboine: A Siouan language of the northern Plains*. Bloomington: Indiana University (Doctoral dissertation).

Davidson, Jill D. 1997. *Prayer-songs to our elder brother: Native American Church songs of the Otoe-Missouria and Ioway*. Columbia: University of Missouri (Doctoral dissertation).

de Reuse, Willem. 1994. Noun incorporation in Lakota (Siouan). *International Journal of American Linguistics* 60. 199–260.

Dorsey, James Owen. 1892. Siouan onomatopes. *American Anthropologist* 5(1). 1–8.

Dorsey, James Owen. n.d. *Chiwere notes*. Box 158 items 292–311, Oversize BAE items 312–313, National Anthropological Archives. Washington: Smithsonian Institution.

Goddard, Ives. 1996. Introduction. In William C. Sturtevant & Ives Goddard (eds.), *Languages* (Handbook of North American Indians 17), 1–16. Washington: Smithsonian Institution.

Goodtracks, Jimm G. (ed.). 1992 – present. *Baxoje, Jiwere-Nyut^aji, Ma^unke Ich^e Wawagaxe: Ioway, Otoe-Missouria language to English dictionary*. Boulder, CO: Center for the Study of Native Languages of the Plains & Southwest. http://iowayotoelang.nativeweb.org.

Goodtracks, Jimm G. n.d. *Welcome to the Ioway, Otoe-Missouria language website*. Topeka, KS: Baxoje Jiwere Ich^e – Ioway Otoe-Missouria Language Project.

Graczyk, Randolph Jan. 1991. *Incorporation and cliticization in Crow morphosyntax*. Chicago: University of Chicago (Doctoral dissertation).

Greer, Jill D. 2008. The Otoe-Missouria Flag Song. In Jie Zhang (ed.), *Proceedings of the 2007 Mid-America Linguistics Conference* (Kansas Working Papers in Linguistics 30). Lawrence: University of Kansas. https://kuscholarworks.ku.edu/bitstream/handle/1808/3907/kwpl-30-Greer.pdf.

Hale, Kenneth & Josie White Eagle. 1980. A preliminary metrical account of Winnebago accent. *International Journal of American Linguistics* 46. 17–34.

Hamilton, Rev. William & Samuel Irvin. 1848. *An Ioway grammar*. Highland, KS: Presbyterian Mission Press.

Heath, Jeffrey. 1998. Pragmatic skewing in 1 ↔ 2 pronominal combinations in Native American languages. *International Journal of American Linguistics* 64(2). 83–104.

Helmbrecht, Johannes. 2002. Nouns and verbs in Hocank (winnebago). *International Journal of American Linguistics* 68(1). 1–27.

Helmbrecht, Johannes & Christian Lehmann (eds.). 2010. *Elements of grammar, learner's dictionary* (Hocąk Teaching Materials 1). Albany: State University of New York – Albany Press.

Hopkins, Jill D. 1988. *Deixis in Chiwere*. Columbia: University of Missouri – Columbia (Master's thesis).

Hopkins, Jill D. & Louanna Furbee. 1989. *Discourse markers in Chiwere*. Paper presented at the 1989 Mid-America Linguistics Conference, Oklahoma State University, Stillwater, OK.

Hopkins, Jill D. & Louanna Furbee. 1991. *The Chiwere Siouan verb-word*. Paper presented at the Conference on American Indian Languages held concurrently with the annual meeting of the American Anthropological Association, November 1991, Chicago.

Larson, Rory. 2004. *Loan words*. Paper presented at the Siouan and Caddoan Languages Conference, Kaw Nation Tribal Center, Kaw City, OK, June 2004.

Marsh, Gordon H. 1936. *Materials for a study of the Ioway language*. Manuscript 4178. Philadelphia: American Philosophical Society.

Miner, Kenneth L. 1979. Dorsey's law in Winnebago-Chiwere and Winnebago accent. *International Journal of American Linguistics* 45(1). 25–33.

Rankin, Robert L. 1977. From verb, to auxiliary, to noun classifier and definite article: Grammaticalization of the Siouan verbs 'sit', 'stand', 'lie'. In Robert L. Brown, Kathleen Houlihan, Larry G. Hutchinson & Andrew MacLeish (eds.), *Proceedings of the 1976 Mid-America Linguistics Conference*, 273–83. Minneapolis: University of Minnesota Linguistics Department.

Rankin, Robert L. 2005. Quapaw. In Heather Hardy & Janine Scancarelli (eds.), *The native languages of the southeastern United States*, 454–498. Lincoln: University of Nebraska Press.

Rankin, Robert L. 2009. *The Siouan verb template continued: Enclitics – a comparative and historical survey*. Presented at the Siouan and Caddoan Languages Conference, University of Nebraska – Lincoln, June 2009.

Rankin, Robert L., John P. Boyle, Randolph Jan Graczyk & John E. Koontz. 2003. A synchronic and diachronic perspective on 'word' in Siouan. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A cross-linguistic typology*, 180–204. Cambridge: Cambridge University Press.

Rood, David S. 1979. Siouan. In Lyle Campbell & Marianne Mithun (eds.), *The languages of Native America: Historical and comparative assessment*, 236–298. Austin: University of Texas Press.

Rood, David S. & Allan R. Taylor. 1996. Sketch of Lakhota, a Siouan language. In William C. Sturtevant & Ives Goddard (eds.), *Languages* (Handbook of North American Indians 17), 440–482. Washington: Smithsonian Institution.

Schweitzer, Marjorie. 2001. Otoe and Missouria. In Raymond J. DeMallie & William C. Sturtevant (eds.), *Plains*, vol. 13 (Handbook of North American Indians), 447–461. Washington: Smithsonian Institution.

Stanley, Lori A. 1993. *The Indian path of life: A life history of Truman Washington Dailey of the Otoe-Missouria tribe*. Columbia: University of Missouri.

Susman, Amelia L. 1943. *The accentual system of Winnebago*. New York: Columbia University (Doctoral dissertation).

Taylor, Allan. 1976. On verbs of motion in Siouan languages. *International Journal of American Linguistics* 42(4). 287–296.
Trechter, Sara. 1995. *The pragmatic functions of gendered clitics in Lakhota*. Lawrence: University of Kansas (Doctoral dissertation).
Ullrich, Jan F. (ed.). 2008. *New Lakota dictionary*. 1st edn. Bloomington, IN: Lakota Language Consortium. http://www.lakotadictionary.org/nldo.php.
Wedel, Mildred W. 2001. Iowa. In Raymond J. DeMallie & William C. Sturtevant (eds.), *Plains*, vol. 13 (Handbook of North American Indians), 432–446. Washington: Smithsonian Institution.
Whitman, William. 1936. *The Oto*. New York: Columbia University Press.
Whitman, William. 1947. Descriptive grammar of Ioway-Oto. *International Journal of American Linguistics* 13(4). 233–48.
Wistrand-Robinson, Lila. 1972. *An Iowa/Otoe-English dictionary*. Unpublished manuscript. Lawrence: University of Kansas.

-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5
PRO	INDEF	LOC	PAT OBJ	AGT Actor	REF	POS	DAT BEN	INS	Arch 2Pers	VERB STEM	CAU	NEG	PL1	PL2	ASP MOOD
hį- 12AGT	*wa*$_{2a}$- 'them, something;	*a-* 'on'	*hį-* 1SG	*ha-*, *he-* 1SG	*k*^{h}*i-* 'self'	*gra-* 'one's own'	*gi-* 'for, to'	*ba-* 'by cutting'	*s-*		*-hi* 'make, cause' [+pers. affixes]	*=sgúñį* 'not'	*=ñe* 'gen' [Whitman's indefinite]	*=wi* 'definite'	*=hna* 'future, incompletive'
wa$_{1a}$- 1PL.PAT	INDEFINITELY EXTENDED OBJECT	*i-* 'at'	*wa*$_{1b}$-	*ra-*, *re-* 2SG				*bo-* 'with a blow'							
	wa$_{2b}$- 'toward', directional	*u-* 'in'	*ri*$_{1}$- 2SG	*a-* 3PL motion verbs				*da-* 'by heat or cold'							
								gi$_{2}$- 'with obj. away from self, pushing with something'							
								nąą 'by foot'							
								ra$_{2}$- 'by mouth/teeth'							
								ri$_{2}$- 'with obj. moving toward self, pulling with something'							
								ru- 'with hand, toward self, pulling'							
								wa$_{3}$- 'with hand away, by pushing with hand'							

Figure 1: Verbal template: Prefix slots in order = verb = suffix slots

Part III

Analyses of individual Siouan languages

Introduction to Part III

Siouan languages have received relatively little attention in general linguistic theory, and so pose a challenge even to theories based on typological generalizations from broad selections of the world's languages. The head-final, partially polysynthetic nature of Siouan languages raises issues for some claims in syntactic theory, and interesting issues arise in phonology as well. This section of the volume comprises five chapters applying formal linguistic theory to problems in the phonological or syntactic structure of a single Siouan language. (Cross-linguistic studies are in the next section, Part IV.)

David Rood ("The phonology of Lakota voiced stops") reexamines a longstanding problem, the phonological status of voiced stops in Lakota. He proposes a new analysis drawing on autosegments and feature geometry to account for the sonorant-like behavior of /b/ and /g/ in lenition, nasalization, and cluster contexts, and concludes that Lakota is not a voicing language but an aspiration language.

John Boyle ("The syntax and semantics of internally headed relative clauses in Hidatsa") analyzes the internally-headed-relative-clause construction in Hidatsa, from both formal syntactic and formal semantic perspectives. Working within the Minimalist framework, he demonstrates that Hidatsa IHRC are nominalized clauses; using Heim's framework, he then presents a formal semantic explanation for the well-known indefiniteness restriction on the heads of IHRC.

This section of the volume concludes with three interrelated chapters with overlapping authors, all dealing with Ho-Chunk (Hocąk)[1] syntax, especially the existence and structure of verb and adjective phrases (VP and AP) in the language.

Meredith Johnson ("A description of verb-phrase ellipsis in Hocąk") demonstrates that Ho-Chunk does have true verb-phrase ellipsisellipsis, verb-phrase, with cross-linguistically typical characteristics. This argues strongly for the existence of VP in Ho-Chunk.

[1] As discussed in the volume preface, numerous spellings exist for the name of this language. *Ho-Chunk* is the usage on the Tribe's web site, but the three authors in this section all opted to use the spelling *Hocąk*.

Bryan Rosen ("On the structure and constituency of Hocąk resultatives") continues the theme of arguing that Ho-Chunk has a full range of syntactic categories, this time including both VP and AP. The claim that Ho-Chunk has adjectives is controversial: nearly all work on Ho-Chunk and other Siouan languages argues or assumes that these words are stative verbs.

Meredith Johnson, Bryan Rosen and Mateja Schuck ("Evidence for a VP constituent in Hocąk") rounds out this part of the volume by cataloguing the arguments for a configurational analysis of Ho-Chunk (and, by extension, other Siouan languages as well). Subjects and objects are shown to behave differently with respect to a number of tests, including scope as well as the elliptical and resultative constructions discussed in the previous two chapters.

Chapter 10

The phonology of Lakota voiced stops

David Rood

Lakota has two phonetic voiced stops, [b] and [g], which do not behave like stops. They occur in clusters with sonorants, and in syllable or morpheme final position as the result of several morphophonemic processes. In the clusters, their behavior is easily described using the distinctive feature [SONORANT VOICE] proposed in Rice (1993), but in coda position they (and the fricatives, and the lateral sonorant [l]) manifest lenition instead. In this lenition environment, both voicing and nasalization are variable, leading to the proposal that Lakota is an aspiration language like English rather than a voicing language like French.

Look at almost any written Lakota text or dictionary, and you will see many instances of the letters <b> and <g>. Look at any Lakota grammar, and you will learn that these letters (for the most part) do not represent contrastive sounds in the language — they represent allophones of /p/ or /ph/ and /k/ respectively. Yet everyone who writes the language uses both <b> and <g> to some extent, so they must seem salient to those who speak, study or record the language. What is their status, really?

To investigate this question, the following discussion will call on phonological theory (particularly autosegmental phonology and feature geometry; see Clements 1985), the concept of two kinds of phonological [VOICE] features (Rice 1993), the typology of the way languages realize the feature [VOICE] (i.e. the difference between aspiration languages and voicing languages – Ringen, Beckman & Jessen 2013), and the history of the language. Also relevant, it turns out, are some of the unusual properties of the feature [NASAL] in this language (and in Siouan in general). Combining these tools gives us a consistent picture of how these voiced stops function in Lakota: they are sonorants, or more precisely, sonorant obstruents, in an aspiration language. Rankin (2001), an unpublished conference paper, examined much of the same data and came to similar conclusions based on historical and comparative evidence and argumentation. I will try to summarize his arguments in parallel with mine in what follows.

David Rood. 2016. The phonology of Lakota voiced stops. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 233–255. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.173 DOI:10.17169/langsci.b94.173

1 Orthographic and structural preliminaries

Before tackling the investigation, however, we need some facts about Lakota and its orthography in order to understand the data. All examples are presented in the orthography of the *New Lakota Dictionary* (Ullrich 2012), with the modification that we sometimes need to use <ʔ> to illustrate a point. The Lakota consonant inventory is presented in this orthography in Table 1. The table shows only one category of aspirated stops, however, though the language uses two kinds of aspiration that are not quite in complementary distribution. Aspiration on stops may therefore sometimes be represented by <ȟ> instead of <h>. Likewise not shown in this table is the symbol <ŋ>, which marks nasalization of a preceding vowel. Occasionally, however, we need to represent phonetically the velar nasal for which the IPA symbol is exactly this one. In those cases, the symbol is enclosed in square brackets, viz. [ŋ].

Another important feature of the presentation of the examples is the so-called ablauting vowel. Some Lakota verbs change their final vowel in certain grammatical contexts. The options are [a, aŋ, e, iŋ]. To indicate that an example is a verb from this category, it is traditional in Lakota lexicography to present the word with the final vowel <A> or <Aŋ>.

We will sometimes have occasion to cite reduplicated forms in what follows. The syllable which is copied cannot be predicted for most words, and whether the copied syllable is inserted before or after its model is also variable.

Finally, since we are discussing phenomena that are only marginally phonemic, if at all, we will often need to indicate the systematic status of a consonant. I have used the traditional slashes (/…/) for underlying or phonemic sounds, and square brackets ([…]) for actual pronunciations. In Table 1, (ʔ) represents the glottal stop, which the standard orthography assumes is predictable between vowels and optional in word-initial position, but which may sometimes be structurally important, if not actually phonemic, in morpheme-initial position, as we will see below. Letters in angle brackets (<…>) are letters, not sounds.

2 The patterns

The letters we want to study are of course meant to indicate voiced stops. They occur in two patterns: as the first members of clusters, and in morpheme final position when an underlying morpheme final vowel is lost. The loss of the vowel restructures the syllables of the word, putting the formerly preceding consonant into the coda position of the formerly penultimate syllable.

Table 1: Lakota consonant inventory using the standard orthography

	Labial	Alveolar	Alveo-palatal	Velar	Laryngeal
Unaspirated stop or affricate	p	t	č	k	(ʔ)
Ejective	p’	t’	č’	k’	
Aspirated	ph	th	čh	kh	
Voiced stop	b			g	
Voiceless fricative		s	š	ȟ	
Voiced fricative		z	ž	ǧ	
Lateral		l			
Nasal	m	n			
Glide	w		y		h

2.1 The distribution of the voiced stops

2.1.1 [b]

The stops in question occur both singly and in clusters. The only cluster including [b] is [bl], but it enjoys very high frequency because it is the first-person agent marker for the large number of common verbs which begin with /y/ (the [l] replaces the /y/). Singly, the sound represented by the letter <b> occurs contrastively in just one borrowed word, one probably onomatopoetic word, three roots, and one conjugated verb:[1]

(1) Words with phonemic /b/
   a. *bébela* ‘baby’
   b. *škíbibila*‘black capped chickadee’
   c. *bá* ‘to blame’ (not widely known)
   d. *ábela* ‘scattered’ (with several derivations)
   e. *kabú* ‘to play the drum’ (*ka-* ‘by hitting’; *bu* ‘make a hollow noise’)
   f. *wahíbu* ‘I left to come’ (archaic) *wa* ‘I’; *hiyu* ‘to start coming’; *y>b* ‘I’)

The borrowed word, presumably from French, is *bébela* ‘baby’ (*-la* is the regular Lakota diminutive). The possibly onomatopoetic bird name, *škíbibila*, is also

[1] Thanks to the anonymous reviewer for reminding me of some of these.

recorded *škípipila*, and the form with the plain voiceless stop is actually the preferred one, according to the dictionary. The roots are *bá* 'to blame, to criticize', *-be-*, found only in adverbs having to do with being scattered, and *bú* 'to make a loud, hollow noise'. The first is seldom used; the second is also recorded *pe* in the older Buechel dictionary (Buechel 1970: 278 under the entry *kaibebekiya*); and the last is used with several derivational prefixes such as *ka-* 'by hitting' in *kabú* 'to play the drum'. Formerly, [b] also occurred as part of the first-person agent marking for the verb *hiyú* 'to start coming', which used to be doubly conjugated, *wahíbu*; today the form is regularized to *wahíyu*.

Rankin (2001: 8) asserts that [b] is a regular allophone of /w/ in that only the former occurs before /u/: we have syllables [wi], [we], [wa], [wo], but [bu]. However, he gives no examples of this [bu], only examples of clusters ending in /w/ (gwu, swu, šwu, ȟwu) where the first element of the cluster prevents the allophone rule from applying.[2] The anonymous reviewer for this chapter also pointed out that [b] occurs in some compounds where we expect [bʔ]. The example cited is *hítobuya* 'four-year-old calf or colt' contracted from *tópa* 'four' and *ʔuyÁ* 'to grow'. The voiced stop is expected if the glottal stop is pronounced, and sometimes voicing remains even if the glottal stop is lost. Cf. discussion of the interaction of voiced stops with glottal stops in §4.2 below.

In all its other occurrences, [b] is in complementary distribution with /p/ and /ph/. In what follows, I will ignore phonemic /b/ (in words like those in example (2.1.1)), since it does not alternate with anything and can be considered an ordinary voiced obstruent. The feature [LARYNGEAL VOICE] is at work here as it is with the fricatives.

We will also see that the [bl] cluster can absorb nasalization from a following vowel and become [mn], another phenomenon that we need to account for.

2.1.2 [g]

The sound represented by <g> is a little different. It is never contrastive, and in non-morpheme-final contexts we find the clusters [gl, gw, gm] and [gn]. A reviewer for this volume as a whole pointed out that there are also examples of the cluster [gž] in reduplications where the [g] represents an underlying /l/,

[2] In the closely related Dakota dialect, the behavior of sonorants is different from what we find in Lakota. For example, the clusters /sw/, etc. are [sb] etc. there. I do not know whether [b] or [w] is older. Although the reviewer suggested I discuss the Dakota patterns as well, it seems to me they are different enough to provide more distraction than enlightenment. For example, corresponding to Lakota [gl] we find Dakota [hd]. There is a long and complicated debate in the phonological literature over whether or not [h] is a sonorant, to which these Dakota data can contribute, but that is a topic for another paper.

following a pattern known as coronal dissimilation discovered by Carter (1974: 225–226) and also discussed by Shaw (1980: 338). Since this [g] presumably inherits its voice from the underlying /l/, the voicing here must be [SV]. See examples in (2), with the reduplication enclosed in angle brackets.

(2) a. *waŋ<žíg>žila* 'one at a time' (reduplication of *waŋžíla* 'one')
b. *o<žúg>žula* 'several to be full', reduplication of *ožúla* 'to be full')

Neither [k] nor [kh] ever occurs in obstruent-sonorant clusters.

Another difference between [g] and [b] is that for unknown reasons [g] never nasalizes in clusters, though /l/ may, changing [gl] to [gn]. Both [gl] and [gn] are frequently the result of a /k/ morpheme occurring before and combining with an initial /y/ of a verb.

(3) Examples of the clusters beginning with [g] and [b]
a. *Oglála* 'Oglala'
b. *wagmú* 'squash', lit. 'curved thing' (*wa-* 'indefinite patient'; *gmu* 'curved')
c. *gnáyaŋ* 'to fool, play tricks on, deceive'
d. *yugwá* 'to shell, as corn; to open vegetable pods' (*yu-* 'by hand'; *gwa* '??')
e. *obláye* 'prairie'
f. *waŋblí* 'eagle'
g. *blé* 'lake'

2.2 Fricatives

As we saw in Table 1, both voiced and voiceless fricatives are phonemic. Examples are in (4). This voice contrast is not reconstructable to Proto-Siouan, but is limited to the Central Siouan languages. Exactly how it arose is unknown. Miner (1979) pointed out that voiced stops and voiceless fricatives form a single phonological class in Dakotan, Dhegiha, Ho-Chunk and Chiwere (i.e. all the Central Siouan languages) for purposes of forming clusters with sonorants.

(4) Voice contrasts in Lakota fricatives
a. *zí* 'yellow' vs. *sí* 'foot'
b. *ǧú* 'burned' vs. *ȟé* 'mountain'
c. *waŋží* 'one' vs. *waší* 'I ordered him to'

Rankin (2001: 4) summarizes what I, too, believe about the fricatives:

> The underlying status of Dakotan root-final fricatives has received a lot of, in my opinion, undeserved attention in the literature. Proto-Siouan, as far as we can tell, had only voiceless fricatives. Present-day voiced fricatives have several sources, not all of them regular, for example the initial /z/ of záptaŋ 'five' which corresponds with /s/ virtually everywhere else in Siouan. The definitive study of voiced fricative origins has yet to be done. One of the primary sources of voiced fricatives in Mississippi Valley Siouan languages including Dakota seems to have been post-accentual voiceless fricatives, i.e.,
>
> s š x̌ > z ž ǧ / V′ __
>
> It is the case, however, that for one reason or another, voiced and voiceless fricatives do contrast in this environment, so, underlyingly, I would leave the fricatives as they are on the surface. In any event, the fricatives in and of themselves do not present any special problems; even if we consider them underlyingly voiced, as do Carter and, at least originally, Shaw, we simply posit a syllable-final (also word-final) devoicing rule of the sort familiar from German, Russian and a host of other garden variety languages. Syllable-final devoicing is also extremely common — only slightly less so than word-final devoicing. So much for the fricatives: they merely do what is expected of obstruents. The real problem is the stops anyway.

2.3 Morpheme-final position or coda position

Lakota nouns and verbs always end in a vowel in their underlying forms. However, there are some very frequently occurring contexts in which the final vowel is lost, often indicating a subordinate or dependent relationship with a following word, or simply as a consequence of fast speech. When the vowel is lost, the consonant preceding it changes in systematic ways: voiced fricatives are devoiced, as we just learned, and labial and velar stops are at least partially voiced. The coronal stop /t/ becomes [l], and the coronal affricate, /č/, changes either to [l] or to [g] depending on the word. Often roots ending with /č/ have some consonant final forms with [l] and others with [g] (see (5)). In what follows, we will try to show that these changes are all manifestations of a single process: morpheme- or syllable-final lenition.

(5) Variable (lexically conditioned) change in morpheme-final /č/
   a. *šičA* 'bad' reduplicates as ***šig****šíča*
   b. *šičA* 'bad' + *yá* 'causative' > ***šil****yá*

The process applies to either plain or aspirated labial stops (I have no aspirated velar examples). Examples of the plain stops abound in what follows, but the aspirated one is rare:

(6) Voicing of underlying aspirated labial
   a. *hakáb* < *hakápȟa* 'youngest; last in line' (*haká* 'behind'; *-pȟa* 'adverbial')
   b. *sáŋm* < *sáŋpȟa* 'beyond; more than' (*saŋ* '?'; *-pȟa* 'adverbial')

All the voiced non-fricative consonants (viz. [b, g, l]) may become nasals if the preceding vowel is nasalized. (Recall that the nasal spread changing /bl/ to [mn] mentioned above was conditioned by a *following* nasal vowel.) The resulting nasals may or may not cease the nasalization before the release of the stop and thus terminate in a non-nasalized voiced segment, i.e. one may hear [b], [^{m}b], [m^{b}] or [m], [g], [^{ŋ}g], [ŋ^{g}] or [ŋ], but almost always [l] or [n] instead of [nl]. There is also great deal of variation in the amount of voicing in this context, a fact that is reflected by the vacillation native speakers demonstrate when writing. Ella Deloria, a native speaker and prolific recorder of Lakota texts in the mid 20th century, for example, always writes the stops as voiceless unless they are nasalized, while other writers may elect to use the letter representing the voiced stop. The New Lakota Dictionary writes some of them voiced, some voiceless. More details about this phenomenon will be included with the analysis of it presented below.

The grammatical contexts for this vowel loss plus consonant modification process include compounding, building of tightly knit phrases which are not necessarily compounds (e.g. serial verb constructions), and the left-hand copy in a reduplicated form. There is also an adverbial derivation of most verbs, used similarly to an English present participial complement, which is derived by deleting the verb-final vowel. Example (7) illustrates both vowel loss in compounding when *šúŋka* becomes *šúŋk-* (lack of voicing in /k/ is optional; see (8c) below) and this participial derivation:

(7) Derivational verb final vowel deletion
   šuŋk'ákaŋyaŋkA 'to ride horseback' > *šuŋk'ákaŋyaŋg* 'on horseback, riding a horse' (*šuŋka* 'horse' + *ʔakaŋ* 'on' + *yaŋkA* 'to sit')

(8) More examples of morpheme-final consonant alternations
   a. Fricatives devoice:
      mas'óphiye 'store' (< *máza* 'metal' + *óphiye* 'box' (literally 'cash box'))

luslúzahe 'swift things' (reduplication of *lúzahe* 'be swift')
čhaȟ'íčazo 'ice skates' (*čháǧa* 'ice' + *i* 'instrumental' + *kazó* 'draw a line')

b. Stops optionally voice (coronals become [l]):
sabsápa or *sapsápa* 'be black' (reduplication of *sápa*)?
šilyá 'to cause something to spoil' (*šíčA* 'be bad' + *yA* 'cause')
tágni 'nothing' (*táku* 'something, anything' + *-ni* 'NEG')
maníl 'out in the wilderness' < *mánitu* 'to be wilderness'

c. Stops nasalize ([ŋ] in square brackets represents the velar nasal; parentheses mark optionality):
šuŋ[ŋ](g)'ákaŋyaŋkA 'ride horseback' (*šúŋka* 'horse' + *akáŋ* 'on' + *yaŋkÁ* 'sit') (also pronounced *šuŋk'ákaŋyaŋke*; see section 4.2 below for explanation)
nuŋmnúŋpa 'two at a time' (reduplication of *núŋpa* 'two')
waŋyáŋ[ŋ](g) wahí 'I have come to see him' (*waŋyáŋkA* 'to see', *wa-* 'I' + *hí* 'arrive coming')

Taking the differences between [b] and [g] into account, we have four contexts to examine: [g] in clusters, [g] in morpheme final position, [b] in the cluster [bl], and [b] in morpheme final position. I will discuss these four contexts in what follows. First, however, I wish to establish the theoretical principles which I apply to the analysis.

2.4 The theoretical model

The model which I will use to describe this behavior is a somewhat eclectic version of feature geometry in which phonological segments are specified only for the marked value of unpredictable features. If a feature is absent or negative in a given sound, it is simply not mentioned until the end of the phonological part of the grammar, when predictable elements are inserted (see Rice 1993 and references there, or Botma 2011, for examples of these proposals in action). As the anonymous reviewer pointed out, this form of underspecification combined with feature geometry is controversial among phonologists, but it seems to work well for our purposes here.

The features themselves are arranged in a hierarchical tree; each labeled place in the tree is called a node. Some of the features are directly dependent on the root node (top of the tree) of the configuration and others are subordinate to one of those nodes; each level is termed a tier. In particular, for this Lakota

discussion, the root node has dependent on it the features [NEG], [LARYNGEAL] and [SONORANT VOICE] ([SV], explained below), and on the next lower tier, [SV] optionally dominates either [LATERAL] or [NASAL], while [LARYNGEAL] may dominate [VOICE]. That does not exhaust the inventory of features (there are also manner features like [STOP]), but those are the only ones we need here. See (10) and (11) below for examples of this tier structure and the feature dependencies.

As we said above, any property that can be predicted from its context can be filled in by default or redundancy rules and must be left out of the underlying representation. So, for instance, in a language that has both /t/ and /d/ in contrast, [d] will have the overt feature [VOICE] in its underlying representation, while [t] has its [LARYNGEAL] node left unmarked. Obviously, this kind of feature notation dispenses with the idea of plus and minus on features, since only plus features are allowed.

Phonological rules for assimilation are modeled as feature sharing between adjacent segments. Features are said to spread from one segment to another, but they may only spread to a segment that has a place for them on the next higher tier. Hence [NASAL] can only spread to an adjacent [SV] node, so only a segment with [SV] can be nasalized. Since [SV] is directly dependent on the [root] tier, however, it can spread to any adjacent segment. In general, when a feature spreads, it brings with it whatever is dependent on it. So we would expect the spreading of [SV] from a nasal or [l] to /p/ and /k/ to bring the features [NASAL] or [LATERAL] along and create total assimilation, which does not happen in Lakota. Instead of spreading, then, we must rely on another mechanism, termed "copying", argued for in Rice & Avery (1991) (as cited in Rice 1993: 316). The grammar can copy from an intermediate tier without including its dependents, as in example (10) below.

Within this model (and others), it is claimed that sonorants are always voiced by default, and therefore [VOICE] is not a feature in their underlying representation — it is added by a redundancy rule (see Botma 2011: 175, 177 or Szigetvári 2008: 105). (What are often termed "voiceless sonorants" are marked with the feature [SPREAD GLOTTIS], which also marks aspiration in stops.) This creates a dilemma for situations like the Lakota [bl] and [gl], since the voicing in the stops is clearly assimilation to a following sonorant, but the sonorant has no [VOICE] feature to spread. Rice's solution, which I believe is appropriate here, is that languages manifest two kinds of voicing, represented by different features. The feature with which a sonorant voices the Lakota stops is [SONORANT VOICE] ([SV]), and the resulting stops are "sonorant obstruents". [SV] is present by definition in the underlying configuration of sonorants, including vowels. Evidence for such a claim is presented in Rice (1993; 2013). In contrast, the feature [VOICE]

dependent on [LARYNGEAL] is used by the fricatives and also by the phonemic [b]s.

To explain the behavior of [b] and [g], then, we must have a properly configured tree with appropriate features specified, and rules to rearrange the underlying configurations to produce the surface forms, as detailed in §4.

3 Analysis

3.1 The clusters

Earlier studies of Lakota phonology took little notice of the cluster phenomenon, explaining it as simple voice assimilation of the stop to the following sonorant (Carter 1974: 37; Shaw 1980: 7; Patterson 1990: 22 fn 4). That works just as well, except for the fact that in the currently used model there is no [VOICE] feature to spread.

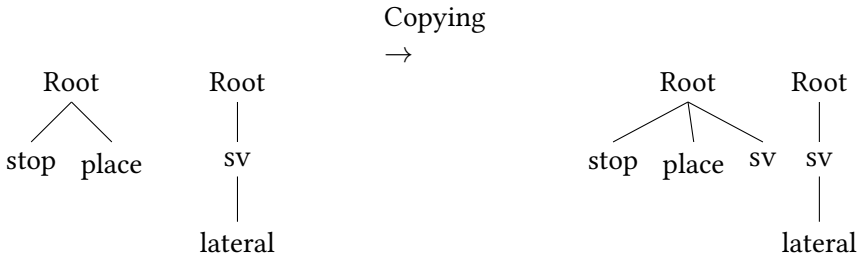
The argument for [sv] in these stops, then, is based first of all on the fact that there is no other choice within the theory. But there is a phonological argument, too. Botma (2011: 180) says, "One type of evidence that is often adduced for the sonorant status of voiced stops is the presence of oral-nasal alternations […]. Implicit in this approach is the claim that only sonorants can be nasalized." As mentioned above, when the [bl] cluster occurs before a nasal vowel, the cluster is nasalized to [mn]. Similarly, but not identically, when a [gl] cluster precedes a nasal vowel, the [l] becomes [n]; however, as noted above, the [g] does not change. Convincing examples of the alternation are provided by some of the motion verbs, as in (9). The verb 'go', *yÁ*, has the first person form *blA* 'I went'. When the final vowel of the verb is ablauted to *iŋ*, however, the form becomes *mníŋ*, e.g. *mníŋ kte* 'I will go'. Similarly, the verb *glÁ* 'be going home' changes to *gníŋ kte* 'he will be going home'.

(9) Example of nasalization of [bl] and [gl]
   a. *yA* 'go' + first person agent > *blA*
   b. *blA* + *iŋ*-ablaut > **bliŋ*[3] > *mniŋ* (example: *mníŋ kte* 'I will go')
   c. *glA* 'go home' + *iŋ*-ablaut > **gliŋ* > *gníŋ* (example: *gníŋ kte* 'he will go home')

[3] Actually, there is one extended family on the Pine Ridge reservation that does say *bliŋ kte*, to the amusement of many others.

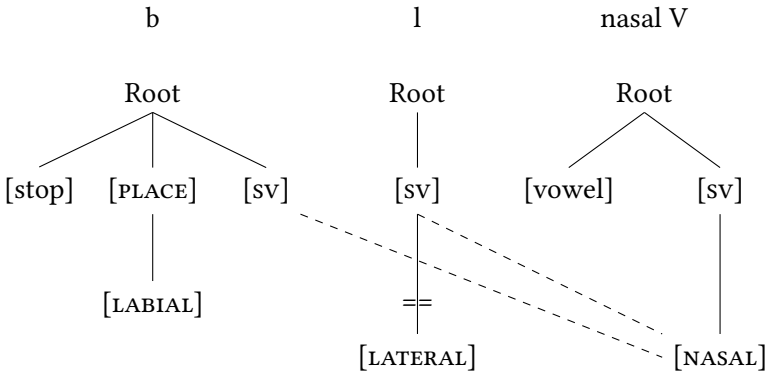
The /stop/-clusters are accounted for by rule (10) (where the mechanism of "copying" mentioned above allows one node to take on a feature from an adjacent node without also taking on the dependents of the source node) and the nasalization (different for /bl/ and /gl/) by rule (11):

(10) Rule for sonorantizing stops preceding /l/ in clusters:

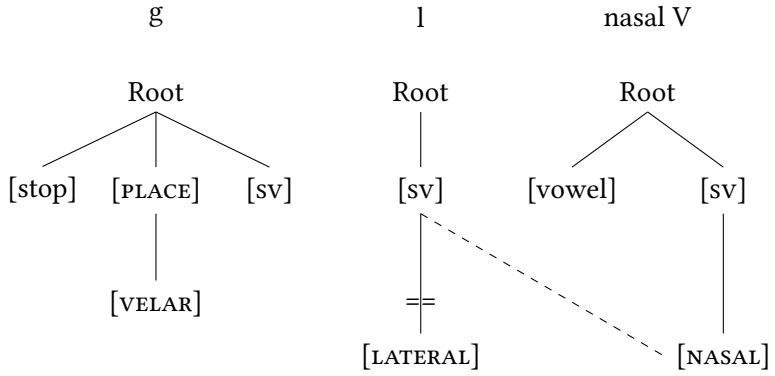


(11) Rules for nasalizing the sonorant obstruents

a.



b.



3.2 Morpheme-final or coda position

If that were the whole story, of course, it would not be very exciting, since under either the old theory or the new one, we have a simple assimilation to describe. However, the other context for [b] and [g] brings up other issues.

We have just established that Lakota [b] and [g] in clusters are sonorant allophones of /p/ or /ph/ and /k/, so we can now use that discovery to analyze those sounds in morpheme final position. Here the phonological phenomenon is not assimilation, however — there is no spreading or copying. Rather, we have a change conditioned by the position of the segment in the word. Let us look a little more closely at exactly what is going on here.

When the underlying final vowel of a morpheme is deleted, the consonant preceding it comes to stand in word or morpheme final position, and at the same time, the syllable structure of the morpheme changes. In the simplest case, what used to be the onset of the final shifts to the coda of the former preceding syllable, as in these examples (<.> is a syllable boundary):

(12) Examples of syllable structure changes when a vowel is lost
   a. *khú.ža* 'be sick' reduplicates *khuš.khú.ža*
   b. *šá.pA* 'be dirty' + *yA* 'cause' > *šab.yÁ*
   c. *šúŋ.ka* 'dog' + *má.ni.tu* 'wilderness' > *šuŋ[ŋ](g).má.ni.tu* 'coyote'

In more complicated cases, e.g. when the element following the deleted vowel itself begins with a vowel, the newly final consonant can remain in onset position, and we usually do not see the consonant modification. However, vowel initial words often begin with a phonetic glottal stop ([ʔ]), and sometimes that glottal

stop remains part of the structure of the word in non-initial position. (When or whether the glottal stop in this language is a phoneme is an unexplored question, but it often seems to be unpredictable.) Examples of this are found in (13):

(13) Varying resyllabification before vowel-initial second elements
  a. *Lakȟóta* 'Lakota' + *iyápi* 'language'
    i. *La.kȟól.ʔi.ya.pi*
    ii. *La.kȟó.ti.ya.pi*
  b. i. *čhaŋ.té* 'heart' + *aglé* 'set against' > *čhaŋ.tágle* 'plot (evil) against'
    ii. *čhaŋ.té* 'heart' + *ʔasní* 'recover' > *čhaŋl.ʔásni* 'calm down from an angry spell'

The conflicting glottal-stop phenomena sometimes result in doublets such as the combination of *Lakȟóta* 'Lakota' plus *iyápi* 'language' (see (13a)). In this example, if the glottal stop remains in the compound, we have the form *La.kȟól.ʔi.ya.pi*, with /t/ > [l] in coda position. However, if the glottal stop is omitted in the compound, we get the form *La.kȟó.ti.ya.pi*, in which the /t/ does not change because it is still in syllable onset position. Both of these constructions are grammatical. A third possibility, illustrated by example (8c) above, *šuŋ.k'á.kaŋ.yaŋ.ke* 'ride horseback', manifests a rule that a voiceless stop followed by a glottal stop may merge the two into an ejective, which then prevents the change from obstruent to sonorant. Which of these three realizations (sonorant-syllable boundary, syllable boundary-voiceless stop, or syllable boundary-ejective) is actually used is lexically and perhaps idiolectically conditioned.

Instead of doublets, we sometimes see lexically conditioned constructions with one or the other syllable structure, e.g. (see 13b) *čhaŋ.té* 'heart' + *a.glé* 'set against' > *čhaŋ.tá.gle* 'plot (evil) against' (no glottal stop beginning the second element of the compound) in contrast with *čhaŋ.té* 'heart' + *ʔa.sní* 'recover' > *čhaŋl.ʔá.sni* 'calm down from an angry spell', in which the /t/ changes its role in the syllable and consequently changes to a sonorant.

Phonologists often discuss changes of this sort using the concepts of segment weakening or strengthening, also called lenition or fortition (but note Honeybone 2008: 10): "Cases of real fortition are vanishingly rare."). These concepts have been around since the mid-19th century at least (Honeybone 2008), but there is often disagreement among scholars as to which contexts are strengthening and which are weakening, or which phonetic phenomena constitute one or the other of those processes. Word final devoicing in German, for example, is sometimes argued to be fortition to mark a word boundary (Iverson & Salmons 2007 as cited

in Harris 2009), but much more often described as lenition (Harris 2009), perhaps as a consequence of vocal tract relaxation at the end of a unit. In support of the latter position in general, not just for German, Szigetvári (2008: 112) asserts "It is a phonological commonplace that consonant lenition is typical in Coda position." Since our target sounds are in coda position, we should therefore look for a way to declare their voicing to be lenition.

There are numerous definitions of lenition in the literature (see Honeybone 2008, Harris 2009, and Szigetvári 2008 for discussion). Sometimes it can be described as moving toward a less marked version of the sound, sometimes as the loss of a distinctive feature (which may amount to the same thing), and sometimes it is defined, utilizing a particular sonority hierarchy, as an increase or decrease in sonority, though there is little consensus about which sounds are more sonorant than others, especially among the obstruents. Another hypothesis, specifically to explain final devoicing, is that the word boundary has the properties of a voiceless obstruent, so final devoicing is assimilation to the voicelessness of the word boundary.

Obviously, assimilation to the word or morpheme boundary is not available to describe adding [VOICE] to the Lakota stops, and since the change in question adds rather than subtracting a feature, the "remove a feature" or "move toward unmarked" categorization is also wrong for the stops here, though it does describe the behavior of the fricatives accurately. That leaves us with the idea that the principle in Lakota for coda consonants is "make them sonorant". Rankin (2001: 5) formulated this as a sound law: syllable final stops become sonorants. He posits that the first version of this sound change was to produce nasals in syllable final position, since nasals are the least marked sonorants (cf. Rice 1993). Then [b] and [g] result from de-nasalization in oral contexts. I do not think we need to posit such a sequence, since we have a theoretical device for directly turning stops into sonorants: add [SV] to the stop. The labial and velar then become sonorant obstruents, and the coronals become [l]. The change from [č] to /g/ requires a change in place as well, but the voicing is now accounted for.

Before we accept this analysis, however, we should take note of another situation where voiced obstruents frequently replace voiceless ones, namely, in intervocalic position. There are many examples of this in the history of the Romance languages. Here the process seems to be voicing assimilation, and it is usually called lenition. If we carry this description of voiced stops as the result of lenition over to word or syllable final position, however, we are faced with a major contradiction: both devoicing, as in German, and voicing, as in Romance, are called lenition. I think we cannot have it both ways, and therefore invoke, again, the

concept of [SV] as different from [LARYNGEAL VOICE]. Loss of [LARYNGEAL VOICE] (fricatives) and addition of [SV] (producing sonorants from stops) can then both be lenition.

Consequently, it seems clear that the combination of analyzing [b] and [g] as sonorants and observing that they are allophones of voiceless and aspirated stops in codas, an environment that characteristically supports weakening, is probably the correct description of just what these sounds are doing in this language. The analysis is strengthened by the addition of [l] to the class that includes [b] and [g], since [l] is an underlying sonorant to begin with. Using the concepts developed by Mielke (2008), we can thus describe an "emergent" phonological class for this language consisting of [b,g,l]. Moreover, the fact that all three of these sounds can undergo the further addition of [NASAL] means that they must have an [SV] or sonorant node to which the feature can spread from the vowel.

As alluded to above, Rankin (2001: 5) came to a similar conclusion using historical reconstruction principles instead of phonological theory. He isolated the voiced variants of the coda stops and speculated that there was a sound change which turned stops into sonorants, perhaps starting with all of them becoming nasals (the least marked sonorants, according to Rice's (1993) analysis), after which [b] and [g] denasalized in non-nasal contexts. He says:

> It seems clear that […] [the] stops systematically become/became the corresponding sonorants. This is relatively clear from the inventory of reflexes […] alone: of these, m, n, ŋ and l are sonorants. Only b and g are stops and phonetic obstruents. But although they are phonetic stops, they have residual voicing left over from their sonorant phonological status. I would claim that, even synchronically, the mysterious voiced stop codas are actually variants not of the underlying, voiceless stops but of voiced sonorants.

4 Issues in need of further investigation

There remain, however, two rather embarrassing complications to this determination, namely the variability in the degree of voicing in the stops in the coda, and the variability in the nasalization in the same context. Furthermore, we have not said anything about the process whereby /y/ becomes [l] in the sonorant clusters. All three of these deserve further study, but some hints about how and where to look are available from both theoretical phonology and Siouan historical studies.

4.1 Variable voicing

The definitions and descriptions of sonorants, including sonorant obstruents, in the literature frequently make the point that these sounds are by definition spontaneously voiced, and fully voiced throughout the duration of their articulation.[4] Note, for example, that Botma (2011: 188) says of several African languages "that sonorant stops are consistently realized with *active* voice facilitation, and that the degree of voice facilitation is such that these sounds are voiced throughout their duration" [emphasis in original]. We have seen, however, that this is not the case for Lakota [b] and [g]. We have been relying on Rice's determination that the kind of "voicing" found in sonorants is phonologically different from the laryngeal voice found in obstruents, but no one has discovered a phonetic difference in the way voice is physically (phonetically) manifested in these two categories. Yet the case for phonological [sv] is compelling enough that we should not abandon the analysis on the basis of the manifestation of [voice], but rather think about how we might accommodate variable voicing in the analysis.

I propose that the explanation can be derived from the observation that languages can use the voiced/voiceless contrast in two different ways. One recent summary of this proposal is found in Harris (2009: 15–19), another in Ringen, Beckman & Jessen (2013). They divide languages into "voicing" and "aspiration" types. In voicing languages, like Dutch and French, there is no Voice Onset Time (VOT) delay after either voiced or voiceless stops in initial position, and the vocal fold vibration accompanying voiced stops is present throughout the duration of the stop closure in both initial and final position unless the language also has a final devoicing rule (Dutch does, French does not). Final voiced stops may even manifest audible voice after the release of the stop; some refer to that as "post-voicing".

In aspiration languages, like German and English, on the other hand, there is a contrast in word initial position between significant VOT delay and its absence. The delay in voice onset is heard as aspiration on the stop. When the stop is not aspirated, phonetic voicing can "leak" into it from the environment. It is a commonplace in beginning phonetics classes, for example, to have students observe that a word initial [b] in English does not start its voicing immediately at the beginning of the stop closure, and a word-final [b] loses its voicing part way through the closure. Although most of us have been taught that this is

[4] Let me remind the reader that what we have for decades been calling voiceless sonorants (e.g. the initial *wh-* in English *which* in some dialects, or the sound spelled <ll> in Welsh) are now considered better described as sonorants with the feature [spread glottis], the same feature that produces aspiration in stops.

assimilative or "leaking" voicelessness into an underlyingly voiced consonant, the new analysis is instead that the stops (spelled <b>, <d>, <g>) are not voiced at all in their underlying forms – this kind of voice is not phonologically significant. What counts for phonology is the presence vs. absence of aspiration. Anecdotal evidence for the importance of aspiration over voice in English can come from the observation that beginning English-speaking phonetics students have to work hard to hear so-called voiceless unaspirated stops, transcribing them as voiced instead. It is clear that Lakota is an aspiration language in this typology, albeit largely without much "leaking" of voice into the plain stops.

I do not know whether the languages that are cited in support of the [sv] hypothesis are voicing or aspiration, but it seems possible that the determination that sonorant obstruents are voiced throughout their duration could correlate with the "voicing" typology. However, Lakota appears to be a language that both has sonorant obstruents and is an aspiration language. Since the phonetic voice which manifests phonological [sv] can only come from the environment, the observed variability in the manifestation of that voice makes sense. Evidence for this comes from the informal observation that [b] and [g] in coda position often seem to match the voice phonetics of the following consonant. For example, *tȟáŋka* 'large' is recorded in the dictionary as reduplicating *tȟaŋktȟáŋka*, but forming the adverb *tȟaŋgyá*. It would seem worthwhile to examine how sonorant obstruents are realized in other aspiration languages.

4.2 Variable nasalization

I have not discovered an equivalently neat solution to the nasal variation, but a recent phonetic study of Lakota nasalization (Scarborough et al. in press) may provide a clue. The question investigated was how nasalization is manifested in Lakota vowels under three circumstances: when the surrounding consonants are oral, when the vowel is oral but a neighboring consonant is nasal, and when the vowel is nasal with neighboring nasal consonants.

> The study examines the timing and degree patterns of acoustic vowel nasality across contrastive and coarticulatory contexts [...] With respect to patterns within the vowel, in the absence of nasal consonant context, contrastive vowel nasalization is generally greatest in degree late in the vowel. Low nasal vowels in carryover contexts parallel this pattern (despite the location of the nasal consonant before the vowel), and low nasal vowels in anticipatory contexts are most nasal at the start of the vowel.

(Scarborough et al. in press).

Siouanists are convinced that historically, vowel nasalization is underlying and nasal consonants are derived when nasalization is spread from the vowel (Rankin, Carter & Jones 1998). This, of course, is also the pattern we have described above for [b] and [g]. In the sonorant clusters, the following vowel spreads nasalization leftward, including /b/ in its target but excluding /g/; in coda position, a preceding nasalized vowel may spread its nasalization rightward to varying degrees. In both cases, the phonetic nasalization is strongest at the edge of the vowel which is further from the nasal consonant. The historical development is always a leftward spread. Schematically, assuming that the phonetic pattern discovered by Scarborough et al. (in press) is also the historical one:

(14) Phonetically gradient nasality in adjacent consonants and vowels

a. In the clusters and historically for all sonorants:

Sonorant(s)	first part of vowel	last part of vowel
fully nasal	less nasal	more nasal
C	Vŋ	

b. In the codas:

Vowel start	vowel end	C start	C end
most nasal	less nasal	most nasal	less nasal
Vŋ		C	

So the cluster pattern matches the historical pattern (and there is no "partially nasal" in the consonants of the clusters), while in the coda, the consonant replicates the vowel pattern of decreasing nasality during the articulation of the segment. The question to be investigated, then, is whether there is a good explanation for the within-vowel or within-sonorant nasalization gradient. I leave the answer to future research.

4.3 The relation between [y] and [l]

The final topic for this discussion is a mystery that may never be solved, but might inspire someone to try additional research. Throughout the preceding, we have assumed the presence of underlying /l/ in the clusters [bl, gl, gn] and [mn]. However, as was noted briefly above, a very frequent occurrence of these clusters is in conjugated verbs, where [bl] signals first person agent of a verb whose stem begins with /y/, and [gl] represents /k/ in several morphemes, also before a verb stem that begins with /y/.

Koontz (1991) treats the y-initial verbs as one subclass of "syncopating stems", reconstructing the [b] as *p and the [g] as *k. We have accounted above for the

voicing in the stops, but nothing has been said about a process that could change /y/ to [l] to provide the conditioning environment for the /p/ > [b] and /k/ > [g] phenomena.

The *Comparative Siouan Dictionary* (Rankin, Carter & Jones 1998; Rankin et al. 2015) point out that proto-Siouan seems to have had five sonorants, which they reconstruct as *r, *y, *w, *R and *W. Informally among themselves, the dictionary editors refer to *R and *W as "funny r" and "funny w". The difference between *r and *R, or between *w and *W, is murky and confusing, since their reflexes are pretty much the same segments (e.g. [l, r, n, d, y, ð], maintaining sonority and coronality for both *r and *R), but distributed differently among the daughter languages. In particular, in Lakota, *r descends as /y/, and *R descends as /l/. (For the curious, but irrelevant here, *y descends to Lakota as /čh/). Our [bl] and [gl] clusters, then, are historically *pr and *kr, and somewhere in the more recent history of the language that *r changed to *R in the conjugated forms but remained *r in the stem. No one has made any suggestions about when or how this could have happened, nor exactly what the development might have been. Of little or no additional help, but also in need of an account, is the second person form of these verbs in which the stem-initial /y/ changes to [l] but there is no other prefix.

5 Conclusion

We started out to learn about the phonological status of the phonemically marginal [b] and completely allophonic [g] in Lakota. We concluded that they are patterning as sonorants for two reasons: first, because they are assimilating to a following sonorant, and only sonorants can acquire an autosegmental feature that would account for them being voiced, and second, because they are sometimes nasal, and only sonorants can be nasals. However, this manifestation of sonority in the obstruents comes from two different directions. On the one hand, both [b] and [g] assimilate (copy) their sonorant feature from a following consonant to create clusters, and on the other, [b] and [g], together with [l], constitute a phonological class that realizes weakening (lenition) in coda position in derived words. In both patterns, an adjacent nasalized vowel can contribute its nasality to the sonorants, but in different ways for each pattern. In the clusters, the conditioning environment follows the sonorant, whereas in the codas, it is a preceding nasal that causes the shift. In the codas, the two stops (but not [l]) show considerable variation in the degree of voicing that they manifest, which may be attributable to the fact that Lakota is an aspiration language in which

obstruent voicing is phonologically absent and phonetically acquired from the environment. Also in the codas, the extent to which the nasalization feature spreads into the consonants is variable, perhaps repeating the pattern of gradient nasality discovered to be phonetically present in nasalized vowels. Finally, we revealed an assimilation problem that takes us into the history of the language, namely the fact that the [l] that we have been studying replaces a /y/ in the verb stem when the verb takes on an agent prefix. That replacement is apparently old, but unexplained. We have thus combined the phonological theories of autosegments and feature geometry, voicing typology, and the history of Siouan to shed light on the question of what the phonological status of the voiced stops is in this language.

Acknowledgments

I would like to thank Keren Rice profusely for considerable help with all aspects of this paper, from the inspiration for the analysis to the formatting of rules and examples and suggestions for theoretical bibliography. I have also profited from the comments of an anonymous reviewer, as acknowledged in the appropriate places above, and in addition I would like to thank editor Catherine Rudin for very useful feedback.

Abbreviations

NEG	negation
SV	Sonorant Voice
VOT	Voice Onset Time

References

Botma, Bert. 2011. Sonorants. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 171–193. Hoboken, NJ: Wiley-Blackwell.

Buechel, Eugene, S. J. 1970. *A dictionary of the Teton Lakota Sioux language*. Pine Ridge, SD: Holy Rosary Mission.

Carter, Richard T., Jr. 1974. *Teton Dakota phonology* (University of Manitoba Anthropology Papers 10). Winnipeg: University of Manitoba.

Clements, George N. 1985. The geometry of phonological features. *Phonology Yearbook* 2. 225–252.

Harris, John. 2009. Why final-obstruent devoicing is weakening. In Kuniya Nasukawa & Phillip Backley (eds.), *Strength relations in phonology*, 10–45. Berlin: de Gruyter Mouton.

Honeybone, Patrick. 2008. Lenition, weakening and consonantal strength: Tracing concepts through the history of phonology. In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.), *Lenition and fortition*, 9–92. Berlin: de Gruyter Mouton.

Iverson, Gregory K. & Joseph C. Salmons. 2007. Domains and directionality in the evolution of German final fortition. *Phonology* 24. 121–145.

Koontz, John E. 1991. *Central Siouan syncopating stems*. Manuscript revision.

Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.

Miner, Kenneth L. 1979. An unusual natural class in Mississippi Valley Siouan. *Siouan and Caddoan Linguistics* 2. 8–9.

Patterson, Trudi. 1990. *Theoretical aspects of Dakota morphology and phonology*. Champaign: University of Illinois – Urbana-Champaign (Doctoral dissertation).

Rankin, Robert L. 2001. *On Dakotan pre-boundary and cluster phonology*. Paper presented at the Siouan and Caddoan Languages Conference, University of Chicago, June 2001.

Rankin, Robert L., Richard T. Carter & A. Wesley Jones. 1998. Proto-Siouan phonology and grammar. In Xingzhong Li, Luis López & Tom Stroik (eds.), *Papers from the 1997 Mid-America Linguistics Conference*, 366–375. Columbia: University of Missouri Press.

Rankin, Robert L., Richard T. Carter, A. Wesley Jones, John E. Koontz, David S. Rood & Iren Hartmann (eds.). 2015. *Comparative Siouan dictionary*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://csd.clld.org (25 September, 2015).

Rice, Keren. 1993. A reexamination of the feature [sonorant]: The status of "sonorant obstruents". *Language* 69. 308–344.

Rice, Keren. 2013. *Sonorant obstruents revisited*. Handout forthe International Conference on Phonetics and Phonology 2013, National Institute for Japanese Language and Linguistics, 25 January.

Rice, Keren & Peter Avery. 1991. *A learning path for phonology: A theory of segmental elaboration*. Paper presented at the Boston University Conference on Child Language Acquisition, October 1991.

Ringen, Catherine, Jill Beckman & Michael Jessen. 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49. 259–284.

Scarborough, Rebecca, Georgia Zellou, Armik Mirzayan & David S. Rood. in press. Phonetic and phonological patterns of nasality in Lakota vowels. *Journal of the International Phonetic Association.*

Shaw, Patricia A. 1980. *Theoretical issues in Dakota phonology and morphology*. New York: Garland Publishing.

Szigetvári, Péter. 2008. What and where? In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.), *Lenition and fortition*, 93–129. Berlin: de Gruyter Mouton.

Ullrich, Jan F. (ed.). 2012. *New Lakota dictionary*. 2nd edn. Bloomington, IN: Lakota Language Consortium. http://www.lakotadictionary.org/nldo.php.

Chapter 11

The syntax and semantics of internally headed relative clauses in Hidatsa

John P. Boyle

Hidatsa is a highly endangered member of the Missouri Valley branch of the Siouan language family. Like all Siouan languages, Hidatsa is a left-branching SOV language with many polysynthetic characteristics. Hidatsa specifically, and Siouan in general, is typologically unusual in that it only allows relative clauses to be internally headed. This paper examines the structure of internally headed relative clauses (IHRCs) in Hidatsa. I vide a syntactic analysis within the Minimalist Program as proposed by Chomsky (1995; 1998; 2001) and show that IHRCs are nominalized sentences that serve as DPs in larger superordinate clauses. This extends previous analyzes of IHRC for Native American languages, most notably Williamson (1987), Cole (1987), Culy (1990), and Basilico (1996) among others. I then show that Hidatsa, like other languages with IHRCs, obeys the indefiniteness restriction first proposed by Williamson (1987) for Lakota. Employing the general framework first proposed by Heim (1982) and developed by Williamson (1987) and Basilico (1996), I provide a semantic explanation for the fact that the head noun of an IHRC cannot be marked as definite. Following Williamson (1987) and Culy (1990), I treat the indefinite determiner that follows the head of the IHRC as a variable and not as a quantifier. I argue that the head noun must move covertly to a [SPEC, CP] position to escape VP-level existential closure.

1 Introduction

Hidatsa is a highly endangered member of the Missouri Valley branch of the Siouan language family and is spoken by approximately 175 speakers, primarily on the Ft. Berthold Indian Reservation in North Dakota. Typologically, Hidatsa is left branching with a complement-head SOV word order and, like most other Siouan languages, it has an active-stative pronominal system. Along with Crow, the other member of the Missouri Valley branch of Siouan, it has many polysynthetic characteristics, most notably productive noun incorporation (Rankin et al.

John P. Boyle. 2016. The syntax and semantics of internally headed relative clauses in Hidatsa. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 255–287. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.174
DOI:10.17169/langsci.b94.174

2003; Boyle 2007). One aspect of Hidatsa that is of particular interest is the structure of its relative clauses.

Like many other Siouan languages (see Drummond 1976 and Cumberland 2005 (Assiniboine); Williamson 1987 and Rood & Taylor 1996 (Lakota); Quintero 2004 (Osage), and Graczyk 1991, 2007 (Crow), among others), relative clauses in Hidatsa are internally headed. This means that the noun that is modified by the relative clause is positioned within the relative clause and is not external to it, as happens in languages like English.

In §2, I will provide data that shows that Hidatsa RCs are nominalized clauses that act as DPs for superordinate predicates. I will also detail how IHRCs in Hidatsa differ with regard to specificity. In §3, I will discuss the notion of HEAD and show that the head of the IHRC in Hidatsa does indeed stay low in the subordinate construction (counter to claims made by Kayne 1994, Bianchi 1999 and Di Sciullo 2005). In §4, I will examine several previous works on IHRC, most notably Williamson (1987) and Culy (1990). In §5, I will provide a syntactic analysis within a minimalist framework showing how the derivation of the IHRC is built in Hidatsa. In §6, I will argue that the head of the clause must move covertly in order to escape the existential closure of the VP. This is motivated by a semantic necessity for full interpretation. This analysis explains Williamson's indefiniteness restriction for IHRCs. §7 will provide a brief conclusion as well as avenues for further research.

2 The status of Hidatsa RCs

2.1 Hidatsa RCs as DPs

In Hidatsa, RCs are nominalized constructions that act as DPs for superordinate clauses. As stated above, Hidatsa has an SOV word order as shown below in (1).

(1) **Common S-O-V sentence with simple DP arguments**

Wíaš	*wacéeš*	*íkaac.*
wía-š	wacée-š	íkaa-c
woman-DET.D	man-DET.D	see-DECL

'The woman saw the man.'

In Hidatsa, RCs can serve in any syntactic role that can be filled by an ordinary DP. In example (2), the RC (in brackets) is subject of the main clause.

(2) **RC serving as subject**
[*Wacée akuwaapáahiš*] *wía* *íkaac.*
[wacée aku-waapáahi-š] wía íkaa-c
[man REL.S-sing-DET.D] woman see-DECL
'The man that sang saw the woman.'

In (3), the RC is the object of the main clause.

(3) **RC serving as object**
Wíašéʔeri [*wacée akuwaapáahiš*] *íkaac.*
wía-šéʔe-ri [wacée aku-waapáahi-š] íkaa-c
woman-DEM-TOP [man REL.S-sing-DET.D] see-DECL
'The woman saw the man that sang.'

In both examples (2) and (3) the word order of the main clause is SOV. RCs can also be possessed as shown in (4). In these types of constructions, the possessed RC takes a possessive prefix, in this case the 3rd person possessive prefix *ita-*.

(4) **Possessed RC**
Ráaruwa [*itaʔaruʔapáxihe*] *híiwareec.*
rée-a-ruwi-a [ita-aru-apáxi-hee] híi-wareec
go-CONT-along-CONT [3.POSS.A-REL.N-stop.rest-3.CAUS.D.SG] arrive-NE
'Going along, he (First Worker) arrived at the place where Sun stops to rest.' (Lowie 1939: I: 21)

RCs in Hidatsa can also function as objects of postpositions, shown in example (5). In these types of constructions, the head of the RC takes a postpositional suffix. In this example, the postpositional suffix is *-wahu* 'inside' (shown in **bold**).

(5) **Object of a postpositional phrase**
Ráaruwaa [*wirawáhu* *arušiipikáatikua*] *hírawa*
ráa-ruwa-a [wira-**wáhu** aru-šiipi-káati-kua] hírawi-a
go-continue-CONT [woods-**inside** REL.N-thick-EMPH-LOC] sleep-CONT
wáakiruk *uʔúšiawareec.*
wáaki-ruk uʔúšia-wareec
be.there-TEMP arrive-NE
'Going along in the woods where it is very thick, he (First Worker) arrived while (Spotted Tail) was still sleeping.' (Lowie 1939, II:54)

In examples (2)–(5), each of the RCs act as a DP for the main clause, serving as an argument in the same manner as a common DP.

2.2 Specificity in Hidatsa RCs

Hidatsa RCs are all formed in a similar manner. They are nominalized clauses that can take zero, one or two overt DP arguments depending on the transitivity of the nominalized verb. This clause then acts as a DP for a superordinate clause. It is the entire RC that takes a final determiner, thus making it available as an argument for another predicate. The overt syntactic structure is shown in (6).

(6) **The Hidatsa RC structure**
[(DP) (DP) REL-verb]-DET

As shown in (6), the verb (or predicate) is relativized by a relative marker. Hidatsa has two relative markers, *aku-* (shown in (2) and (3)) and *aru-* (shown in (4) and (5)). The first marker, *aku-*, marks for specificity and speakers prefer to use it for animate entities. An example of this is shown in (7).

(7) **The Hidatsa specific relative marker**
Wašúka aku?áwakaš *wakítʰaac.*
wašúka **aku**-ʔ-áwaka-š wa-kía-tʰaa-c
dog REL.S-EPE-1A.see-DET.D 1A-fear-NEG-DECL
'I am not afraid of that dog that I see.'

The second marker, *aru-*, marks nonspecific arguments and is often used for inanimate objects or entities. An example of this is shown in (8).

(8) **The Hidatsa non-specific relative marker**
Wašúka aru?awákaš *wakítʰaac.*
wašúka **aru**-ʔ-awáka-š wa-kía-tʰaa-c
dog REL.N-EPE-1A.see-DET.D 1A-fear-NEG-DECL
'I am not afraid of a dog that I see.'

The overriding attribute of these markers is one of specificity and not animacy. It should be noted that neither of these markers is a relative pronoun. There is no gap in the relative clause, as in English, nor do they act as an argument projected by the verb. This is to say, they do not satisfy the subcategorization frame of the verb. The subcategorization frame of the verb can only be satisfied by an overt

DP argument or signaled by a pronominal prefix, which agrees with a *pro*.[1] The difference in specificity does nothing to alter any aspect of the syntax and as such, they can be treated identically in the syntactic analysis presented below.

3 The notion of head and internal heads

The term *internally headed relative clause* was first coined by Gorbet (1976) in his description of Diegueño nominals. In some of the subsequent literature, there has been confusion about the notion of *head*. Cole, Herbert & Hermon (1982) and Weber (1983) (among others) have referred to IHRCs as *headless relative clauses'*. This was based on the notion that the head noun that was being modified by the relative clause was not external to the RC as in languages like English. In this paper, the notion of *head* is a semantic one. It refers to the noun being modified by the relative clause. It is not to be confused with the *head* in an X-bar construction (i.e. a X^0 category), which is a syntactic notion. As IHRCs are sentences, and sentences never have DPs as their head, the head of the IHRC cannot be the syntactic head of the sentence. The term *headless relative clause*, as used by the authors mentioned above, is misleading for another reason as well. This reason is that the head noun modified by the clause may be null. An example of a RC in Hidatsa with an overt head is shown in (9).

(9) **Hidatsa relative clause with an overt head**
*Mary **uuwáki** akuhíriš warúcic.*
Mary **uuwáki** aku-híri-š wa-rúci-c
Mary **quilt** REL.S-make-DET.D 1A-buy-DECL
'I bought the quilt that Mary made.'

This construction can be juxtaposed with the example in (10), where the head that is modified by the RC is null.

(10) **Hidatsa relative clause with a null head**
*Mary **e** akuhíriš warúcic.*
Mary **e** aku-híri-š wa-rúci-c
Mary **e** REL.S-make-DET.D 1A
'I bought (what/something/it) Mary made.'

[1] Hidatsa is an active-stative language with regards to its pronominal system and only the first and second person pronominal prefixes are overt. The third person marker is null. In this paper, I treat the pronominal prefixes as agreement markers on the verb that agree with a null *pro* that occupies an argument position.

As stated above, the subcategorization frame of the verb is satisfied by the projection of a *pro*. There is also no overt agreement marker on the verb as the third person agreement prefix is a null affix. This is not possible in English as is clearly reflected in the gloss with the insertion of the English third person indefinite (*what/something/it*). For the purposes of this paper, the example shown in (10) is an example of a *headless* relative clause (i.e. the head that is modified by the RC is null or not phonologically overt). Given this evidence, we can now define an IHRC (following Culy 1990: 27) as: "A (restrictive) internally headed relative clause is a nominalized sentence, which modifies a nominal, overt or not, internal to the sentence."

3.1 Internal heads

In languages with internally headed relative clauses, the head of the RC is always internal to the RC and not part of the superordinate clause. This is counter to the claims of Kayne (1994), which were elaborated on by Bianchi (1999) and Di Sciullo (2005). Kayne claims that the heads of these clauses are always the left-most element, and thus are really outside of the relative construction itself. While this may appear to be correct for IHRCs with intransitive verbs and transitive verbs where the head noun corresponds to the subject of the RC, it cannot account for IHRCs where the head noun corresponds to the object of the RC. Following the copy theory of movement, it could be claimed that the copy at the relative head position is deleted in IHRCs, in contrast to EHRCs, in which the copy in the tail of the movement chain is deleted. This would suggest that there is a choice in pronunciation between the head and the tail of the chain in the formation of relative clauses. However, there is no evidence or motivation for the head of the RC to move out of the clause to the left periphery and then not be pronounced in that external position, contravening the typical pattern for copy deletion. This is particularly problematic when the head noun is an internal argument in the clause.[2] A clear example of this is shown in (11).

(11) **Hidatsa internally headed relative clause**

[*Wacéeš*	***wašúkawa*** *akutíheeš*]	*šipíšac.*
[wacée-š	**wašúka-wa** aku-tí-hee-š]	šipíša-c
[man-DET.D	**dog-DET.I** REL.S-die-3.CAUS.D.SG-DET.D]	black-DECL

'The dog [that the man killed] is black.'

[2] A more detailed critique of Kayne's analysis of IHRCs can be found in Boyle (2007).

In this sentence, the relative clause *wacéeš wašúkawa akutíheeš* contains the noun that is modified, namely *wašúkawa* 'a dog'. In the English translation, the relative clause *that the man killed* does not contain the noun that is modified by the relative clause. In English, the head is outside of the clause. This is not the case in Hidatsa and languages like it. In these languages, the head stays low and is internal to the subordinate clause, hence the name *internally headed relative clause*.

This type of relative clause is clearly subordinate to the matrix (or superordinate) clause. Number marking on the different verbs in (12) provides evidence of this subordinate relationship. In Hidatsa, verbs agree in number (singular or plural) with their subject.

(12) **Plural agreement in a Hidatsa relative clause**

Alex [*wíaaku?o*	*uuwáki akuhíra?aš*]	*íkaac.*
Alex [wía-aku-?o	uuwáki aku-híri-?a-š]	íkaa-c
Alex [woman-DET.SPEC-PL.I	quilt REL.S-make-PL.D-DET.D]	see-DECL

'Alex saw [those women who made the quilt].'

In this sentence, the matrix verb *íkaa-* 'see' agrees with the subject of the main clause and is marked as singular in number. This is shown with zero marking, and it agrees with the subject *Alex*. The subordinate verb in the relative clause *híri-* 'make' shows plural agreement. This is shown with the definite plural marker *-?a-*. The verb *híri-* agrees with its subject, *wíaaku?o* 'those women'. This argument, *wíaaku?o*, is also the head of the relative clause, which is marked with the indefinite plural marker *-?o-*. In IHRCs, the head occupies an argument position that is determined by its role in the subordinate relative clause, in this case the subject. In this example there is agreement in number between the subjects and the verbs in both clauses. The entire RC, not merely the head noun, then acts as an argument for a superordinate verb. In (12) the RC is the object of the main verb.

4 The syntax of IHRCs

Although a number of people[3] have worked on IHRCs, there are two analyses that I will review and build upon. These are Williamson (1987) who developed

[3] These include Hale & Platero (1974), Gorbet (1976), Fauconnier (1979), Cole (1987), Culy (1990), Kayne (1994) Bianchi (1999), Citko (2001) among others.

the notion that IHRCs have an indefiniteness restriction, and Culy (1990) who argues for a *wh* element that triggers movement of the internal head to [SPEC, CP] at LF.

4.1 The indefiniteness restriction

In her 1987 article, Williamson observed that in Lakota the head in an IHRC cannot be marked as definite.[4] This "indefiniteness restriction" has been generalized to IHRCs in general and to date no evidence has emerged to counter this claim. This indefiniteness restriction holds true for Hidatsa. Examples (13) and (14) are both grammatical in Hidatsa.

(13) *Wacéewa akuʔawákaš maapáahic.*
wacée-wa aku-awáka-š maa-páahi-c
man-DET.I REL.S-1A.see-DET.D INDEF-sing-DECL
'The man that I saw sang (something).'

(14) *Wacée akuʔawákaš maapáahic.*
wacée aku-awáka-š maa-páahi-c
man REL.S-1A.see-DET.D INDEF-sing-DECL
'The man that I saw sang (something).'

It is important to note that the English gloss of both (13) and (14) uses the definite article for the head of the relative clause: *the man*. This is because *the man* gets its definiteness from the determiner that nominalizes the predicate *awáka* 'see', which in both examples is the definite determiner -*š*. While both (13) and (14) are grammatical in Hidatsa, example (15) is not.

(15) *[*]Wacéeš akuʔawákaš waapáahic.*
wacée-š aku-awáka-š waa-páahi-c
man-DET.D REL.S-1A.see-DET.D INDEF-sing-DECL
(Intended: 'The man that I saw sang.')

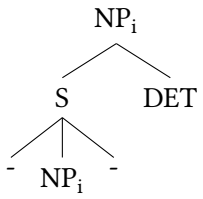
In (15) we see that Williamson's claim that the head of an IHRC cannot be marked definite holds true for Hidatsa. The definite determiner on the head of the IHRC makes the sentence ungrammatical.

[4] It should be noted that while the head of an IHRC cannot be marked as definite, the head may take either an indefinite determiner or no determiner at all.

4.2 Williamson's analysis of Lakota

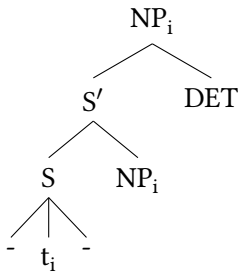
Williamson's (1987) analysis of the relative structure is that shown in (16). In this structure the S^5 has a determiner for a sister.

(16) **Williamson's (1987) model at S-structure**



This analysis is what Williamson posits for a representation of the overt syntax, or S-Structure. We can see that the IHRC acts as a sentence or subordinate clause in its own right. This entire structure then acts as an argument or DP for the matrix verb. Williamson posits that the internal head obligatorily moves outside of the IHRC at LF, giving the structure shown in (17).

(17) **Williamson's (1987) model after movement at LF**



Williamson further posits a cyclic rule that co-indexes the internal head with the DP dominating the IHRC. Although Williamson's syntactic analysis is adequate, both Hoeksema (1989) and Culy (1990) point out that the movement at LF is unmotivated in general. Nonetheless, many researchers have adopted Williamson's analysis. Williamson's major insight into the structure of IHRCs is

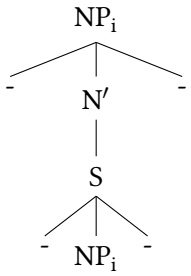
[5] Although I will use a Minimalist framework and the associated terminology for my analysis, I will employ the older terminology used by the various authors cited so as to not confuse their arguments. For example many authors use S for TP and S′ for CP. This older terminology will be retained when showing older examples written in early P&P or pre-P&P frameworks.

her observation and explanation of the fact that the head of an IHRC cannot be marked with a definite determiner.

4.3 Culy's analysis of IHRCs

Culy (1990) describes the syntactic structure of an IHRC as that shown in (18).

(18) **Culy's (1990) structure of an IHRC**.



The main problem with this structure is that it violates the endocentric constraint of X-bar syntax, which states that all phrases must project a head and all heads must project a phrase. In the schema proposed above in (18), the N′ exhaustively dominates an S, not an N. Culy claims, "while this structure is unusual, it is similar to a rule proposed by Jackendoff (1977) that allows 'category-switching' ". That rule is:

$$x^i \rightarrow af \rightarrow y^i$$

The difference, according to Culy, is that in an IHRC construction the two categories N′ and S are not at the same level. Culy's argument for this structure is as follows: N′ dominates S because IHRCs are nominalized sentences and as such occur with the elements of a NP that also occur with N′. That is to say, IHRCs have an internal structure of S but an external distribution of N′.

Culy also states that there does not seem to be any framework-independent evidence about whether the IHRC is an S (TP) or an S′ (CP) and that there is no strong evidence that IHRCs have overt complementizers.

Culy makes one other important prediction about IHRCs that is relevant to this discussion. He proposes that IHRCs have a *wh* element which is not overt. He uses this notion to provide motivation for the movement at LF that Williamson proposes. As shown in (18), Culy (like Williamson) argues that there is coindexation between the DP dominating the IHRC and the internal DP being modified

(the head). This coindexation is similar to that which exists between a relative pronoun and its antecedent. Culy bases his argument on Safir (1986) who proposes that in English relative clauses without relative pronouns, there is a null *wh* operator[6] which functions like a relative pronoun. Given this assumption, Culy shows that the relative clause in (19a) has the S-structure shown in (19b).

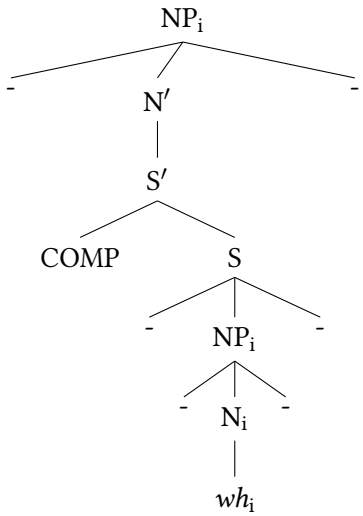
(19) **English relative clause without a relative pronoun**

a. the dog I ran away from

b. $[_{NP}$ the dog_i $[_{S'}$ wh_i $[_S$ I ran away from e_i]]]

Culy proposes that any common noun can optionally act as a *wh* operator in languages with IHRCs. This operator then moves from the head to COMP (or C) at LF via the rule of *wh* construal, just as in-situ *wh* elements must move at LF. Culy then proposes a tree structure for both D-structure and S-structure that is the same in all relevant respects. This is shown in (20).

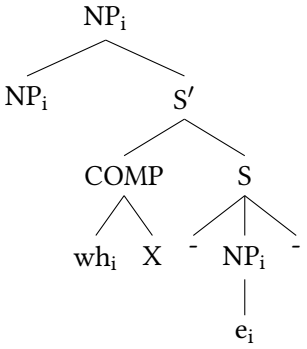
(20) **D- and S-structure of an IHRC (Culy 1990)**



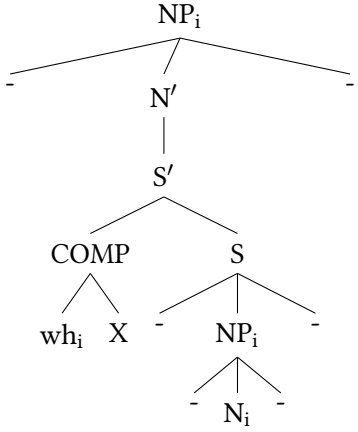
Additionally, to account for the construction in (20), Culy proposes a general rule that coindexes the NP (that is the IHRC) and the *wh* operator. To do this he proposes the LF structures for EHRCs shown in (21) and the LF structure for IHRCs shown in (22).

[6] Safir terms this operator an "abstract A′ binder", which he then represents as a *wh*.

(21) **LF structure for an EHRC (Culy 1990)**

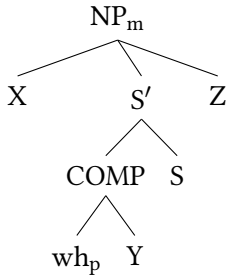


(22) **LF structure of an IHRC (Culy 1990)**



Culy points out that these structures both have an NP dominating an S′ with a *wh* element in its COMP that is coindexed with the NP internal to the clause. He further states that while it has generally been assumed that coindexation in EHRCs is between a *wh* element and an NP that is the head, this is not a necessary assumption. The same effect can be accomplished by coindexing the *wh* element with the NP dominating the relative clause, since this NP will have the same index as its daughter NP by general feature passing conventions (i.e. that a head and its mother share the same features). By taking this approach, Culy subsumes coindexation in EHRCs and IHRCs under the same rule, which he formalizes as the Relative Coindexing Constraint, shown in (23).

(23) **Relative Coindexing Constraint (RCC) (Culy 1990)**
It must be the case that m = p.



This allows a generalization about the coindexing that occurs in both EHRCs and IHRCs, thus providing a unifying account of both internally headed and externally headed relative clauses.

5 The syntax of the Hidatsa IHRC

After having reviewed the above approaches to IHRCs, a theoretical analysis of the Hidatsa relative clause is now possible. As shown above in §4.1, Hidatsa obeys the indefiniteness restriction first proposed by Williamson. In this respect, it is typical of IHRCs in other languages. However, Hidatsa is unusual in that its RCs are prefixed with a relative marker that distinguishes specificity of the nominalized clause. In §5.1, I will show that these prefixes are not relative pronouns. In §5.2, I will argue that they are complementizers and provide a syntactic analysis of IHRCs in a Minimalist framework.

5.1 The syntactic status of the Hidatsa relative markers

The Hidatsa RC is a nominalized sentence with the structure shown above in (6) and repeated here as (24).

(24) [(DP) (DP) REL-verb]-DET

In addition to overt DPs, the argument positions can also be filled with a *pro*. Like full DPs, *pro* triggers number and person agreement on the verb. Given this fact, the *aku-* and *aru-* relative markers cannot be relative pronouns in the common sense. That is to say, they do not serve the same function as relative pronouns in languages like English. They do not serve as arguments. Consider again example (9) repeated here as (25).

(25) *Mary uuwáki akuhíriš* *warúcic.*
Mary uuwáki aku-híri-š wa-rúci-c
Mary quilt REL.S-make-DET.D 1A-buy-DECL
'I bought the quilt that Mary made.'

In this example, we see that the matrix verb *rúci-* 'buy' projects two arguments and two θ roles. The argument positions are filled by a *pro* subject, which is first person (shown by the agreement marker *ma-*), and the relative clause itself. Likewise, the verb in the relative clause, *híri-* 'make', also projects two arguments and θ roles. These are filled by the subject of the clause, *Mary*, and by the object, *uuwáki* 'quilt'. If *aku-*, and by extension *aru-*, were relative pronouns, they would fill an argument slot, but this is not possible since both verbs project only two arguments each, and these are filled with either full DPs or *pro*. It is thus not possible for either of the Hidatsa relative markers to be a relative pronoun. If this was their role, it would be a violation of the theta criterion (Chomsky 1981: 36). As they cannot be relative pronouns, I propose that they are complementizers. They signal that the RC is a complement clause of a superordinate verb and in that role they function like the English complementizer *that*.

5.2 The Hidatsa relative markers as complementizers

Culy (1990) claims that 1) no languages have overt complementizers in IHRC constructions; and 2) as a result of this lack of overt complementizers, there is no framework independent evidence as to whether the IHRC is an S (TP) or an S′ (CP).

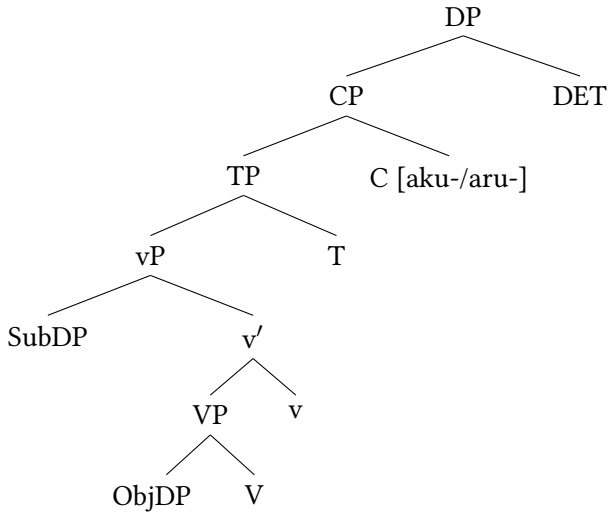
All Siouan languages have an extensive set of morphemes that serve as clause final markers. In Hidatsa, a predicate must have one of these markers to be grammatical. In main clauses, these morphemes show illocutionary force: declarative *-c*, emphatic *-ski*, speculative *-tóok*, past definite *-št*, and permission *-ahka* among others (Matthews 1965; Boyle 2007). Following Rizzi (1997), I assume that complementizers are the syntactic elements that express this type of illocutionary information. As illocutionary force markers, these morphemes indicate the clause type (Cheng 1997) or Force (Chomsky 1995) of the sentence and must be projected in C. In subordinate clauses, these morphemes can include complementizers that include conditional and temporal subordinators. As in matrix clauses, these morphemes are also projected in C.[7]

[7] Hidatsa also has co-subordinate clauses (Van Valin & LaPolla 1997). These clauses are connected with switch-reference (SR) markers. Boyle (2007) argues that these are coordinate structures with the SR markers coordinating vPs (same-subject markers) or TPs (different subject

Only relative clauses do not have a clause final illocutionary force marker. While there is a determiner that nominalizes the entire clause, which appears to be clause final, I propose that this is not the case. That determiner is the head of the DP that takes the RC as its complement. The determiner is thus outside of the RC. It is not a complementizer. To unify the syntactic structure of Hidatsa, I believe we can analyze the RC markers *aku-* and *aru-* as complementizers that mark the clause type as relative (REL). This feature satisfies the underspecified clause type feature in T. If this analysis is correct, we now have evidence that IHRCs are CPs, not TPs. The mechanics of how these complementizers function will be detailed below.

While Culy needs the structure presented in (18) above for his analysis, it is highly unusual as it is exocentric — that is to say, the head does not project a phrasal level. In (18), the S projects an N′ and then NP. Culy's analysis could have been simplified but he employs older syntactic notations that do not allow him to capture greater generalities. The analysis presented here follows from the structure proposed by Williamson with only slight alterations. An IHRC in Hidatsa has the base-generated structure shown in (26).

(26) **Proposed structure for IHRCs in Hidatsa**



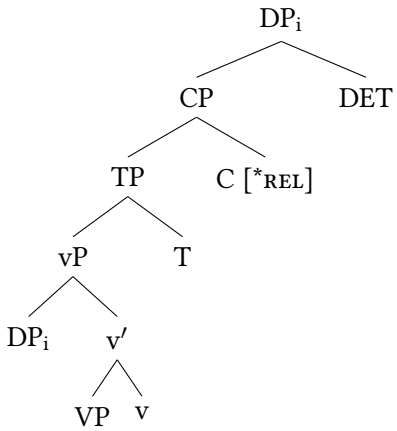
markers). These co-subordinate clauses receive their illocutionary force from the matrix verb. See Gordon & Torres (2012) for a similar analysis of SR marking in Mandan.

This structure allows for a straightforward syntactic analysis without postulating an exocentric construction.[8] It captures the insights of both Williamson and Culy. In addition, it shows the placement of the Hidatsa complementizers, *aku-* and *aru-*, as projections of C. These relative complementizers surface as prefixes to the verb. The motivation for this ordering is discussed below. The overt syntactic data provides framework-independent evidence for the CP status of IHRC. If this analysis is correct, it is not consistent with Culy's claims presented at the beginning of this section.

5.3 Hidatsa *aku-* and *aru-* as strong features in C

Above I have argued that Hidatsa has overt complementizers in relative clauses, namely the *aku-* and *aru-* morphemes. In Hidatsa IHRCs, these morphemes carry strong features, and as such they trigger movement of the verbal complex to T and then C. It is important to note that this type of cyclic roll-up is head-to-head movement, rather than phrasal movement. Consider the structure (shown in example (27) of the Hidatsa IHRC, here simplified for only one DP in the relative clause.

(27) **Proposed structure for IHRCs in Hidatsa**



Culy (like Williamson) argues that there is coindexation between the DP dominating the IHRC and the internal DP being modified (the head). In Hidatsa, the head of the RC is coindexed with the DP that dominates the entire IHRC.

[8] In Hidatsa DPs select NPs or nominalized verbs, which include relative clauses, as complements.

Evidence of coindexation between the head of the IHRC and the DP dominating the IHRC can be seen in number agreement.[9] An example of this agreement is shown in (28) and (29).

(28) **Plural agreement with overt subject**
wíaku?o *uuwáki akuhíra?aš* *íiwia?ac*
wía-aku-?o uuwáki aku-híri-?a-š íiwia-?a-c
woman-DET.S-PL.I quilt REL.S-make-PL.D-DET.D cry-PL.D-DECL
'[Those women who made the quilt] cried.'

(29) **Plural agreement with null subject**
uuwáki akuhíra?aš *íiwia?ac*
uuwáki aku-híri-?a-š íiwia-?a-c
quilt REL.S-make-PL.D-DET.D cry-PL.D-DECL
'(The ones who) made the quilt cried.'

In these examples, we see that the matrix verbs agree with their subjects, which are IHRC. The head in each of these clauses, whether overt (as in 28) or null (as in 29), is plural and this plural number is marked on both the matrix and subordinate verbs. Thus, both the head and the IHRC must have the same number feature. Coindexation is the usual way for two DPs to have the same number marking. The overt interaction of these elements offers convincing evidence that there is indeed a coindexing relationship between the head and the IHRC.

5.4 Move and Merge in the Hidatsa IHRC

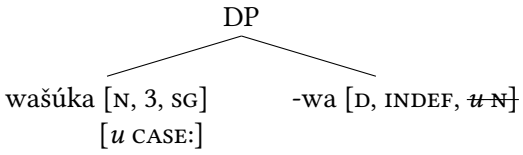
Consider again example (11) repeated here as (30).

(30) **Hidatsa internally headed relative clause**
[*wacéeš* *wašúkawa* *akutíheeš*] *šipíšac*
[wacée-š wašúka-wa aku-tí-hee-š] šipíša-c
[man-DET.D dog-DET.I REL.S-die-3.CAUS.D.SG-DET.D] black-DECL
'The dog [that the man killed] is black.'

The direct object is built by the indefinite determiner selecting for a NP. These two elements merge to form a DP as shown in (31).

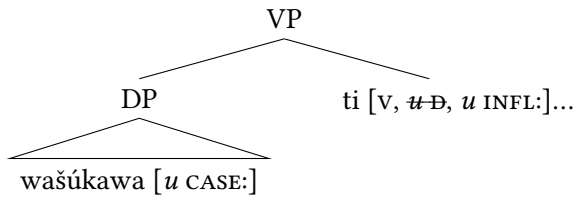
[9] In some languages, person agreement also provides evidence for this coindexation. While Williamson shows this evidence in Lakota, it is not seen in Hidatsa since the type of person agreement morphology that exists in Lakota does not exist in Hidatsa.

(31)



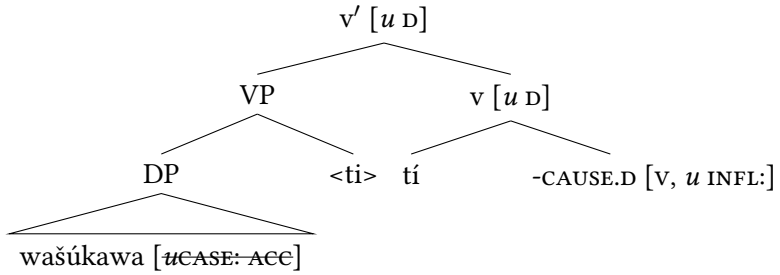
The verb *ti-* 'die' then selects for an argument, satisfying the [*u* D] feature of the verb as shown in (32).

(32)



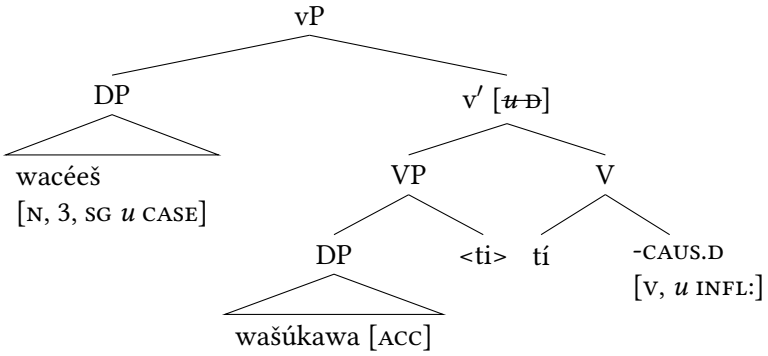
In Hidatsa, the direct causative is projected as the head of V. Direct causatives take stative verbs as their complements and add an agentive argument to the subcategorization of the verb. The lexical verb then moves up to adjoin to the v forming the derived verb *ti+CAUS.D-* and projecting a v′. The verb *ti+CAUS.D-* in Hidatsa is literally 'cause to die'. Accusative [ACC] case is then licensed and checked on the DP *wašukawa*, making it a direct object as shown in (33).

(33)



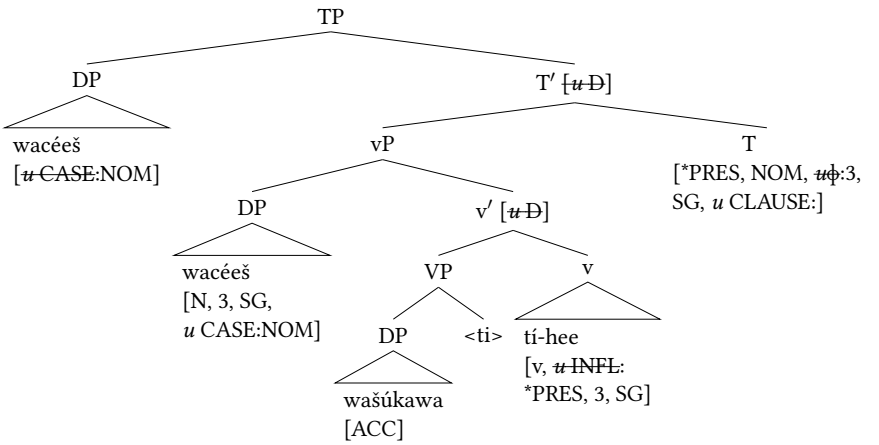
The [*u* D] feature in v is passed up to v′ and projects a SPEC position where the agent is merged, thus checking the [*u* D]. This is shown in (34).

(34)]



The T head then merges with vP. T has the features [*PRES, NOM, *EPP ([*u* D]), *u* ϕ: , *u* CLAUSE:]. This gives us the structure shown in (35).

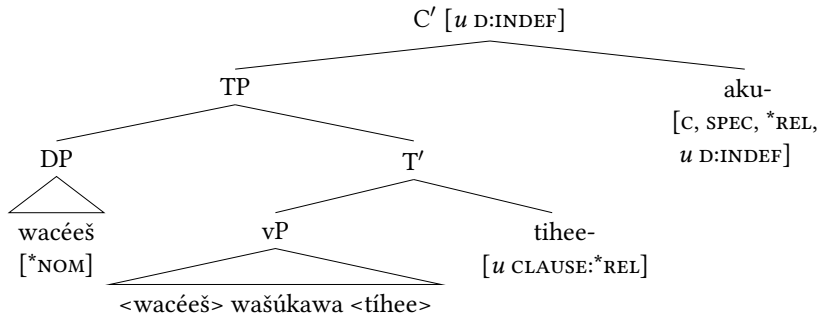
(35)



In this structure, nominative case [NOM] is passed down from T to value the [*u* CASE:] feature of the DP in [SPEC, VP]. This agentive DP, *waceeš* 'the man', then passes its phi-features up to value the [*u* ϕ:] in T as [3, SG]. The features [*PRES, 3, SG] are then passed down to value the [*u* INFL:] feature on the verb. Once the CAUS.D is valued for person and number, it is realized as *hee-* at Spellout. The strong [*PRES] tense feature then attracts the verb to T where the tense features are checked (shown below in 36). The EPP feature attracts the agentive DP to [SPEC, TP] where the [NOM] case feature is checked and the [EPP (*u* D]) feature is satisfied. These are all straightforward Move and Merge operations.

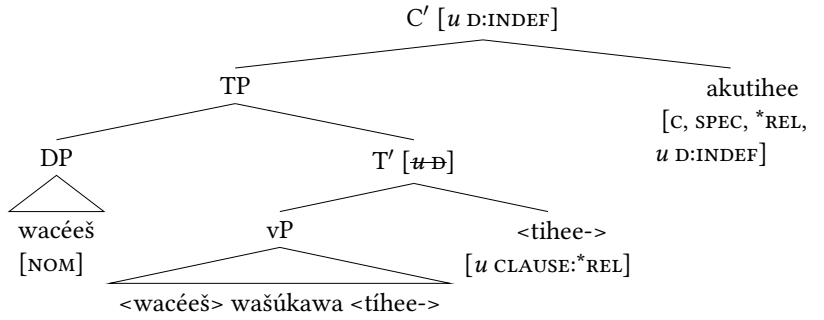
The TP then merges with a C. Normally, if there is no [*wh* feature] in C, there is no [SPEC, CP] position projected. Complementizers for IHRCs are different. I propose that the Hidatsa relative complementizers have the features [C, *REL, *u* D:INDEF]. Additionally, the *aku-* complementizer has a value of [SPEC] (specific) and the *aru-* complementizer has a value of [NON.SPEC] (nonspecific). This gives us the structure shown in (36).

(36)



As the [*REL] illocutionary feature in C is strong it triggers Move of the verbal head in T, so that this feature can be checked. The verb moves to C where the *aku-* is prefixed to it. This is shown in (37).

(37)

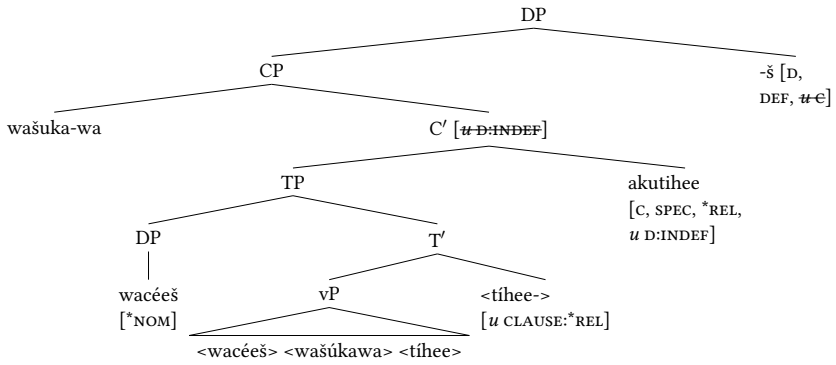


Normally, illocutionary features in C are weak and don't trigger movement. Because of this the illocutionary force markers concatenate onto the verbal complex as a suffix due to roll-up movement.

The [*u* D:INDEF] feature in C moves up to C′ and creates a [SPEC, CP] position. This is a weak feature, so movement of the indefinite DP is triggered after Spell-

out at LF. The semantic motivations for this will be discussed below in §6. The determiner then nominalizes the clause, selecting for a C rather than N, giving us the complete relative clause structure. This is shown in (38).

(38)



6 Semantic constraints and motivations

Prior to Williamson (1987) there was very little in the way of semantic explanation for IHRCs. Williamson and most linguists after her built on Heim's (1982) ideas about definites and indefinites and applied them to IHRCs. According to Heim, indefinites act not as quantifiers, but as variables. In this section, I will review Heim's basic assumptions. I will then review Williamson's account of IHRC (1987) as well as Basilico's (1996) account, which employs not only Heim's ideas but also Diesing's mapping theory (1990; 1992a,b) to account for how IHRCs function at LF. Adopting Heim's framework, and building on Williamson and Basilico, I develop an account of IHRCs that simplifies previous work and provides an explanation as to the nature of IHRCs and their heads with regard to definiteness and movement at LF. I then show how this account unifies our understanding of IHRCs and EHRCs.

6.1 Heim's account of indefinite determiners

Heim's (1982) dissertation has proven to be very important in the theoretical explanation for the semantics of IHRCs. This work explores how the logical form of a sentence is constructed. Although Heim accepts the commonly held view that noun phrases headed by a common noun are generalized quantifiers, her major

contribution to semantic analysis (particularly for IHRCs) is that indefinites are variables, not quantifiers. So a sentence like (39a) will have the semantic representation seen in (39b):

(39) **DPs as Generalized Quantifiers**
   a. Every dog is barking
   b. every [(dog(x)) (is-barking(x))]

In (39b), 'every' is the quantifier, '(dog(x))' is the restriction of the quantifier, and '(is barking(x))' is the nuclear scope of the quantifier.

According to Heim (1982), definites and indefinites can be distinguished by three properties. First, only indefinites can undergo Operator Indexing. This means that an indefinite can be indexed with another element in the sentence. Definites, along with proper names and pronouns, are not subject to Operator Indexing. Second, only indefinites are constrained by the Novelty Condition. The Novelty Condition states that an indefinite NP must not have the same index as any NP to its left. Third, definites, but not indefinites, presuppose their descriptive content, if they have any. That is to say, a definite presupposes the existence of an entity with the properties of its descriptive content, while an indefinite does not. The crucial property in understanding the distribution of determiners in IHRCs is that only indefinites undergo Operator Indexing. Lastly, as indefinites are variables and not quantifiers, they must be bound by an existential operator inserted in the sentence by an operation she calls Existential Closure.

6.2 Williamson's treatment of IHRCs

Williamson (1987) claims that all languages that have IHRCs will have an indefiniteness restriction. According to this restriction, only indefinite NPs may be heads in an IHRC. This is to say that the head of an IHRC cannot be marked with a definite determiner. Williamson claims that the indefiniteness restriction cannot be attributed to some inherent (i.e. lexical) property requiring wide scope of the indefinite NP. In addition, one cannot attribute this restriction to the traditional distinction between quantifiers, on the one hand, and proper nouns and definite NPs, on the other. To understand the indefiniteness restriction, Williamson claims that we must understand that both simple declaratives containing an indefinite and RCs indicate the intersection of two sets. The traditional view of this can be seen in (40) and (41).

(40) I bought a dog.
   $\exists$ x (Dog (x) & Buy (I, x))

(41) dog that I bought
(Dog (x) & Buy (I, x))

In (40) we see a proposition with a bound variable and in (41) we see a propositional function with a free variable. Williamson suggests that we reconsider the traditional view of indefinites as existential quantifiers. Following Heim (1982), Williamson proposes that indefinites are 'quantifier-free'. That is to say that they are essentially free variables. This then gives the example in (40) the semantic interpretation of the example in (41). The quantifier force of indefinites in simple declaratives is determined by the rule of Existential Closure. Thus, IHRCs have the interpretation of a propositional function. Williamson claims that universal quantifiers are excluded as heads because semantically such a quantifier is interpreted as a restrictive term. A definite is familiar (known) and presupposes the content of its predicate. This property is at variance with the meaning of restrictive RCs, for if the head is already familiar to the hearer, further specification by the RC is, at best, unnecessary. While I agree with Williamson's analysis of IHRCs, she does not provide a motivation for why it is true.

6.3 Basilico's account of IHRCs and Diesing's mapping theory

Basilico (1996) notes that most theorists working in a transformational framework posit that the internal head moves to an external position at some point in the derivation. Examples of this have been shown above with the work of Cole (1982; 1987), Williamson (1987), Culy (1990), Kayne (1994) and Bianchi (1999), among others. With the exception of Kayne and Bianchi, most researchers working on IHRCs have posited that this movement takes place at LF (or its predecessor D-Structure).

Prior to Basilico (1996) there were two general approaches to head movement in IHRCs. In the first approach, advocated by Broadwell (1985; 1987), Cole (1987), Lefebvre & Muysken (1988), and Cole & Hermon (1994), the head moves to a position external to the CP of the relative clause. The second approach, advocated by Williamson (1987), Brass et al. (1989), and Bonneau (1992), postulates that the head moves to the [SPEC, CP] of the RC but not out of the clause itself. While all of these works have arguments supporting the nature of the movement, none of them provide a detailed explanation as to why the head needs to move.

Basilico presents evidence that in some languages with IHRCs, movement of the head occurs in the overt syntax. He argues that the head need not necessarily move to a position external to the clause and that while the head is not in its usual place it nevertheless remains within the RC in the overt syntax. Drawing from the previous work of Williamson (1987), Jelinek (1987), and Culy (1990), Basilico

adopts the notion that IHRCs are not cases of relativization semantically, but cases of quantification. Following Heim (1982), Basilico argues that IHRCs are associated with quantificational elements that bind variables within the subordinate clause itself. The sentential part of the IHRC is interpreted semantically as an open sentence. According to Basilico the Hidatsa relative clause in (42) would have the semantic interpretation shown in (43).

(42) *wáceewa akuʔawákaaš*
wácee-wa aku-awákaa-š
man-DET.I REL.S-1A.see-DET.D
'the man that I saw'

(43) ι x [man (x) & I saw (x)]

In this example, the sentential part of the IHRC 'man I saw' should be interpreted semantically as *man (x) & I saw (x)*, an open sentence with two unbound variables. According to Basilico, the definite determiner *-š* functions as an (iota) operator that binds the variables within the relative clause.[10] Following Culy (1990), the sentential part of the IHRC functions as the restriction on the operator associated with the relative clause.

In this analysis, one of the variables associated with the sentence is provided by the head noun. The importance of this, namely the indefiniteness restriction on the head NP, was first noted by Williamson (1987). She showed that the head NP in IHRCs is not allowed to be marked as definite. According to Basilico, this follows from Heim's (1982) analysis that indefinite NPs are not associated with quantificational force (as presented above) and Kratzer's (1989) Prohibition Against Vacuous Quantification. In a similar manner to Culy (1990), Basilico follows Heim (1982) in treating indefinites as having no quantificational force. He argues that they provide only a variable, which must be bound by another operator in the representation. In IHRCs this operator is the determiner associated with the entire IHRC itself; it comes to bind the variable associated with the indefinite head. Basilico argues that if there was a definite marker on the head then the variable provided by the head would be unavailable for binding. Since the operator associated with the IHRC would not bind a variable, this would be a violation of the prohibition against vacuous quantification (as shown in 44).

(44) For every quantifier Q, there must be a variable x such that Q binds an occurrence of x in both its restrictive clause and its nuclear scope (Kratzer 1989).

[10] See Jelinek (1987) for the proposal concerning the use of the iota operator with IHRCs.

Since the sentential part of the IHRC forms the restriction on the operator, there would be no variable for the operator to bind if there were no indefinite within the subordinate sentence to provide this variable.

Basilico then goes on to apply Diesing's (1990; 1992a; 1992b) Mapping Hypothesis to the head movement in IHRCs. Her mapping hypothesis, which holds at LF, proposes two notions:

(45) a. Material from VP (vP) maps into nuclear scope, which is the domain of existential closure.
b. Material from TP maps into a restrictive clause.

A restrictive clause is that part of the representation which forms the restriction on some operator. That is, an indefinite that restricts some operator will be in a different syntactic position at LF than an indefinite that receives an existential interpretation by VP-level existential closure. The former indefinite NPs (the heads of the IHRCs) must not be within the VP at LF, while the latter must be in the VP at LF (Basilico 1996). Therefore, the indefinite head of an IHRC must move out of VP simply because it is indefinite. RCs are quantificational and are selected by the determiner. The head must move out of its argument position in order for the quantificational operator that is associated with it (as shown by the final determiner of the RC) to bind the variable introduced by the head (which is indefinite or null).

Basilico argues that there is an operator associated with the IHRC, which must come to bind the variable associated with the indefinite head. In order for an indefinite to become bound by an operator and not undergo existential closure, it must move out of the VP by LF. Thus, the quantificational approach to IHRCs and the mapping hypothesis provide a motivation for head movement. The head must move in order to be bound by the operator associated with the IHRC. If there is no head movement, and the head remains in the VP, then there will be no variable to bind, and as a result, this will violate the prohibition against vacuous quantification.

Basilico, like others before him, claims that IHRCs are DPs. Like other DPs, IHRCs can appear as arguments. For Basilico the difference between IHRCs and other DPs lies in what the head D of the DP takes as its complement. Noun phrase DPs must take NPs as the complement to the head D; this NP functions as the restriction on the head of D. IHRC DPs on the other hand take sentences (TPs) as their complements and these sentences function as the restriction on the head D (Basilico 1996).

Unfortunately, Basilico is only examining languages (Diegueño, Mojave, and Cocopa) that have evidence of movement in the overt syntax of their IHRCs. As a result of this, the structures he posits show movement taking place prior to Spellout. In these structures, either the head or the entire IHRC move, thus allowing the head to escape existential closure.

By only examining IHRCs that show some evidence of movement overtly in the syntax, Basilico avoids the more general consideration of what happens in languages with IHRCs that show no evidence of movement. As a result, he need not posit any structure for the majority of languages with IHRCs where head movement is done covertly at LF.

6.4 The semantic motivation for movement at LF

Consider again example (11) repeated here as (46).

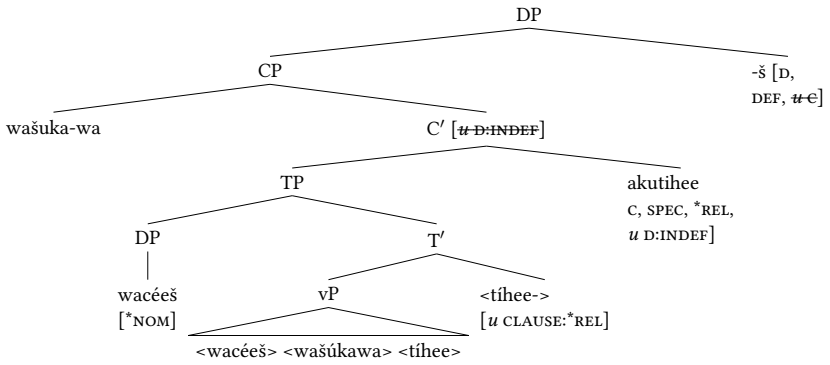
(46) **Hidatsa internally headed relative clause**

[*Wacéeš*	*wašúkawa* *akutíheeš*]	*šipíšac.*
[wacée-š	wašúka-wa aku-tí-hee-š]	šipíša-c
[man-DET.D	dog-DET.I REL.S-die-3.CAUS.D.SG-DET.D]	black

'[The dog that the man killed] is black.'

This is an unambiguous IHRC. In this sentence the head of the RC is *wašúkawa* 'a dog'. The structure of this sentence given above in example (38) is repeated here as (47).

(47)



As Basilico (1996) has shown, relative clauses are quantificational. A determiner selects a relative clause as its complement. In addition, all IHRCs are re-

strictive, which means it is part of the representation which forms the restriction on some operator. An indefinite that restricts some operator must be in a different syntactic position at LF than an indefinite that receives existential interpretation by VP-level existential closure. Given this, the head of the IHRC, which must be indefinite, must not be within VP at LF. The head must move out of its argument position in order for the quantificational operator that is associated with it (as shown by the final determiner on the RC) to bind the variable introduced by the head (which is indefinite or null). The indefinite determiner is not associated with any quantificational force: it is an identity function. The semantics of the Hidatsa indefinite can be seen in (48).

(48) [-wa] = λ P<et>.P

Given this, the head must move; if it does not, it will not escape the existential closure of the VP. The head must be indefinite (or generic with null morphology) if it is to be bound by this outside operator. Given that the head of the IHRC is in-situ at Spellout, any movement must take place at LF. If this movement does not take place, the derivation will crash.

Although Basilico (1996) postulates structures with either IP or VP adjunction, this cannot be correct for Hidatsa, as IHRCs show no evidence of movement. However, I have postulated that Hidatsa does show clear evidence for a CP structure as the overt relative markers act as complementizers. This overt evidence for a complementizer shows that Culy (1990) was correct in postulating the complementizer position for IHRCs and while it is rarely filled overtly in many of the world's languages that have these structures, it is in Hidatsa. In Hidatsa the head of the IHRC moves to the [SPEC, CP] position at LF.

6.5 A unified account of IHRCs and EHRCs

Given the above analysis, we can see that IHRCs and EHRCs are remarkably similar. EHRCs serve as complements (restrictive RCs) or adjuncts (non-restrictive RCs) to NPs, which are arguments of a predicate. IHRCs are DPs that serve as arguments of a predicate. In EHRCs, the head that is modified by the clause is coindexed with an element inside the RC. This is either a relative pronoun or an operator. This relative pronoun or operator is coindexed with the head outside of the RC. In IHRCs, the head stays inside of the RC, but it is coindexed with the determiner that takes the IHRC as its complement. This head must move to [SPEC, CP] at LF to escape existential closure, just as the relative pronoun or operator moves to [SPEC, CP] in an EHRC.

7 Conclusion

In this paper I have shown that Hidatsa has IHRCs. I have examined previous attempts at describing their syntactic structure and provided a new one based on data from Hidatsa. In addition, I have provided theory external evidence for the possibility of complementizer in these clauses (namely the *aku-* and *aru-* markers). Following Culy (1990), I have argued that this complementizer has a strong feature, and that in Hidatsa this is different from all other complementizers triggering movement. This accounts for the morpheme order in Hidatsa relative clauses. Following previous work on IHRCs (most notably Williamson 1987 and Culy 1990) I have expanded and simplified how the semantics of IHRCs functions. I have provided motivation for the movement of the head of the IHRC at LF in addition to explaining why Williamson's indefiniteness restriction holds true.

Acknowledgment

I would like to thank the late Alex Gwin for the many hours spent discussing Hidatsa. He was a wise and generous man. I would also like to thank Jason Merchant and Catherine Rudin for their comments on an earlier draft of this paper as well as Brian Agbayani's comments on a later draft. Thank you also to two anonymous reviewers for their helpful comments. Lastly, I would like to thank Robert L. Rankin, who was always so encouraging and helpful to students and junior scholars.

Abbreviations

1.A	first person active	EPE	epenthetic consonant
3.CAUS.D.SG	third person causative	IHRC	internally headed relative clause
3.POSS.A	third person alienable possessive	INDEF	indefinite
CONT	continuative	LOC	locative
DECL	declarative	NE	narrative ending
DET.D	definite determiner	NEG	negative
DET.I	indefinite determiner	PL.D	plural definite
DET.S	specific determiner	PL.I	plural indefinite
EHRC	externally headed relative clause	RC	relative clause
EMPH	emphatic	REL.N	nonspecific relative
		REL.S	specific relative
		TEMP	temporal
		TOP	topic

References

Basilico, David. 1996. Head position and internally headed relative clauses. *Language* 72(3). 498–532.

Bianchi, Valentina. 1999. *Consequences of antisymmetry: Headed relative clauses* (Studies in Generative Grammar). New York: de Gruyter Mouton.

Bonneau, José. 1992. *The structure of internally headed relative clauses*. Montréal: McGill University (Doctoral dissertation).

Boyle, John P. 2007. *Hidatsa morphosyntax and clause structure*. Chicago: University of Chicago (Doctoral dissertation).

Brass, Andrew, Kenneth Hale, Ellavina Perkins & Margaret Speas. 1989. Aspects of logical form in Navajo. In Eung-Do Cook & Keren Rice (eds.), *Studies in athabaskan syntax*, 317–34. Berlin: de Gruyter Mouton.

Broadwell, George Aaron. 1985. Internally headed relative clauses in Choctaw. *Kansas Working Papers in Linguistics* 6. 16–24.

Broadwell, George Aaron. 1987. A-bar anaphora and relative clauses. *Proceedings of the North East Linguistics Society* 16. 47–56.

Cheng, Lisa Lai Shen. 1997. *On the typology of Wh-questions* (Outstanding Dissertations in Linguistics). New York: Garland Publishing.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1998. *Minimalist inquiries: The framework* (MIT Occasional Papers in Linguistics). Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Beyond explanatory adequacy. *MIT Occasional Papers in Linguistics* 20. 1–141.
Citko, Barbara. 2001. Deletion under identity in relative clauses. In Minjoo Kim & Uri Strauss (eds.), *Proceedings of the North East Linguistics Society*, vol. 31, 1–15.
Cole, Peter. 1982. *Imbabura Quechua* (Lingua Descriptive Studies 5). Amsterdam: Elsevier.
Cole, Peter. 1987. The structure of internally headed relative clauses. *Natural Language and Linguistic Theory* 5(2). 277–302.
Cole, Peter, Wayne Herbert & Gabriella Hermon. 1982. Headless relative clauses in Quechua. *International Journal of American Linguistics* 48(2). 113–124.
Cole, Peter & Gabriella Hermon. 1994. Is there LF wh-movement? *Linguistic Inquiry* 25. 239–262.
Culy, Christopher. 1990. *The syntax and semantics of internally headed relative clauses*. Stanford: Stanford University (Doctoral dissertation).
Cumberland, Linda A. 2005. *A grammar of Assiniboine: A Siouan language of the northern Plains*. Bloomington: Indiana University (Doctoral dissertation).
Di Sciullo, Anna Maria. 2005. *Asymmetry in morphology*. Cambridge, MA: MIT Press.
Diesing, Molly. 1990. *The syntactic roots of semantic partition*. Amherst: University of Massachusetts – Amherst (Doctoral dissertation).
Diesing, Molly. 1992a. Bare plural subjects and the derivation of logical representations. *Linguistic Inquiry* 23. 353–380.
Diesing, Molly. 1992b. *Indefinites*. Cambridge, MA: MIT Press.
Drummond, Valerie E. 1976. *Relative clauses in Sioux Valley Santee and Carry-the-Kettle Assiniboine*. Toronto: University of Toronto (Master's thesis).
Fauconnier, Gilles. 1979. *Theoretical implications of some global phenomena in syntax*. New York: Garland Publishing.
Gorbet, Larry. 1976. *A grammar of Diegueño nominals*. New York: Garland Publishing.
Gordon, Zachary & Silvana Karin Torres. 2012. *Reanalysis of the Mandan verbal suffix complex through the complementizer phrase*. Paper presented at the annual meeting of the Society for the Study of Indigenous Languages of the Americas and the Linguistic Society of America, Portland, OR.

Graczyk, Randolph Jan. 1991. Relative clauses in Crow. In Frances Ingemann (ed.), *Proceedings of the 1990 Mid-America Linguistics Conference and Siouan and Caddoan Languages Conference*, 490–504. Lawrence: University of Kansas.

Graczyk, Randolph Jan. 2007. *A grammar of Crow*. Lincoln: University of Nebraska Press.

Hale, Kenneth Locke & Paul Platero. 1974. Aspects of Navajo anaphora: Relativization and pronominalization. *Diné Bizaad Náníl'iih (Navajo Language Review)* 1. 9–28.

Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Amherst: University of Massachusetts – Amherst (Doctoral dissertation).

Hoeksema, Jack. 1989. Review of "The representation of (in)definiteness". *Language* 65(1). 115–125.

Jackendoff, Ray. 1977. Constraints on phrase-structure rules. In Peter W. Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 249–283. New York: Academic Press.

Jelinek, Eloise. 1987. Headless relatives and pronominal arguments: A typological survey. In Paul Kroeber & Robert E. Moore (eds.), *Native American languages and pronominal arguments*, 136–148. Bloomington: Indiana University Linguistics Club.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kratzer, Angelika. 1989. Stage-level and individual-level predicates. In Emmon Bach, Angelika Kratzer & Barbara Partee (eds.), *Papers on quantification*, 225–264. Amherst: University of Massachusetts – Amherst.

Lefebvre, Claire & Pieter Muysken. 1988. *Mixed categories: Nominalization in Quechua*. Dordrecht: Kluwer.

Lowie, Robert H. 1939. *Hidatsa texts (with grammatical notes and phonograph transcriptions by Zellig Harris and C. F. Voegelin)*. Vol. 1 (Prehistory Research Series 6). Indianapolis: Indiana Historical Society.

Matthews, George Hubert. 1965. *Hidatsa syntax*. Berlin: de Gruyter Mouton.

Quintero, Carolyn. 2004. *Osage grammar* (Studies in the Anthropology of North American Indians). Lincoln: University of Nebraska Press.

Rankin, Robert L., John P. Boyle, Randolph Jan Graczyk & John E. Koontz. 2003. A synchronic and diachronic perspective on 'word' in Siouan. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A cross-linguistic typology*, 180–204. Cambridge: Cambridge University Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 213–276. Dordrecht: Kluwer.

Rood, David S. & Allan R. Taylor. 1996. Sketch of Lakhota, a Siouan language. In William C. Sturtevant & Ives Goddard (eds.), *Languages* (Handbook of North American Indians 17), 440–482. Washington: Smithsonian Institution.

Safir, Kenneth. 1986. Relative clauses in a theory of binding and levels. *Linguistic Inquiry* 17(4). 663–89.

Van Valin, Robert D., Jr. & Randy J. LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.

Weber, David. 1983. *Relativization and nominalized clauses in Huallaga (Huanuco) Quechua* (University of California Publications in Linguistics 103). Berkeley: University of California Press.

Williamson, Janis Shirley. 1987. An indefiniteness restriction for relative clauses in Lakhota. In Eric Reuland & Alice ter Meulen (eds.), *The representation of (in)definiteness*, 168–190. Cambridge, MA: MIT Press.

Chapter 12

A description of verb-phrase ellipsis in Hocąk

Meredith Johnson

In this paper, I argue that Hocąk displays verb-phrase ellipsis (VPE) and provide the first thorough description of this phenomenon. VPE in Hocąk displays the two defining characteristics of VPE cross-linguistically: it targets all VP-internal material, and it is licensed by a functional head. In the case of Hocąk, I propose that the licensing head is active *v*. Furthermore, Hocąk VPE also shows many other traits of VPE in other languages: the ellipsis site can be found in coordinated and adjacent clauses in addition to embedded and adjunct clauses, VPE is insensitive to the the contents of the VP, and VPE gives rise to both strict and sloppy readings. Lastly, I argue that VPE in Hocąk is derived by deletion of a full-fledged VP, and that the ellipsis site cannot be analyzed as a null *pro* form.

1 Introduction

This purpose of this paper is to both argue that Hocąk displays verb-phrase ellipsis (VPE) and to provide the first thorough description of this phenomenon. In VPE constructions, a VP goes unpronounced when there is an appropriate antecedent VP and a licensing head that identifies the gap. Both of these properties can be seen in the examples of VPE in Hocąk in (1) below. In each example, the VP in the second conjunct is interpreted as identical to the VP in the first conjunct, even though the former has no phonological realization. Instead, the light verb *ųų* takes the place of the VP.

(1) a. *Cecilga* [VP *važątirehižą ruwį*] *kjane anąga nee šge* [*ha'ųų*]
Cecil-ga *važątire-hižą* ∅*-ruwį* *kjane anąga nee šge* *ha-ųų*
Cecil-PROP car-INDEF 3S/O-buy FUT and I also 1S-do
kjane.
kjane
FUT
'Cecil will buy a car, and I will too.'

Meredith Johnson. 2016. A description of verb-phrase ellipsis in Hocąk. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 287–313. Berlin: Language Science Press. https : / / doi . org / 10 . 17169 / langsci . b94 . 175
DOI:10.17169/langsci.b94.175

b. *Cecilga* [$_{VP}$ *xjanąre waši*] *anąga Bryanga šge* [*ųų*].
Cecil-ga *xjanąre* *∅-waši* *anąga Bryan-ga* *šge* *∅-ųų*
Cecil-PROP yesterday 3s-dance and Bryan-PROP also 3s-do
'Cecil danced yesterday, and Bryan did too.'

Throughout this paper, I rely on the set of diagnostics of VPE established by Goldberg (2005), and subsequently used for Indonesian by Fortin (2007). Goldberg (2005) uses characteristics of English VPE to establish a typology of VPE crosslinguistically, noting that "English VP Ellipsis has a characteristic set of behavioral traits, the confluence of which is not found in other types of null anaphora" (Goldberg 2005: 27).

Goldberg developed this set of traits in order to diagnose verb-stranding verb-phrase ellipsis (VVPE) in a variety of languages, including Hebrew, Irish and Swahili. In VVPE, the verb has undergone raising before the remainder of the VP is elided. On the surface, VVPE can be ambiguous between a null object analysis or VPE analysis; thus, some of her diagnostics serve to distinguish these two approaches. An example of VVPE in Hebrew is provided in (2):

(2) *Tazmini et Dvora la-misiba? Kvar hizmanti.*
invite.2FUT Dvora to.the-party already invite.1PST
'Will you invite Dvora to the party? I already invited (Dvora to the party).'
(Goldberg 2005: 14)

Hocąk VPE does not face this problem: there is an overt light verb standing in for the VP, much like English VPE. Nonetheless, the data from Hocąk are consistent with all of the characteristics that Goldberg argues are diagnostic of VPE crosslinguistically. Furthermore, I show that these traits also distinguish VPE from other elliptical phenomena found in Hocąk, including gapping, stripping and null complement anaphora.

This paper is structured as follows. In §2, I establish that putative VPE in Hocąk displays the two most important characteristics of VPE: it targets all VP-internal material, and it is licensed by a functional head. In the case of Hocąk, I propose that the licensing head for VPE is active *v*. In §3, I show that Hocąk VPE displays other traits that have been attributed to VPE crosslinguistically. §4 demonstrates that VPE in Hocąk must be analyzed as a deletion process, rather than a null *pro* form. §5 concludes the paper.

2 Establishing the presence of VPE in Hocąk

In this section, I show that the construction that I argue instantiates VPE in Hocąk displays the two defining characteristics of VPE. In §2.1, I demonstrate that the ellipsis site includes all VP-internal material. In §2.2, I show that VPE is subject to the presence of an appropriate licensing head.

2.1 *ųų* targets the VP

VPE is possible with both intransitive and transitive verbs, as seen in (3-4) below. (3a) and (3b) show that *ųų* can target intransitive VPs. In the examples in (4) with transitive verbs, the direct object is also included in the ellipsis site.

(3) a. *Cecilga* [$_{VP}$ *kere*] *anąga Matejaga šge* [*ųų*].
Cecil-ga *∅-kere,* *anąga Mateja-ga* *šge* *∅-ųų*
Cecil-PROP 3s-leave and Mateja-PROP also 3s-do
'Cecil left, and Mateja did too.'

b. *Meredithga* [$_{VP}$ *nįįp*] *anąga Sarahga* *šge* [*ųų*].
Meredith-ga *∅-nįįp* *anąga Sarah-ga* *šge* *∅-ųų*
Meredith-PROP 3s-swim and Sarah-PROP also 3s-do
'Meredith swam, and Sarah did too.'

(4) a. *Matejaga* [$_{VP}$ *waisgap sguuhižą rook'į*] *anąga Sarahga šge*
Mateja-ga *waisgap sguu-hižą ∅-rook'į* *anąga Sarah-ga* *šge*
Mateja-PROP cake-INDEF 3s/O-bake and Sarah-PROP also
[*ųų*].
∅-ųų
3s-do
'Mateja baked a cake, and Sarah did too.'

b. *Meredithga* [$_{VP}$ *waaruchižą hogiha*] *anąga Bryanga šge*
Meredith-ga *waaruc-hižą ∅-hogiha* *anąga Bryan-ga* *šge*
Meredith-PROP table-INDEF 3s/O-paint and Bryan-PROP also
[*ųų*].
∅-ųų
3s-do
'Meredith painted a table, and Bryan did too.'

VPE can also target other internal arguments. Both indirect objects and resultative phrases are typically analyzed as VP-internal (see e.g., Larson 1988 and

Levin & Rappaport Hovav 1995), and they are also subject to VPE. The ditransitive example in (5) shows that both a direct object and indirect object can be contained in the ellipsis site.

(5) *Cecilga* [VP *Meredithga* *wiiwagaxhižą* *hok'ų*] *anąga*
Cecil-ga *Meredith-ga* *wiiwagax-hižą* *∅-hok'ų* *anąga*
Cecil-PROP Meredith-PROP pencil-INDEF 3S/O-give and
Matejaga *šge* [*ųų*].
Mateja-ga *šge* *∅-ųų*
Mateja-PROP also 3S-do
'Cecil gave Meredith a pencil, and Mateja did too.'

In (6), we see examples of VPE with resultative constructions in which the direct object and result have both been elided.

(6) a. *Cecilga* [VP *ważątirehiža* *šuuc hogiha*] *anąga Bryanga* *šge*
Cecil-ga *ważątire-hiža* *šuuc ∅-hogiha* *anąga Bryan-ga* *šge*
Cecil-PROP car-INDEF red 3S/O-paint and Bryan-PROP also
[*ųų*].
∅-ųų
S-do
'Cecil painted a car red, and Bryan did too.'

b. *Meredithga* [VP *mąąshiža* *paras gistak*] *anąga Matejaga*
Meredith-ga *mąąs-hiža* *paras ∅-gistak anąga Mateja-ga*
Meredith-PROP metal-INDEF flat 3S/O-hit and Mateja-PROP
šge [*ųų*].
šge *∅-ųų*
also 3S-do
'Meredith hit metal flat, and Mateja did too.'

VPE also targets various adjuncts. (7) shows that VPE targets VPs containing temporal adjuncts.

(7) a. *Cecilga* [VP *xjanąre* *waši*] *anąga Bryanga* *šge* [*ųų*].
Cecil-ga *xjanąre* *∅-waši* *anąga Bryan-ga* *šge* *∅-ųų*
Cecil-PROP yesterday 3S-dance and Bryan-PROP also 3S-do
'Cecil danced yesterday, and Bryan did too.'

b. *Meredithga* [VP *hąąpte'e kšeehižą* *ruuc*] *anąga*
Meredith-ga *hąąpte'e kšee-hižą* *∅-ruuc anąga*
Meredith-PROP today apple-INDEF 3S/O-eat and Mateja-PROP

Matejaga šge [ųų].
Mateja-ga šge ∅-ųų
also 3s-do
'Meredith ate an apple today, and Mateja did too.'

In (8), locative adjuncts are included in the ellipsis site.

(8) a. *Cecilga [$_{VP}$ hosto eja waši] kjane anąga Bryanga*
Cecil-ga hosto eja ∅-waši kjane anąga Bryan-ga
Cecil-PROP gathering there 3s-dance FUT and Bryan-PROP
šge [ųų] kjane.
šge ∅-ųų kjane
also 3s-do FUT
'Cecil will dance at the gathering, and Bryan will too.'

b. *Cecilga [$_{VP}$ ciinąk eja wažątirehižą ruwį] anąga*
Cecil-ga ciinąk eja wažątire-hižą ∅-ruwį anąga
Cecil-PROP city there car-INDEF 3s/O-buy and
Bryanga šge [ųų].
Bryan-ga šge ∅-ųų
Bryan-PROP also 3s-do
'Cecil bought a car in the city, and Bryan did too.'

(9) exemplifies VPE with a comitative.

(9) *Cecilga [$_{VP}$ hinųkra hakižu waši] anąga Bryanga šge*
Cecil-ga hinųk-ra hakižu ∅-waši anąga Bryan-ga šge
Cecil-PROP woman-DEF be.with 3s-dance and Bryan-PROP also
[ųų].
∅-ųų
3s-do
'Cecil danced with the woman, and Bryan did too.'

(10) demonstrates that various manner adverbs can also be subject to VPE.

(10) a. *Bryanga [$_{VP}$ teejąki nįįtašjak taaxu racgą] anąga Sarahga*
Bryan-ga teejąki nįįtašjak taaxu ∅-racgą anąga Sarah-ga
Bryan-PROP often coffee 3s-drink and Sarah-PROP
šge [ųų].
šge ∅-ųų
also 3s-do
'Bryan often drinks coffee, and Sarah does too.'

b. *Cecilga* [$_{VP}$ *pįįhį mąąnį*] *anąga Bryanga šge* [*ųų*].
Cecil-ga pįįhį Ø-mąąnį anąga Bryan-ga šge Ø-ųų
Cecil-PROP quietly 3s-walk and Bryan-PROP also 3s-do
'Cecil walked carefully/quietly, and Bryan did too.'

c. *Meredithga* [$_{VP}$ *hikųhe nįįtašjak taaxu racgą*] *anąga Bryanga*
Meredith-ga hikųhe nįįtašjak taaxu Ø-racgą anąga Bryan-ga
Meredith-PROP quickly coffee 3s-drink and Bryan
šge [*ųų*].
šge Ø-ųų
also 3s-do
'Meredith drank coffee quickly, and Bryan did too.'

In all of the examples in (7)–(10), the adjunct in the antecedent VP is interpreted as being present in the ellipsis site, indicating that *ųų* targets the entire VP rather than just the object(s).

Lastly, complement clauses can also be included in VPE. The example in (11) has two possible interpretations: either that Meredith also bought a car, or that Meredith also said that Cecil bought a car. Under the second reading, VPE targets the matrix clause, eliding the verb and its complement clause.

(11) *Bryanga* [*vp Cecilga wažątirehižą ruwįže ee*] *anąga*
Bryan-ga Cecil-ga wažątire-hižą Ø-ruwį-že Ø-ee anąga
Bryan-PROP Cecil-PROP car-INDEF 3S/O-buy-COMP 3S-say and
Meredithga šge [*ųų*].
Meredith-ga šge Ø-ųų
Meredith-PROP also 3S-do
'Bryan said that Cecil bought a car, and Meredith did too.'

2.2 Licensing of VPE

The main characteristic that distinguishes VPE from other elliptical processes is the presence of an overt licensing head located in the inflectional domain above the VP. VPE in English can be licensed by a variety of functional elements, such as *do* in (12a), *be* in (12b), *have* in (12c), *can* in (12d) and *will* in (12e). The obligatory presence of an inflectional head has led previous researchers to argue that VPE is licensed by T/Infl (Bresnan 1976; Sag 1976; Zagona 1988; Lobeck 1995).

(12) a. Lily wore a skirt, and Molly **did** too.
b. Lily is reading a book, and Molly **is** too.

c. Lily hasn't finished the book, but Molly **has**.
d. Lily can ride a bike, and Molly **can** too.
e. Lily will leave, and Molly **will** too.

In contrast, there is no such inflectional head found with stripping or gapping. Stripping is an elliptical phenomenon in which an entire clause is elided except for a single element that is stranded. This is illustrated in (13a). In gapping constructions, the verb (and other potential material) is left unpronounced, while there are two elements that are stranded. An example of gapping can be seen in (13b).

(13) a. Lily came over, and Molly too.
b. Lily brought bagels, and Molly danishes.

In Hocąk, the licensing requirement on VPE is different: VPE is conditioned solely by the presence of the light verb *ųų*. We have seen this in all of the instances of VPE given above. The examples in (14)–(18) illustrate that *ųų* is indeed a light verb: it productively combines with both nouns and verbs to create complex predicates. Based on its distribution, I assume that *ųų* realizes the functional head *v* (Hartmann 2012: Examples 14–18).

(14) a. *mąąnąąpeja*
'warrior'
b. *mąąnąąpeja ųų*
'be in the military'

(15) a. *nąąwąǧoǧo*
'fiddle'
b. *nąąwąǧoǧo ųų*
'play the fiddle'

(16) a. *waruc*
'food'
b. *waruc ųų*
'cook, prepare food'

(17) a. *waagax*
'paper, letter'
b. *waagax ųų*
'write (a letter)'

(18) a. *hooxiwi*
'cough' (verb)

b. *hooxiwi ųų*
'have a cold'

Tense and modals can be present in VPE constructions; however, they are never obligatory. When present, tense and modals always co-occur with the light verb *ųų*. (19a) shows that the future tense marker *kjane* can follow *ųų*, while (19b) and (19c) demonstrate that the modals *ną* and *s'aare* can also appear after *ųų*. When *ųų* is omitted, the result is ungrammatical.

(19) a. *Cecilga wažątirehižą ruwį kjane anąga nee šge *(ha'ųų)*
Cecil-ga wažątire-hižą ∅-ruwį kjane anąga nee šge ha-ųų
Cecil-PROP car-INDEF 3S/O-buy FUT and I also 1S-do
kjane.
kjane
FUT
'Cecil will buy a car, and I will too.'

b. *Meredithga hąąke wažątirera pįį'ų ruxuruknį nųnįge*
Meredith-ga hąąke wažątire-ra ∅-pįį'ų ruxuruk-nį nųnįge
Meredith-PROP NEG car-DEF 3S/O-fix-NEG but
*Matejaga *(ųų)* ***ną***.
Mateja-ga ∅-ųų ***ną***
Mateja-PROP 3S-do can
'Meredith can't fix the car, but Mateja can.'

c. *Meredithga hąąke nįįtašjak taaxu ruwįnį nųnįge*
Meredith-ga hąąke nįįtašjak taaxu ∅-ruwį-nį nųnįge
Meredith-PROP NEG coffee 3S/O-buy-NEG but
*Matejaga *(ųų)* ***s'aare***.
Mateja-ga ∅-ųų ***s'aare***
Mateja-PROP 3S-do must
'Meredith didn't buy coffee but Mateja must have.'

Thus, we see that T/Infl does not play the same role in VPE licensing in Hocąk as it does in other languages. However, VPE in Hocąk is constrained by the type of predicate. As the examples in (20) show, VPE is not licensed with non-agentive verbs:

(20) a. * *Meredithga kšee gipį anąga Bryanga šge ųų.*
Meredith-ga kšee ∅-gipį anąga Bryan-ga šge ∅-ųų
Meredith-PROP apple 3S-like and Bryan-PROP also 3S-do
(Intended: 'Meredith likes apples, and Bryan does too.')

b. * *Cecilga wįįxra waaja anąga Meredithga šge ųų.*
Cecil-ga wįįx-ra wa-∅-haja anąga Meredith-ga šge ∅-ųų
Cecil-PROP duck-DEF 3O.PL-3S-see and Meredith-PROP also 3S-do
(Intended: 'Cecil saw the ducks, and Meredith did too.')

c. * *Meredithga hoišą anąga Bryanga šge ųų.*
Meredith-ga ∅-hoišą anąga Bryan-ga šge ∅-ųų
Meredith-PROP 3S-busy.STAT and Bryan-PROP also 3S-do
(Intended: 'Meredith is busy, and Bryan is too.')

d. * *Cecilga hįįcge nųnįge Bryanga hąąke ųųnį.*
Cecil-ga ∅-hįįcge nųnįge Bryan-ga hąąke ∅-ųų-nį
Cecil-PROP 3S-tired.STAT but Bryan-PROP NEG 3S-do-NEG
(Intended: 'Cecil is tired, but Bryan isn't.')

Like other Siouan languages, Hocąk exhibits an active-stative alignment pattern: the active set of verbal person markers is used to index the subject of transitive verbs and active intransitive verbs, while the stative set is used to index the object of transitive verbs and the subject of stative intransitive verbs. This alignment pattern interacts with VPE in revealing ways. While VPE is banned with most stative intransitive verbs, such as those in (20c) and (20d), VPE is possible with certain stative intransitives when they have an agentive reading. *Hokąre* 'to fall in' is normally a stative intransitive verb, but it is possible to use it in VPE contexts if the subject falls *deliberately*, as in (21). In this context, *ųų* takes the *active* person marker set. In (21b), the verb takes the second person active marker *š-*; the stative marker *nį-* is not permitted. The marker *nį-* is the one that would typically be found on the verb *hokąre*, as shown in (22).

(21) a. *Meredithga nįį eeja hokąre anąga Bryanga šge ųų.*
Meredith-ga nįį eeja ∅-hokąre anąga Bryan-ga šge ∅-ųų
Meredith-PROP water there 3S-fall.in and Bryan-PROP also 3S-do
'Meredith fell into the water (deliberately), and Bryan did too.'

b. *Meredithga nįį eeja hokąre anąga (nee) šge š'ųų/*nį'ųų.*
Meredith-ga nįį eeja ∅-hokąre anąga nee šge š-'ųų/nį-'ųų
Meredith-PROP water there 3S-fall.in and you also 2S-do
'Meredith fell into the water (deliberately), and you did too.'

(22) *Honįkąre.*
<nį>hokąre
<2S>fall.in.STAT
'You fell in(to something).' (Hartmann 2012)

This restriction on VPE is not due to lexical properties of *ųų*: when *ųų* functions as a light verb, it can form non-agentive verbs, as in (23):

(23) a. *hooxiwi ųų* 'have a cold' (stative intransitive)

b. *roo taakac ųų* 'have a fever' (stative intransitive)

c. *paaxšišik ųų* 'have an upset stomach' (stative intransitive) (Hartmann 2012)

To formalize this restriction on VPE in Hocąk, I adopt Merchant's (2001) proposal that ellipsis takes place when a so-called "[E]-feature" is present on the relevant licensing head. In the case of Hocąk, I propose that an [E]-feature is present only on the agentive *v* head.[1] This accounts for the fact that VPE is solely conditioned by the presence of the light verb (or *v*) *ųų*, and furthermore that only agentive verbs can be elided: if there is a non-agentive *v* present, then ellipsis will not be licensed.

This conclusion is in line with other research that argues that *v* is responsible for licensing VPE crosslinguistically. Many recent approaches to ellipsis have argued for a link between phases and elliptical phenomena (Holmberg 2001, van Craenenbroeck 2004, Gengel 2007, Yoshida & Gallego 2008, Gallego 2009, among others). Specifically, they propose that ellipsis results when a phasal head (e.g. *v*, C, D) licenses deletion of its complement. These theories are a natural development of Chomsky's (2000; 2001; 2004) theory of phases: if ellipsis is PF-deletion, it follows that the units that are sent cyclically to the PF interface are precisely the ones that can be targeted for deletion. More concretely, Rouveret (2012) adopts the phasal analysis of ellipsis, and puts forward a theory to predict which languages permit VPE. He argues that *v* always has an uninterpretable [tense] feature, and that, in languages with VPE, the [tense] feature is valued on *v* phase-internally. Rouveret proposes that the elements that license VPE are all merged in *v*, and subsequently move to Infl. All of these approaches are compat-

[1] This agentivity requirement on a process that affects the VP is not completely unique to Hocąk. For example, Hallman (2004) notes that English *do so* replacement is restricted to agentive VPs, even though other uses of *do* are not subject to this constraint (e.g., *Max loves studying French, and Mary does (*so) too.*) Rouveret (2012) also shows that VPE in Welsh is licensed uniquely by the light verb *gweund*, and furthermore that VPE is not permitted with stative predicates. The only possibility with stative VPs is VVPE. However, Rouveret also shows that *gweund* is also incompatible with stative predicates in its non-elliptical uses. This contrasts with the behavior of *ųų* in Hocąk and *do* in *do so* in English.

ible with the Hocąk data, with the caveat that VPE is more restricted in Hocąk: it is only licensed by active *v*.

3 Crosslinguistic characteristics of VPE

In the previous section, I demonstrated that Hocąk displays the two defining characteristics of VPE: the elliptical process in question targets the entire VP, and is conditioned by the presence of a licensing head. Goldberg (2005) discusses five other characteristics of VPE that are not shared by other elliptical phenomena, which are listed in (24):

(24) a. Possible in both coordinated and adjacent CPs
b. Insensitive to contents of elided VP
c. Ellipsis site can be in a syntactic island
d. Ellipsis site can be embedded
e. Presence of strict and sloppy readings

In the subsections that follow, I show that Hocąk VPE also generally conforms to this typology. In the areas where Hocąk appears to differ from English, I demonstrate that this is due to other differences between the two languages that are independent of ellipsis.

3.1 Ellipsis licensed in both coordinated and adjacent CPs

Goldberg (2005) notes that English VPE is possible with a variety of sentence types. VPE is licit when the antecedent VP and elided VP are found in conjoined CPs (25a), in adjacent CPs uttered by the same speaker (25b), and when the antecedent is in a question and the ellipsis site in the answer (25c). In this section, I show that the same is true in Hocąk.

(25) a. Lily hates beets, but Molly doesn't.
b. Lily hates beets. Molly does too.
c. Who hates beets? Molly does.

All of the examples of VPE we saw in §2 involved two clauses joined by the coordinator *anąga* 'and'. VPE is also possible with disjunction, as seen in (26) with *nųnįge* 'but'.

(26) a. *Cecilga wažątirehižą ruwį nųnįge nee hąąke ha'ųųnį.*
Cecil-ga wažątire-hižą ∅-ruwį nųnįge nee hąąke ha-ųų-nį
Cecil-PROP car-INDEF 3S/O-buy but I NEG 1S-do-NEG
'Cecil bought a car, but I didn't.'

b. *Sarahga hąąke haas gihinį nųnįge Matejaga ųų.*
Sarah-ga hąąke haas ∅-gihi-nį nųnįge Mateja-ga ∅-ųų
Sarah-PROP NEG berry 3S-pick-NEG but Mateja-PROP 3S-do
'Sarah didn't pick berries, but Mateja did.'

(27) shows that VPE is also licit in adjacent CPs. In each example, the antecedent VP is found in the first sentence while the ellipsis site is in the second sentence.

(27) a. *Meredithga waaruchižą hogiha. Bryanga šge ųų.*
Meredith-ga waaruc-hižą ∅-hogiha Bryan-ga šge ∅-ųų
Meredith-PROP table-INDEF 3S/O-paint Bryan-PROP also 3S-do
'Meredith painted a table. Bryan did too.'

b. *Meredithga hąąke waisgap sguu xuwuxuwuhižą ruucnį.*
Meredith-ga hąąke waisgap sguu xuwuxuwu-hižą ∅-ruuc-nį
Meredith-PROP NEG cookie-INDEF 3S/O-eat-NEG
Bryanga ųų.
Bryan-ga ∅-ųų
Bryan-PROP 3S-do
'Meredith didn't eat a cookie. Bryan did.'

Lastly, VPE also occurs in question-answer pairs in Hocąk. In (28a), a yes-no question contains the antecedent VP and the answer contains the gap. (28b) demonstrate that the same holds of *wh*-questions.

(28) a. Question: *Nįįtašjak taaxu šuruwį?* Answer: *Ha'ųų.*
nįįtašjak taaxu šu-ruwį ha-ųų
coffee 2S-buy 1S-do
Q: 'Did you buy coffee?' A: 'I did.'

b. Question: *Peežega Cecilga gišja hii?* Answer: *Bryanga ųų.*
peežega Cecil-ga ∅-gišja hii Bryan-ga ∅-ųų.
who Cecil-PROP 3S/O-visit Bryan-PROP 3S-do
Q: 'Who visited Cecil?' A: 'Bryan did.'

3.2 Ellipsis and the contents of the VP

Goldberg (2005) distinguishes VPE from null complement anaphora (NCA) based on the type of constituent that is elided. In NCA, a matrix verb is stranded and its complement is elided. However, NCA is constrained by the contents of the VP: only propositions can be elided. This is illustrated by the contrast between the grammatical NCA examples in (29a) and (29c) and the ungrammatical examples in (29b) and (29d):

(29) a. Pat doesn't know **that Terry is moving to Japan**, but Robin knows.
b. * Pat doesn't know **how to speak Inuktitut**, but Robin knows.
c. Pat forgot **to close the door**, but Robin remembered.
d. * Pat forgot **the answer**, but Robin remembered. (Fortin 2007: 245)

In contrast, the grammaticality of VPE is not dependent on the contents of the VP. The examples in (30) show that VPE is possible regardless of whether the complement of the VP expresses a proposition or not.

(30) a. Pat doesn't know **that Terry is moving to Japan**, but Robin does.
b. Pat doesn't know **how to speak Inuktitut**, but Robin does.
c. Pat forgot **to close the door**, but Robin didn't.
d. Pat forgot **the answer**, but Robin didn't.

As Fortin (2007) points out, this diagnostic does not serve to distinguish VPE from NCA in languages with null objects. Hocąk allows both null subjects and objects, as seen in (31b):

(31) a. *Wijųkra šųųkra hoxataprookeeja haja.*
Wijųk-ra šųųk-ra hoxatap-rook-eeja ∅-haja
cat-DEF dog-DEF woods-inside-there 3S/O-see
'The cat saw the dog in the woods.
b. *Hoxataprookeeja haja.*
hoxatap-rook-eeja ∅-haja
woods-inside-there 3S/O-see
'[The cat] saw [the dog] in the woods.' (Johnson, Rosen & Schuck 2013: 7)

Thus, it is not surprising that both propositional and non-propositional verbal complements can be null in Hocąk. In (32), the complement of the verb *hiperes* 'know' can be null both when it is a proposition (32a) or an embedded question (32b). Likewise, both propositional (33a) and DP object (33b) complements of *wakikųnųnį* 'forget' surface as null.

(32) a. *Sarahga Meredithga rookhožura ruucra hiperes,*
Sarah-ga Meredith-ga rookhožu-ra ∅-ruuc-ra ∅-hiperes
Sarah-PROP Meredith-PROP pie-DEF 3S/O-eat-COMP 3S-know
anąga Matejaga šge hireperešąną.
anąga Mateja-ga šge ∅-hiperes-šąną
and Mateja-PROP also 3S-know-DECL
'Sarah knows that Meredith ate the pie, and Mateja knows too.'

b. *Sarahga jaagu'ų Meredithga kerera hiperes, anąga*
Sarah-ga jaagu'ų Meredith-ga ∅-kere-ra ∅-hiperes anąga
Sarah-PROP why Meredith-PROP 3S-leave-COMP 3S-know and
Matejaga šge hiperešąną.
Mateja-ga šge ∅-hiperes-šąną
Mateja-PROP also 3S-know-DECL
'Sarah knows why Meredith left, and Mateja knows (why Meredith left) too.'

(33) a. *Bryanga nįįtašjak taaxu ruwįra wakikųnųnį, nųnįge*
Bryan-ga nįįtašjak taaxu ∅-ruwį-ra ∅-wakikųnųnį nųnįge
Bryan-PROP coffee 3S-buy-COMP 3S-forget but
Meredithga hąąke wakikųnųnįnį.
Meredith-ga hąąke ∅-wakikųnųnį-nį
Meredith-PROP NEG 3S-forget-NEG
'Bryan forgot to buy coffee, but Meredith didn't forget.'

b. *Bryanga waisgap sguura wakikųnųnį, nųnįge Meredithga*
Bryan-ga waisgap sguu-ra ∅-wakikųnųnį nųnįge Meredith-ga
Bryan-PROP cake-DEF 3S/O-forget but Meredith-PROP
hąąke wakikųnųnįnį.
hąąke ∅-wakikųnųnį-nį
NEG 3S/O-forget-NEG
'Bryan forgot the cake, but Meredith didn't forget (the cake).'

Thus, this particular diagnostic does not work for Hocąk due to independent factors. The complement of verbs like 'know' and 'forget' can always be null, presumably due to the availability of object *pro* drop.[2]

3.3 Ellipsis in syntactic islands

Goldberg (2005) notes that the ellipsis site in VPE constructions can be inside an adjunct island, while gapping is not permitted in adjuncts. This is shown by the contrast between (34a) and (34b) below:

(34) a. Lily finished her sandwich before Molly did.
b. * Lily finished the sandwich before Molly the pizza.

The same contrast is found in Hocąk. The examples in (35) show that the gap in VPE constructions can be found inside adjunct clauses (which precede the main clause in these examples). In (35a), the ellipsis site is in the clause headed by 'if', in (35b) the ellipsis site is in the clause headed by 'because', and in (35c) it is in the clause headed by 'before.'

(35) a. *Bryanga ųų kjanegi Meredithga Hunterga (nišge)*
Bryan-ga ∅-ųų kjane-gi Meredith-ga Hunter-ga (nišge)
Bryan-PROP 3S-do FUT-if Meredith-PROP Hunter-PROP also
gišja hii kjane.
∅-gišja hii kjane
3S/O-visit FUT
'Meredith will visit Hunter if Bryan will.'

b. *Bryanga hąąke ųųnįge Meredithga (nišge) hąąke*
Bryan-ga hąąke ∅-ųų-nį-ge Meredith-ga (nišge) hąąke
Bryan-PROP NEG 3S-do-NEG-because Meredith-PROP also NEG
Hunterga gišja hiinį.
Hunter-ga ∅-gišja hii-nį
Hunter-PROP 3S/O-visit-NEG
'Meredith didn't visit Hunter because Bryan didn't.'

[2] A full comparison of NCA and VPE is not possible in Hocąk. VPE with verbs like 'know' and 'forget' is ungrammatical (examples omitted for space purposes) since these verbs are non-agentive.

c. *Keenį Sarahga ųųnį Matejaga*
keenį Sarah-ga ∅-ųų-nį Mateja-ga
before Sarah-PROP 3S-do-NEG Mateja-PROP
waisgap sguu xuwuxuwuhižą ruucšąną.
waisgap sguu xuwuxuwu-hižą ∅-ruuc-šąną
cookie-INDEF 3S/O-eat-DECL
'Mateja ate a cookie before Sarah did.'

In contrast, gapping is ungrammatical in adjuncts. (36) illustrates that the the gap cannot be located in an adjunct clause headed by 'if' (36a), 'because' (36b) or 'before' (36c).

(36) a. * *Matejaga rookhožuhižągi Meredithga waisgap sguuhižą*
Mateja-ga rookhožu-hižą-gi Meredith-ga waisgap sguu-hižą
Mateja-PROP pie-INDEF-if Meredith-PROP cake-INDEF
rook'į kjane.
∅-rook'į kjane
3S/O-bake FUT
(Intended: 'Meredith will bake a cake if Mateja will bake a pie.')

b. * *Sarahga važą honąkipįnįhižąge Matejaga važątirehižą*
Sarah-ga važą honąkipįnį-hižą-ge Mateja-ga važątire-hižą
Sarah-PROP bicycle-INDEF-because Mateja-PROP car-INDEF
ruwį.
∅-ruwį
3S/O-buy
(Intended: 'Mateja bought a car because Sarah bought a bicycle.')

c. * *Keenį Bryanga waisgap sguu xuwuxuwuhižąnį Meredithga*
keenį Bryan-ga waisgap sguu xuwuxuwu-hižą-nį Meredith-ga
before Bryan-PROP cookie-INDEF-NEG Meredith-PROP
kšeehižą ruucšąną.
kšee-hižą ∅-ruuc-šąną
apple-INDEF 3S/O-eat-DECL
(Intended: 'Meredith ate an apple before Bryan ate a cookie.')

3.4 Ellipsis in embedded clauses

Goldberg (2005) also shows that the ellipsis site in VPE constructions can be inside an embedded clause, while this is not true of other types of ellipsis. (37a)

demonstrates that VPE is licit in an embedded clause, while (37b)–(37c) illustrate that neither gapping nor stripping are possible in an embedded clause.

(37) a. Lily went to the zoo, and I think (that) Molly did too.
b. * Lily went to the zoo, and I think (that) Molly the aquarium.
c. * Lily went to the zoo, and I think (that) Molly too.

In Hocąk, VPE is licit in the complement clause of various matrix verbs, including 'know' (38a), 'want' (38b), 'think' (38c) and 'say' (38d).

(38) a. *Bryanga hąąke nįįtašjak taaxu ruwįnį, nųnįge Meredithga*
Bryan-ga hąąke nįįtašjak taaxu ∅-ruwį-nį nųnįge Meredith-ga
Bryan-PROP NEG coffee 3S-buy-NEG but Meredith-PROP
ųųra yaaperesšąną.
∅-ųų-ra <ha>hiperes-šąną
3S-do-COMP <1S>know-DECL
'Bryan didn't buy coffee, but I know Meredith did.'

b. *Meredithga hąąke Hunterga gišja hiinį nųnįge*
Meredith-ga hąąke Hunter-ga ∅-gišja hii-nį nųnįge
Meredith-PROP NEG Hunter-PROP 3S/O-visit-NEG but
Bryanga ųų roogų.
Bryan-ga ∅-ųų ∅-roogų
Bryan-PROP 3S-do 3S-want
'Meredith didn't visit Hunter, but Bryan wants to.'

c. *Matejaga hąąke važą honąkipįnįhižą ruwįnį, nųnįge*
Mateja-ga hąąke važą honąkipįnį-hižą ∅-ruwį-nį nųnįge
Mateja-PROP NEG bicycle-INDEF 3S/O-buy-NEG but
Cecilga ųųže yaare.
Cecil-ga ∅-ųų-že <ha>hire
Cecil-PROP 3S-do-COMP <1S>think
'Mateja didn't buy a bicycle, but I think Cecil did.'

d. *Sarahga hąąke waarucra hogihanį, nųnįge Meredithga*
Sarah-ga hąąke waaruc-ra ∅-hogiha-nį nųnįge Meredith-ga
Sarah-PROP NEG table-DEF 3S/O-paint-NEG but Meredith-PROP
ųųže ee.
∅-ųų-že ∅-ee
3S-do-COMP 3S-say
'Sarah didn't paint the table, but Meredith said she did.'

Unlike English, Hocąk does not exhibit any constraint on gapping in embedded contexts. (39a) and (39b) show that the gap can be embedded under the verbs *hire* 'think' and *ee* 'say', respectively.

(39) a. *Meredithga wažą honąkipįnįhižą ruwį anąga Bryanga*
Meredith-ga wažą honąkipįnį-hižą ∅-ruwį anąga Bryan-ga
Meredith-PROP bicycle-INDEF 3S/O-buy and Bryan-PROP
wažątirehižą yaare.
wažątire-hižą <ha>hire
car-INDEF <1S>think
'Meredith bought a bicycle, and I think that Bryan bought a car.'

b. *Meredithga kšeehižą ruuc anąga Matejaga*
Meredith-ga kšee-hižą ∅-ruuc anąga Mateja-ga
Meredith-PROP apple-INDEF 3S/O-eat and Mateja-PROP
wažązihižą hihe.
wažązi-hižą <ha>ee
orange-INDEF <1S>say
'Meredith ate an apple and I said that Mateja ate an orange.'

The examples in (40) show that Hocąk exhibits stripping. (40a) illustrates stripping with an object remnant after the coordinator 'and', while the example in (40b) has an object remnant with disjunction. (40c) shows that stripping is also possible with a subject remnant after the coordinator.

(40) a. *Sarahga šųųkhižą haja, anąga wijukhižą šge.*
Sarah-ga šųųk-hižą ∅-haja anąga wijuk-hižą šge
Sarah-PROP dog-INDEF 3S/O-see and cat-INDEF also
'Sarah saw a dog, and a cat too.'

b. *Meredithga hąąke kšeehižą ruucnį, nųnįge*
Meredith-ga hąąke kšee-hižą ∅-ruuc-nį nųnįge
Meredith-PROP NEG apple-INDEF 3S/O-eat NEG
waisgap sguu xuwuxuwuhižą.
waisgap sguu xuwuxuwu-hižą
but
'Meredith didn't eat an apple, but a cookie.'

c. *Bryanga nįįtašjak taaxu racgą, anąga Matejaga šge.*
Bryan-ga nįįtašjak taaxu ∅-racgą anąga Mateja-ga šge
Bryan-PROP coffee 3S-drink and Mateja-PROP also
'Bryan drank coffee, and Mateja too.'

As is the case in English and other languages, stripping is ungrammatical in embedded clauses in Hocąk. This is shown in (41a) for an object remnant with conjunction, (41b) for an object remnant with disjunction and (41c) for a subject remnant with conjunction.

(41) a. * *Matejaga wažahe gipį, anąga kšeexete šge yaare.*
Mateja-ga wažahe ∅-gipį anąga kšeexete šge <ha>hire
Mateja-PROP banana 3S-like and pineapple also <1S>think
(Intended: 'Mateja likes bananas, and I think (she likes) pineapple too.')

b. * *Bryanga hąąke wažątirehižą ruwįnį, nųnįge Cecilga*
Bryan-ga hąąke wažątire-hižą ∅-ruwį-nį nųnįge Cecil-ga
Bryan-PROP NEG car-INDEF 3S/O-buy-NEG but Cecil-PROP
wažą honąkipįnįhižą ee.
wažą honąkipįnį-hižą ∅-ee
bicycle-INDEF 3S-say
(Intended: 'Bryan didn't buy a car, but Cecil said (he bought) a bicycle.')

c. * *Sarahga waisgap sguuhižą rook'į, anąga Bryanga*
Sarah-ga waisgap sguu-hižą ∅-rook'į anąga Bryan-ga
Sarah-PROP cake-INDEF 3S/O-bake and Bryan-PROP
Meredithga šge ee.
Meredith-ga šge ∅-ee
Meredith-PROP also 3S-say
(Intended: 'Sarah baked a cake, and Bryan said Meredith (baked a cake) too.')

To conclude, the possibilities of having an ellipsis site in embedded contexts differ between English and Hocąk: VPE and gapping are not differentiated by embedding, but VPE and stripping are. However, gapping and VPE are still distinguished in adjunct clauses: as we saw in 3.3, VPE is grammatical in adjunct clauses (35) while gapping is not (36).

3.5 Presence of strict and sloppy readings

Another characteristic of VPE is the fact that elided pronouns and anaphors give rise to two different identity readings. The English example in (42) is ambiguous. Under the so-called "strict" reading, the referent of the pronoun is identical in both the antecedent and elided VP. Under the "sloppy" reading, the pronoun

behaves like a variable, and the referent of the anaphor is different for each conjunct.

(42) Lily saw herself in the mirror, and Molly did too.
Strict reading: Molly saw Lily in the mirror.
Sloppy reading: Molly saw herself in the mirror.

Fortin (2007) shows that stripping also gives rise to both strict and sloppy readings, as in the example in (43). However, there is another possible interpretation for the second conjunct: the remnant can be interpreted as the object of the stripped clause. Fortin terms this additional reading the "object reading." This third reading is unique to stripping constructions, as the remnant DP in VPE is always interpreted as the subject of the elided constituent.

(43) Lily saw herself in the mirror, and Molly too.
Strict reading: Molly saw Lily in the mirror.
Sloppy reading: Molly saw herself in the mirror.
Object reading: Lily saw Molly in the mirror.

In Hocąk, strict and sloppy readings are available with both VPE and stripping, while the additional "object reading" is possible only with stripping. In the examples in (44), the antecedent VP contains a possessed object. (44a) is an instance of VPE, and the second conjunct has two possible interpretations: either Hunter visited Bryan's mother (strict reading) or Hunter visited his own mother (sloppy reading). In (44b), the second conjunct contains a stripping ellipsis site. Both the strict and sloppy readings are available, but the object reading is also possible: the sentence could mean that Bryan visited Hunter.

(44) a. *Bryanga hi'ųnį hiira homąkįnį anąga Hunterga šge*
Bryan-ga hi'ųnį ∅-hii-ra ∅-homąkįnį anąga Hunter-ga šge
Bryan-PROP mother 3S-POSS-DEF 3S/O-visit and Hunter-PROP also
ųų.
∅-ųų
3s-do
'Bryan visited his mother, and Hunter did too.'

b. *Bryanga hi'ųnį hiira homąkįnį anąga Hunterga šge.*
Bryan-ga hi'ųnį ∅-hii-ra ∅-homąkįnį anąga Hunter-ga šge
Bryan-PROP mother 3S-POSS-DEF 3S/O-visit and Hunter-PROP also
'Bryan visited his mother, and Hunter too.'

The examples in (45) show that the same readings are possible with reflexives. The second conjunct of (45a) contains a VPE gap, and it has two interpretations: either Meredith hit Mateja (sloppy) or Meredith hit herself (strict). In the stripping example in (45b), both strict and sloppy readings are possible, but so is the "object reading" under which Mateja hit Meredith.

(45) a. *Matejaga hokijį anąga Meredithga šge ųų.*
Mateja-ga ∅<kii>hojį anąga Meredith-ga šge ∅-ųų
Mateja-PROP 3S<REFL>hit and Meredith-PROP also 3S-do
'Mateja hit herself, and Meredith did too.'

b. *Matejaga hokijį anąga Meredithga šge.*
Mateja-ga ∅<kii>hojį anąga Meredith-ga šge
Mateja-PROP 3S<REFL>hit and Meredith-PROP also
'Mateja hit herself, and Meredith too.'

Thus, while strict and sloppy readings are available with both VPE and stripping, stripping constructions have the additional reading that Fortin (2007) calls the "object reading".

4 Deletion vs. pro-form analysis

In the previous two sections, I presented arguments that Hocąk exhibits VPE. In this section, I further argue that VPE in Hocąk is derived by a deletion process. There are two main approaches to any given elliptical phenomena: the ellipsis site is either a deleted phrase or a null *pro*-form. Here, I extend two arguments in favor of a deletion approach of English VPE to Hocąk. First, I show that extraction from the ellipsis site is possible. Second, I demonstrate that ellipsis sites can contain the antecedent to a pronoun outside of the gap.

Fiengo & May (1994) argue that English VPE is best analyzed as VP deletion. Their argument is based on cases of object extraction from the ellipsis site. In (46a), we see that the object of the second clause has undergone *wh*-movement out of the ellipsis site. (46b) illustrates the phenomenon known as antecedent-contained deletion (ACD). In ACD constructions, the ellipsis site is found inside of a relative clause and is licensed under identity with the matrix VP. The head of the relative clause (here, *everyone*) is the object of the elided VP. In both (46a) and (46b), movement of the object in the elided VP has taken place. This is not expected under a *pro*-form analysis of VPE: a *pro*-form has no internal structure, and thus there should be no object position inside the ellipsis site that the extracted object could have originated in. In contrast, a deletion analysis posits a

full-fledged VP in the ellipsis site which undergoes deletion at a later stage in the derivation. In the examples in (46), the object originated inside the elided VP, and underwent movement before deletion took place.

(46) a. I know which book Max read, and **which book** Oscar didn't.
b. Dulles suspected everyone who Angleton did. (Fiengo & May 1994: 229, 257)

Likewise, Hocąk constructions with *ųų* cannot be analyzed as a *pro*-form, as object extraction is permitted. (47a) shows that focused elements can be extracted from the ellipsis site, and (47b) exemplifies the movement of *wh*-words from the ellipsis site.[3]

(47) a. *Meredithga waagaxra ruwį, nųnįge **wiiwagaxra** hąąke*
Meredith-ga waagax-ra ∅-ruwį, nųnįge wiiwagax-ra hąąke
Meredith-PROP paper-DEF 3S/O-buy but pencil-DEF NEG
ųųnį.
∅-ųų-nį
3S-do-NEG
'Meredith bought the paper, but the pencil, she didn't.'
b. *ɟaagu Bryanga ruwįra yaaperesšąną, nųnįge **jaagu***
ɟaagu Bryan-ga ∅-ruwį-ra <ha>hiperes-šąną nųnįge jaagu
what Bryan-PROP 3S/O-buy-COMP <1S>know-DECL but what
Hunterga ųųra hąąke yaaperesnį.
Hunter-ga ∅-ųų-ra hąąke <ha>hiperes-nį
Hunter-PROP 3S-do-COMP NEG <1S>know-NEG
'I know what Bryan bought, but I don't know what Hunter did.'

As the example in (48) shows, ACD is also grammatical in Hocąk. ACD would not be possible if *ųų* were a *pro*-form, since the head of the relative clause is the object of the elided VP.

(48) *Bryanga ruwį, jaagu Meredithga ųųra.*
Bryan-ga ∅-ruwį jaagu Meredith-ga ∅-ųų-ra
Bryan-PROP 3S/O-buy what Meredith-PROP 3S-do-COMP
'Bryan bought what(ever) Meredith did.'

[3] Like other Siouan languages, Hocąk is a *wh*-in-situ language. However, *wh*-words can undergo focus driven movement.

The second argument in favor of a deletion analysis of VPE in Hocąk comes from so-called "missing antecedents." Hankamer & Sag (1976) demonstrate that the gap in English VPE constructions can contain the antecedent to a pronoun. In the non-elliptical example in (49a), the DP *a camel* in the second conjunct serves as the antecedent for the pronoun *it* in the third conjunct. In (49b), the VP in the second conjunct is elided, resulting in a missing antecedent for the pronoun *it*. Nonetheless, the sentence is still grammatical. It is important to note that the instance of *a camel* in the first conjunct cannot be the antecedent for the pronoun *it*: as (50) shows, DPs under the scope of negation cannot serve as antecedents for pronouns.

(49) a. I've never ridden a camel, but Ivan's ridden a camel$_i$, and he says it$_i$ stank horribly.

b. I've never ridden a camel, but Ivan has, and he says it$_i$ stank horribly. (Hankamer & Sag 1976: 403)

(50) *I've never ridden a camel$_i$, and it$_i$ stank horribly. (Hankamer & Sag 1976: 404)

Hankamer & Sag (1976) argue that the grammaticality of the example in (49b) points to a deletion analysis of VPE. These facts are not readily explained under a *pro*-form analysis: since the ellipsis site would not have internal structure at any point in the derivation, the elided VP in (49b) would never contain the antecedent for the following pronoun.

Examples of VPE with missing antecedents are also grammatical in Hocąk. In (51a), the DP *kšeexetehižą* 'a pineapple' in the second conjunct is the antecedent for the null pronominal subject of the verb *sguu* 'sweet'. In (51b), the VP containing the antecedent is elided, and the resulting sentence is grammatical. Like English, a pronoun cannot find its antecedent in a negated clause (52).

(51) a. *Hąkaga kšeexetehižą haacnį, nųnįge Matejaga*
hąkaga kšeexete-hižą ∅<ha>ruuc-nį nųnįge Mateja-ga
never pineapple-INDEF 3S<1S>eat-NEG but Mateja-PROP
kšeexetehižą ruuc, anąga sguu ee.
kšeexete-hižą ∅-ruuc anąga ∅-sguu ∅-ee
pineapple-INDEF 3S/O-eat and 3S-sweet 3S-say
'I never ate a pineapple, but Mateja ate a pineapple, and she said it was sweet.'

b. *Hąkaga kšeexetehižą haacnį, nųnįge Matejaga ųų,*
hąkaga kšeexete-hižą ∅<ha>ruuc-nį nųnįge Mateja-ga ∅-ųų
never pineapple-INDEF 3s<1s>eat-NEG but Mateja-PROP 3s-do
anąga sguu ee.
anąga ∅-sguu ∅-ee
and 3s-sweet 3s-say
'I never ate a pineapple, but Mateja did, and she said it was sweet.'

(52) *Hąkaga kšeexetehižą haacnį anąga sguu.*
hąkaga kšeexete-hižą ∅<ha>ruuc-nį anąga ∅-sguu
never pineapple-INDEF 3s<1s>eat-NEG and 3s-sweet
'I never ate a pineapple, and it was sweet.'

Both the extraction facts and pronoun antecedent facts point to an analysis in which the contents of elided VPs in Hocąk are present syntactically, and that the omission of elided VPs is due to a deletion process.

5 Conclusion

In this paper, I examined an elliptical phenomenon that I argue instantiates VPE in Hocąk. This process targets all VP-internal material, including direct objects, indirect objects, result phrases, temporal adjuncts, locative adjuncts, comitatives, manner adverbs and complement clauses. VPE is conditioned by the presence of a licensing head, which I showed is the light verb *ųų* in Hocąk. However, Hocąk VPE is constrained in that the antecedent verb must be active. I propose that this restriction is due to the fact that active *v* is the licenser. This elliptical process displays many other traits that Goldberg (2005) and Fortin (2007) demonstrate are characteristic of VPE crosslinguistically. I also briefly discussed that Hocąk VPE should be analyzed as VP deletion, rather than a VP *pro*-form. This paper constitutes the first in depth description of VPE in Hocąk, and contributes to the literature on the properties of VPE crosslinguistically.

Acknowledgment

First and foremost, I would like to thank Cecil Garvin for sharing his language with me and kindly answering my endless questions. Thanks also to Yafei Li, Bryan Rosen and Rand Valentine for discussion and comments at various stages of this project. A portion of this work was presented at the 87th Annual Meeting

of the Linguistic Society of America. I thank the audience there for questions and comments. I would also like to thank an anonymous reviewer and Catherine Rudin for comments on an earlier draft of this paper. Lastly, thanks to Iren Hartmann for generously providing me access to her Lexique Pro Hocąk dictionary to assist with my fieldwork.

Abbreviations

1, 2, 3	first, second, third person	PL	plural
		POSS	possessive
COMP	complementizer	PROP	proper
DECL	declarative		name
DEF	definite	PST	past tense
FUT	future	REFL	reflexive
INDEF	indefinite	S	subject
NEG	negative		agreement
O	object agreement	STAT = stative verb.	

References

Bresnan, Joan. 1976. On the form and functioning of transformations. *Linguistic Inquiry* 7. 3–40.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–51. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures*, vol. 2, 104–131. Oxford: Oxford University Press.

Fiengo, Robert & Robert May. 1994. *Indices and identity*. Cambridge, MA: MIT Press.

Fortin, Catherine Rose. 2007. *Indonesian sluicing and verb phrase ellipsis: Description and explanation in a minimalist framework*. Ann Arbor: University of Michigan (Doctoral dissertation).

Gallego, Ángel J. 2009. *Ellipsis by phase*. Paper presented at the 19th Colloquium on Generative Grammar.

Gengel, Kirsten. 2007. Phases and ellipsis. In Emily Elfner & Martin Walkow (eds.), *Proceedings of the North East Linguistics Society*, vol. 37, 233–246. Amherst, MA: GLSA Publications.

Goldberg, Lotus Madelyn. 2005. *Verb-stranding VP ellipsis: A cross-linguistic study*. Montréal: McGill University (Doctoral dissertation).

Hallman, Peter. 2004. Constituency and agency in VP. In Vineeta Chand, Ann Kelleher, Ángelo J. Rodríguez & Benjamin Schmeiser (eds.), *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, 304–317. Somerville, MA: Cascadilla Press.

Hankamer, Jorge & Ivan A. Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7. 391–426.

Hartmann, Iren. 2012. *Hoocąk Lexique Pro database*.

Holmberg, Anders. 2001. The syntax of 'yes' and 'no' in Finnish. *Studia Linguistica* 55. 140–174.

Johnson, Meredith, Bryan Rosen & Mateja Schuck. 2013. *How configurational is Hocąk?* Paper presented at the 33rd Siouan and Caddoan Languages Conference, Sitting Bull College, Fort Yates, ND.

Larson, Richard K. 1988. On the double-object construction. *Linguistic Inquiry* 19. 335–392.

Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.

Lobeck, Anne. 1995. *Ellipsis: Functional heads, licensing and identification*. Oxford: Oxford University Press.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands and the theory of ellipsis*. Oxford: Oxford University Press.

Rouveret, Alain. 2012. VP ellipsis, phases and the syntax of morphology. *Natural Language and Linguistic Theory* 30. 897–963.

Sag, Ivan A. 1976. A note on verb phrase deletion. *Linguistic Inquiry* 7. 664–671.

van Craenenbroeck, Jeroen. 2004. *Ellipsis in Dutch dialects*. Leiden: Universiteit van Leiden (Doctoral dissertation).

Yoshida, Masaya & Ángel J. Gallego. 2008. *Phases and ellipsis*. Paper presented at the 27th West Coast Conference on Formal Linguistics.

Zagona, Karen. 1988. Proper government of antecedentless VP in English and Spanish. *Natural Language and Linguistic Theory* 6. 95–128.

Chapter 13

On the structure and constituency of Hocąk resultatives

Bryan Rosen

Abstract: This paper explores the structure and constituency of Hocąk (Siouan) "adjectival" resultatives. I argue that Hocąk resultatives project a phrasal XP as the complement of the verb in a Larsonian "VP-shell" (Larson 1988), while the object of the resultative is in Spec,VP. First, I show that the result is an XP and is not a full clause (i.e., a CP). Second, I provide evidence that the result is in a VP-internal position. While the focus of this paper is the structure of resultatives in Hocąk, resultatives as a construction tend to highlight other important characteristics of a language's grammar. I argue that the result predicate is an AP. This puts Hocąk resultatives in line with English adjectival resultatives. The data from resultatives thus suggest that Hocąk has the lexical category *adjective*, contra the previous descriptions of lexical categories in Hocąk (see Lipkind 1945 and Helmbrecht 2006). The goal of this paper is therefore to present new Hocąk data, provide a structural analysis of resultatives, and then explore the adjectival nature of resultative predicates in the language.

1 Introduction

This paper explores the structure and constituency of Hocąk "adjectival" resultatives. In Hocąk resultatives, the result predicate appears to the left of the verb, as exemplified in (1) with *paras* 'flat' and *šuuc* 'red'.[1]

(1) a. *Meredithga mąąsra paras gistakšąną.*
Meredith-ga mąąs-ra paras ∅-gistak-šąną
Meredith-PROP metal-DEF flat 3S/O-hit-DECL

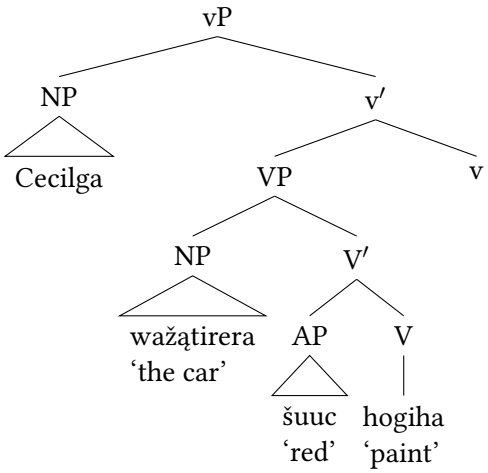
b. *Cecilga ważątirera šuuc hogiha.*
Cecil-ga ważątire-ra šuuc ∅-hogiha
Cecil-PROP car-DEF red 3S/O-paint

[1] Unless noted otherwise, the data comes from elicitation with Cecil Garvin. My methodology follows the standard techniques of translation and acceptability judgment tasks (see Matthewson 2004 for more details).

Bryan Rosen. 2016. On the structure and constituency of Hocąk resultatives. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 313–338. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.176
DOI:10.17169/langsci.b94.176

The analysis of examples like those in (1) is as follows: I propose that Hocąk resultatives project a phrasal AP as the complement of the verb in a Larsonian "VP-shell" (i.e., a recursive VP structure; Larson 1988). The object of the resultative is in the specifier of VP. Thus, the sentence in (1b) has the basic structure in (2).[2]

(2)



While the focus of this paper is to propose a structure of resultatives in Hocąk, resultatives as a construction tend to highlight other important characteristics of a language's grammar. Hocąk resultatives are no exception. I argue that the result predicate is an AP. This puts Hocąk resultatives in line with English adjectival resultatives. The data from resultatives thus suggest that Hocąk has the lexical category adjective, contra the previous descriptions of lexical categories in Hocąk (see Lipkind 1945, Susman 1943, and Helmbrecht 2006). The goal of this paper is therefore to present new Hocąk data, provide a structural analysis of resultatives, and then explore the adjectival nature of resultative predicates in the language. The rest of this paper is organized as follows: §2 provides background on Hocąk syntax and resultatives in Hocąk. §3 examines the constituency of Hocąk resultatives. In §4, I give a syntactic representation of resultatives in

[2] I assume the Principles and Parameters framework (see also the Minimalist Program and X-bar theory; Chomsky 1995). A phrase in this framework consists of three basic layers. The head (X^0) specifies the syntactic type or lexical category of the phrase (e.g., V for verb, N for noun, and A for adjective). Complements are arguments (e.g., objects) of X^0 and are sisters to the X head. Specifiers (Spec for short) are often reserved for subjects of the the phrase. They are sisters to X′.

Hocąk. In §5, I argue that the result predicate projects as an AP. §6 concludes the paper.

2 Overview of Hocąk syntax

In this section, I first present background information on word order in Hocąk, and then I discuss some preliminary characteristics of Hocąk resultatives.

2.1 Word order in Hocąk

Unmarked word order in Hocąk is SOV, as in (3). Variation in word order has discourse effects: a rightward displaced noun phrase is interpreted as discourse-old in (4a), while a leftward moved noun phrase serves a different discourse function (e.g., topic or focus) in (4b). Note that the interpretation in (4b) with OSV word order is possible because there is a pause (represented by the comma) that offsets the fronted object.

(3) *Wijukra šųųkra haja*
wijuk-ra šųųk-ra ∅-haja
cat-DEF dog-DEF 3S/O-see
'The cat saw the dog.'

(4) a. *Wijukra haja, šųųkra*
wijuk-ra ∅-haja šųųk-ra
cat-DEF 3S/O-see dog-DEF
'The cat saw something, the dog.'

b. *Šųųkra, wijukra haja*
šųųk-ra wijuk-ra ∅-haja
dog-DEF cat-DEF 3S/O-see
'The dog, the cat saw (it).'

In double object constructions, the canonical word order is subject–indirect object–direct object–verb. This is shown below in (5).

(5) *Hinųknįkhižą hocįcįhižą wiiwagaxhižą hok'ų.*
hinųknįk-hižą hocįcį-hižą wiiwagax-hižą ∅-hok'ų
girl-INDEF boy-INDEF pencil-INDEF 3S/O-give
'A girl gave a boy a pencil.'

In Hocąk, word order is crucial to disambiguate the subject from the object: the first argument is interpreted as the subject. In (6), the first interpretation of the sentence (although pragmatically unlikely) is the only one with neutral intonation; however, the second interpretation is possible if there is a pause after 'car'.

(6) *Wažątirera hinųkra ruwį.*
wažątire-ra hinųk-ra ∅-ruwį
car-DEF woman-DEF 3S/O-buy
'The car bought the lady.' OR 'The lady bought the car.'

Johnson & Rosen (2014) argue that Hocąk is underlying head-final, by providing evidence from quantifier scope and postverbal predicates. Thus, I represent Hocąk as head-final here.

2.2 Resultatives in Hocąk: Some preliminaries

Resultatives are complex predicates that put together a means predicate (always a verb) and a result predicate, where neither is licensed by a conjunction or an adposition (Williams 2008). In (7), the result *šuuc* 'red' immediately precedes the means *hogiha* 'paint', and the direct object *wažątirera* 'the car' surfaces to the left of the result. Since the result is typically analyzed as the complement of the means (Li 1999, Williams 2008), the result-means order would be expected in a head-final language.

(7) *Cecilga wažątirera šuuc hogiha.*
Cecil-ga wažątire-ra šuuc ∅-hogiha
Cecil-PROP car-DEF red 3S/O-paint
'Cecil painted the car red.'

The word order of resultatives and sentences with object-internal attributive modifiers is similar. Compare the position of the result phrase in (7) with the position of the attributive modifier in (8).

(8) *Cecilga wažątire šuucra hogiha.*
Cecil-ga wažątire šuuc-ra ∅-hogiha
Cecil-PROP car red-DEF 3S/O-paint
'Cecil painted the red car.'

In (8), the modifier *šuuc* 'red' is located to the right of the noun it modifies, *važątire* 'car'. This attributive modifier cannot be to the right of the definite article *-ra*. This entails that *šuuc* 'red' in (8) in an NP-internal position. By comparison, the result in (7) (*šuuc* 'red') is to the right of the definite article *-ra*, which indicates that the result is in an NP-external position.

Moreover, the result AP can "scramble," or move leftward, to a position before the object or subject, as illustrated in (9). In contrast, attributive modifiers do not have this option, as in (10). This contrast demonstrates that resultative predicates are not treated as part of the NP-object, and provides further evidence that they are not in an NP-internal position.

(9) a. *Cecilga* ***šuuc****, važątirera woogiha.*
Cecil-ga šuuc važątire-ra wa-∅-hogiha
Cecil-PROP red car-DEF 3O.PL-3S-paint
'Cecil painted the cars red.'

b. ***Šuuc****, Cecilga važątirera woogiha.*
šuuc Cecil-ga važątire-ra wa-∅-hogiha
red Cecil-PROP car-DEF 3O.PL-3S-paint
'Cecil painted the cars red.'

(10) * *Meredithga* ***šuuc****, wiišgacra ruwį.*
Meredith-ga šuuc wiišgac-ra ∅-ruwį
Meredith-PROP red toy-DEF 3S/O-buy
(Intended: 'Meredith bought the red toy.')

It should be noted that resultative constructions have been categorized cross-linguistically based on whether the lexical semantics of the verb and the result are independent of each other. In his typology of Japanese and English resultative predicates, Washio (1997) presents two types of resultatives: *weak* and *strong*. When the lexical semantics of the verb entails a change, it is called a weak resultative. When the verb in resultative constructions does not entail a change, Washio refers to this class as a strong resultative. In other words, the classification between weak and strong resultatives depends on whether the matrix verb denotes a result. Consider the two English examples in (11).

(11) a. Sam painted the wall red.
b. Alex pounded the metal thin.

In (11a), the verb *paint* entails that there is some change, since *to paint* means to apply color. *Paint* represents an example of a weak resultative. In (11b), however,

the verb *pound* does not entail that the object being pounded will become flat. That is, pounding metal could result in the metal being bumpy. Thus, there is no entailed change with *pound*. The verb *pound* is an example of a strong resultative. In Hocąk, resultatives are possible when the verb lexically specifies a change, as with *hogiha* 'paint' in (7) above and with *gižap* 'polish' in (12) below.

(12) *Meredithga mąąsra gišįnįšįnį gižapšąną.*
Meredith-ga mąąs-ra gišįnįšįnį ∅-gižap-šąną
Meredith-PROP metal-DEF shiny 3S/O-polish-DECL
'Meredith polished the metal shiny.'

A verb like *gižap* 'polish' strongly denotes an activity whereby its object (theme) changes its state to become 'shiny.' Because *gižap* implies this change of state, it is considered a weak resultative.

They can also be formed with verbs that do not specify a change, as with *gistak* 'hit' and *rucgis* 'cut' in (13).

(13) a. *Meredithga mąąsra paras gistakšąną.*
Meredith-ga mąąs-ra paras ∅-gistak-šąną
Meredith-PROP metal-DEF flat 3S/O-hit-DECL
'Meredith hit the metal flat.'

b. *Matejaga peešjįra žiipįk rucgisšąną.*
Mateja-ga peešjį-ra žiipįk ∅-rucgis-šąną
Mateja-PROP hair-DEF short 3S/O-cut-DECL
'Mateja cut the hair short.'

Similar to *pound* in English, *gistak* 'hit' in Hocąk does not denote an event whereby its object results in a particular state (e.g., flat). Thus we can consider this verb a strong resultative. The verb *rucgis* 'cut' belongs to the class of strong resultatives for the same reasons: the event denoted by *rucgis* 'cut' does not contain the notion of being short. Thus, Hocąk exhibits both strong and weak resultatives. With this background in mind, I turn to the next section, where I discuss more about the constituency of Hocąk resultatives.

3 The constituency of Hocąk resultatives

This section outlines some diagnostics that support the structure presented in (2) for Hocąk resultatives. In §3.1, I provide evidence that the result is a phrase and not a clause, while in §3.2 I show that the result is in a VP-internal position.

3.1 The result predicate as a phrase

In this subsection, I show that the result is an XP and is not a full clause (i.e., a CP). First, it should be noted that the result is not a head that forms a compound with the matrix verb; that is, the verb and the result in the construction should not be considered a single lexical unit, such as V^0 or A^0. The result can include adverbial modifiers, such as *hikųhe* 'quickly' in (14a), and the intensifier suffix *-xjį* in (14b).

(14) a. *Meredithga mąąsra paras **hikųhe** gistakšąną.*
Meredith-ga mąąs-ra paras hikųhe ∅-gistak-šąną
Meredith-PROP metal-DEF flat quickly 3S/O-hit-DECL
'Meredith hit the metal flat quickly.'

b. *Meredithga mąąsra parasxjį gistakšąną.*
*Meredith-ga mąąs-ra paras-**xjį** ∅-gistak-šąną*
Meredith-PROP metal-DEF flat-very 3S/O-hit-DECL
'Meredith hit the metal very flat.'

A piece of evidence that the result predicate is not a clause comes from the fact the result phrase cannot take declarative (15a), or complementizer (15b) suffixes.

(15) a. * *Matejaga peešjįra žiipįkšąną rucgisšąną.*
Mateja-ga peešjį-ra žiipįk-šąną ∅-rucgis-šąną
Mateja-PROP hair-DEF short-DECL 3S/O-cut-DECL
(Intended: 'Mateja cut the hair short.')

b. * *Matejaga peešjįra žiipįkra rucgisšąną.*
Mateja-ga peešjį-ra žiipįk-ra ∅-rucgis-šąną
Mateja-PROP hair-DEF short-COMP 3S/O-cut-DECL
(Intended: 'Mateja cut the hair short.')

The result also cannot take the future tense marker *kjane*, as in (16), even though the hair becoming short would necessarily take place after cutting it.

(16) * *Matejaga peešjįra žiipįk ikjane rucgisšąną.*
Mateja-ga peešjį-ra žiipįk kjane ∅-rucgis-šąną
Mateja-PROP hair-DEF short FUT 3S/O-cut-DECL
(Intended: 'Mateja cut the hair short.')

In addition, the result cannot bear the negation suffix *-nį*. Negation in Hocąk is bipartite: the free particle *hąąke* and the suffix *-nį* are both required to form the

negative. The example in (17a) shows that *-nį* attaches to the matrix verb, while (17b) illustrates that the result cannot appear with *-nį*.

(17) a. *Meredithga hąąke mąąsra paras gistaknį.*
*Meredith-ga hąąke mąąs-ra paras ∅-gistak-**nį***
Meredith-PROP NEG metal-DEF flat 3S/O-hit-NEG
'Meredith did not hit the metal flat.'

b. * *Meredithga hąąke mąąsra parasnį gistak.*
*Meredith-ga hąąke mąąs-ra paras-**nį** ∅-gistak*
Meredith-PROP NEG metal-DEF flat-NEG 3S/O-hit
(Intended: 'Meredith did not hit the metal flat.' OR
'Meredith hit the metal such that its surface didn't get fully flat.')

If the result could take one of these suffixes, this would mean that it would have the syntactic status of a clause. Since the examples in (15)–(17) are ungrammatical, the result must not be a clause.

Third, Hocąk resultatives respect the *Direct Object Restriction* (DOR): the result predicate must be predicated on the NP in object position (Levin & Rappaport Hovav 1995). That is, the result must be predicated of a transitive object or the subject of an unaccusative, but not the subject of a transitive or an unergative verb.[3] This restriction is shown in (18) with the transitive verb *gistak* 'hit'.

(18) *Rockyga wanįra šuuc gistakšąną.*
Rocky-ga wanį-ra šuuc ∅-gistak-šąną
Rocky-PROP meat-DEF red 3S/O-hit-DECL
= 'Rocky hit the meat red.'
≠ 'Rocky hit the meat and he was red as a result.'

As seen in (18), since *wanįra* 'the meat' is in object position, it can be the subject of the result, while the subject of the matrix verb, *Rocky*, cannot. The contrast in (18) points to the fact that the result is not a clause (i.e., a CP). I follow Li (1999) and assume that when the result can be linked to either the subject or the object and the result plus the means predicate is not formed in the lexicon (i.e., they do not form a compound), the resultative phrase is a clause with a

[3] Note that the DOR can also apply to so-called "fake" objects (e.g., reflexives) of unergative verbs. For example, the result phrase *hoarse* can be predicated on *herself* in (i). See Carrier & Randall (1992), Li (1999), and Wechsler (2005) for more details on the DOR. (i) The woman sang herself hoarse.

pro-controlled subject (see also Song 2005). According to Chomsky (1982), *pro* is an empty category of the type [+pronominal, −anaphoric], and Binding Theory states that it cannot be bound within its governing category. Thus, *pro* could be bound by either the matrix subject or object. Since the result in (18) cannot be linked to the subject, the result cannot be a clause.

Moreover, Hocąk resultatives show a contrast in availability between prototypical unaccusative and unergative verbs. Perlmutter (1978) defines unaccusative verbs as ones where the single argument is an underlying object, whereas the argument of an unergative verb is an underlying subject. Typically, unaccusative verbs denote change (e.g., *break, melt*) while unergative verbs indicate manner of motion (e.g., *run*) or other bodily functions (e.g., *cry*). In Hocąk, intransitive verbs that take stative agreement morphemes correspond to unaccusatives, and the set of intransitive verbs that bear active agreement morphemes are parallel to unergative verbs (see e.g., Williamson 1984, Woolford 2010). Prototypical unaccusatives in (19), such as *ziibre* 'melt' and *taaxu* 'burn', can serve as the matrix verb of resultatives. On the other hand, prototypical unergative verbs in (20), such as *nąąwą* 'sing' and *nąąk* 'run', cannot.

(19) a. *Xaigirara sgaasgap ziibre.*
xaigira-ra sgaasgap ∅-ziibre
chocolate-DEF sticky 3S-melt
'The chocolate melted sticky.'

b. *Waisgapra seep taaxu.*
waisgap-ra seep ∅-taaxu
bread-DEF black 3S-burn
'The bread burned black.'

(20) a. * *Hinųkra nįįra teek nąąwą.*
hinųk-ra nįį-ra teek ∅-nąąwą
woman-DEF throat-DEF sore 3S-sing
(Intended: 'The woman sang her throat sore.')

b. * *Henryga wagujirera paras nąąkšąną.*
Henry-ga wagujire-ra paras ∅-nąąk-šąną
Henry-PROP shoe-DEF flat 3S-run-DECL
(Intended: 'Henry ran the shoe(s) flat.')

Note that the restriction with unergative verbs also holds when the reflexive morpheme *kii-* denotes the so-called 'fake' reflexive/object of the predicate; see (21).[4]

(21) * *Hunterga hoix'įk kiinąqkšąną.*
Hunter-ga hoix'įk <***kii***>*∅-nąąk-šąną*
Hunter-PROP tired <REFL>3S-run-DECL
(Intended: 'Hunter ran himself tired.')

The DOR states the result must be predicated of the object. If we assume that the subjects of the verbs in (20) are underlying objects, we can maintain the DOR. On the other hand, since unergative verbs do not have an underlying object, no resultative interpretation is possible in (20) and (21).[5]

3.2 VP-internal status of the result predicate

In this subsection, I argue that the result predicate is the complement of the verb. I first show that the result predicate must be VP-internal, and then I provide evidence that resultatives in Hocąk project as a binary structure. Levin & Rappaport Hovav (1995) use VP-ellipsis in order to show that resultatives are VP-internal, and that they are part of the eventuality of the VP. Hocąk has a type of VP-ellipsis shown in (22) and (23): the light verb *ųų* can replace either a minimal VP or a multi-segmental VP, resulting from adjunction to VP. Example (22) shows an example of VP-ellipsis that targets on the object and the verb, while in (23), VP-ellipsis targets a VP-level adjunct, such as *xjanąre* 'yesterday'.

(22) *Cecilga* [$_{VP}$ *wažątirehižą ruwį*] *kjane anąga nee šge* [*haųų*] *kjane.*
Cecil-ga wažątire-hižą ∅-ruwį kjane anąga nee šge ha-ųų kjane
Cecil-PROP car-INDEF 3S/O-buy FUT and I also 1S-do FUT
'Cecil will buy a car, and I will too.' (Johnson 2013: 5)

[4] Under Washio's (1997) typology, intransitive resultatives are a type of weak resultative. For example, resultatives with an unergative verb like 'run' can form a weak resultative. Recall that Hocąk has transitive strong resultatives (see (13) above). Hocąk resultatives thus present a counterexample to Washio's typology: Hocąk has transitive strong resultatives but lacks intransitive strong resultatives. I leave further discussion of these examples with respect to Washio's typology for future work.

[5] The DOR holds consistently in English for transitive objects. In the case of unergative verb phrases, a fake reflexive/object ensures that there is an object that the result can be linked to. (See the translations in (20) and (21)).

(23) *Cecilga* [$_{VP}$ *xjanąre waši*] *anąga Bryanga šge* [*ųų*].
Cecil-ga xjanąre ∅-waši anąga Bryan-ga šge ∅-ųų
Cecil-PROP yesterday 3s-dance and Bryan-PROP also 3s-do
'Cecil danced yesterday, and Bryan did too.' (Johnson 2013: 6a)

As shown in (24b), it is not possible to "strand" the result predicate *šuuc* 'red' under VP-ellipsis. It thus follows that the result is inside the VP, rather than adjoined to VP.

(24) a. *Hunterga* [$_{VP}$ *nąąju seep hogiha*] *anąga Bryanga šge*
Hunter-ga nąąju seep ∅-hogiha anąga Bryan-ga šge
Hunter-PROP hair black 3s/O-dye and Bryan-PROP too
[*ųų*].
∅-ųų
3s-do
'Hunter dyed the hair black and Bryan did, too.'

b. * *Hunterga nąąju seep hogiha anąga Bryanga šge* ***šuuc***
Hunter-ga nąąju seep ∅-hogiha anąga Bryan-ga šge šuuc
Hunter-PROP hair black 3s/O-dye and Bryan-PROP too red
ųų.
∅-ųų
3s-do
(Intended: 'Hunter dyed the hair black and Bryan did red, too.')

Example (24) contrasts with (25). (25) contains the adverb *wasisik* 'energetically' as a depictive. Since depictives are typically analyzed as adjuncts that occupy a VP-external position (Levin & Rappaport Hovav 1995), they can be stranded.

(25) *Bryanga* [$_{VP}$ *waarucra hoix'įk waža*] *anąga Meredithga*
Bryan-ga waaruc-ra hoix'įk ∅-waža anąga Meredith-ga
Bryan-PROP table-DEF tired 3s/O-wipe and Meredith-PROP
wasisik [*ųų*].
wasisik ∅-ųų
energetic 3s-do
'Bryan wiped the table tired(ly) and Meredith did energetically.'

As we saw in (22), *ųų* affects the verb and its complement. Since a result predicate is not strandable with *ųų*, it must be the case that the result is inside the

minimal VP, and thus is part of the core eventuality of the VP. In other words, it follows that the result is inside the verb phrase.

Another option for the structure of resultatives could be that the verb, the result, and direct object are all sisters in a flat structure. Carrier & Randall (1992) provide such a ternary analysis for English resultatives. However, Bowers (1997) argues that a ternary structure cannot account for structures involving Across the Board movement. This type of movement describes a situation when a syntactic element moves from multiple base positions to a single terminal position. In this conjunctive test, the object and result of both conjuncts form a single constituent (see also Li 1999). An Across the Board structure is possible with Hocąk resultatives, as seen in (26), where the verb is moving across conjuncts.

(26) a. *Meredithga mąqsra paras gistak anąga waisgap pereįk.*
Meredith-ga mąqs-ra paras ∅-gistak anąga waisgap pereįk
Meredith-PROP metal-DEF flat 3S/O-hit and bread thin
'Meredith hit the metal flat and the bread thin.'

b. *Meredithga mąqsra gišįnįsįnį gižap anąga ważątirera*
Meredith-ga mąqs-ra gišįnįsįnį ∅-gižap anąga ważątire-ra
Meredith-PROP metal-DEF shiny 3S/O-polish and car-DEF
sgee.
sgee
clean
'Meredith polished the metal shiny and the car clean.'

The ability of Hocąk resultatives to participate in Across the Board movement is consistent with an analysis that argues for a binary structure (Bowers 1997). I conclude that Hocąk resultatives are straightforwardly analyzable under a binary branching approach. This provides another argument that the result is in a VP-internal position.

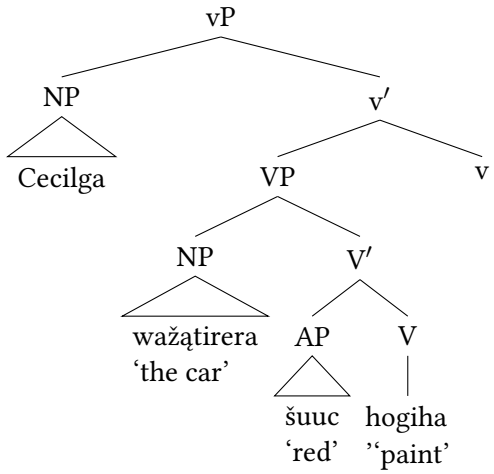
4 Syntactic representation of Hocąk resultatives

In this section, I propose that resultatives are in a Larsonian VP-shell structure (Larson 1988): a VP structure takes another VP as its complement. This approach follows Li's (1999) structure for English resultatives (cf. Hoekstra 1988; Carrier & Randall 1992; Levin & Rappaport Hovav 1995). Larson's (1988) VP-shells are intended to accommodate the double-object construction, where the left-most

object is in a higher position than the right-most. If we maintain a binary branching structure, then a resultative has the same structure as the double-object construction. I claim that the structure for Hocąk resultatives is depicted in (27). The result predicate is the complement of the verb, and I assume that the object is base-generated in Spec,VP. The subject is generated in Spec,vP, where "little v" is a semi-functional head that licenses external arguments (Chomsky 1995).

(27) a. *Cecilga wažątirera šuuc hogiha.*
Cecil-ga wažątire-ra šuuc ∅-hogiha
Cecil-PROP car-DEF red 3s/O-paint
'Cecil painted the car red.'

b.



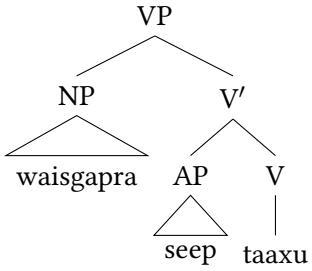
The structure in (27b) straightforwardly explains the facts with respect to Hocąk resultatives. First, the result is not a head that forms a compound with the matrix verb since adverbs and intensifiers can intervene. The structure in (27b) shows that the result is an AP and not a CP. This accounts for why the result cannot take complementizer, tense, negation, or declarative suffixes: the result is an AP, which does not contain clause-level heads or morphology. This property of result APs is also reflected in the fact that Hocąk resultatives obey the DOR. In (18), only the object 'meat' can be modified by the result 'red'. This restriction predicts that the result is not a clause. If the result were a clause, the subject in (18) could also be modified by 'red' because resultative phrases that project as CPs have a *pro* subject, which could be linked to the matrix subject. However, this is not the case. To formalize the relationship between the NP object and

the adjective, I follow Li's (1999) analysis. The AP can assign its theta-role to the object through mutual m-command.[6] In the case of (31), AP and the NP in Spec,VP are both dominated by the same VP node, and they do not dominate each other. Thus, the AP and the object NP mutually m-command each other. On the other hand, the AP does not hold a mutual m-command relationship with the subject in Spec,vP; thus, the AP cannot assign its theta-role to the subject. This results in the DOR effect.

This situation applies to resultatives with unaccusative matrix verbs, as depicted in (28b).

(28) a. *Waisgapra seep taaxu.*
waisgap-ra seep ∅-taaxu
bread-DEF black 3s-burn
'The bread burned black.'

b.



The AP in (28b) has the same position that it has in (27b); that is, it is the complement of the verb. Thus, the AP maintains the same relationship with the object in Spec,VP whether the verb is transitive or intransitive. Consequently, the AP *seep* 'black' and the object *waisgapra* 'the bread' are within the same VP, and the DOR effect is preserved. Data from VP-ellipsis has also demonstrated that the result phrase is inside the VP. This is in contrast to depictive phrases, where the depictive can be stranded by VP-ellipsis. Assuming the structure presented above, this contrast falls out naturally. Depictives have been analyzed as VP-adjuncts in English (Levin & Rappaport Hovav 1995); thus, I suggest that a depictive phrase, such as *wasisik* 'energetically' in (25), is adjoined to the upper VP-shell (i.e., vP) in (27b).[7]

[6] I assume that m-command refers to a syntactic relation where X m-commands Y if and only if the first maximal projection that dominates X also dominates Y and X does not dominate Y. In (31b), X is the NP *važątirera* 'the car', and Y is the AP *šuuc 'red'*.

[7] In this paper, I leave it open whether depictives can adjoin to the lower VP-shell.

To summarize, I have argued that the resultative secondary predicate is the complement to the main verb, and is a phrase. This accounts for a constellation of facts that concern the properties of Hocąk resultatives, including the DOR.

5 The result predicate and adjectives in Hocąk

Thus far I have assumed without comment that the result predicate is an adjective phrase. This section provides evidence that the result is in fact an AP, and thus that Hocąk has adjectives. Traditional grammars (e.g., Lipkind 1945 and Susman 1943) and more recently Helmbrecht (2006) have claimed that Hocąk lacks the lexical class adjectives since there is no distinct inflectional morphology for adjectives and verbs. Instead these works claim that adjectives are a class of stative verbs. For reasons of space, I consider only two of these arguments in detail.

First, Helmbrecht (2006) shows that there is no category establishing morphology with respect to adjectives. Recall that Hocąk has an active-stative split between intransitive verbs. Helmbrecht notes that purported adjectives and stative verbs exhibit parallel agreement morphology, as shown in (29) and (30), respectively.

(29) a. *hį-xete*
1-big
'I am big.'

b. nį-xete
2-big
'You are big.'

c. xete-ire
big-3S.PL
'They are big.'

(30) a. *hį-šiibre*
1-fall
'I fell.'

b. nį-šiibre
2-fall
'You fell.'

c. šiibre-ire
fall-3S.PL
'They fell.'

Example (29) illustrates that the stative set of agreement markers may be used with adjectives: in (29a,b), the prefixes *hiį-* and *niį-* mark 1st and 2nd person respectively, and in (29c) *-ire* encodes third person plural. The example in (30) with the stative verb *šiibre* 'fall' shows that this verb bears the same agreement markers. Since Hocąk is an active-stative language, the similarities between (29) and (30) follow if apparent adjectives are stative verbs. Second, apparent adjectives can be used predicatively without any morphological modification or without the help of auxiliaries, as seen in (31a). Helmbrecht (2006) asserts that the lack of auxiliaries is possible for all adjectives in Hocąk. This possibility extends to verbs as well. (31a) shows an example of the verb *nįį* 'swim'.

(31) a. *Wijukra seepšąną.*
wijuk-ra ∅-seep-šąną
cat-DEF 3S-black-DECL
'The cat is black.'

b. *Hocįcįkra nįį eeja nįįpšąną.*
hocįcįk-ra nįį eeja ∅-nįįp-šąną
boy-DEF water there 3S-swim.ACT-DECL
'The boy swam in the lake.'

Thus, since verbs and purported adjectives may also be the main predicate of the clause, there is no structural difference between adjectives and verbs.

In the following subsections, I will present two arguments that the resultative phrase projects as an AP in Hocąk resultatives.[8] In the first subsection, I argue that the linear ordering of the result and the matrix verb indicates that the result is an AP. In the second subsection, I turn to the fact that (stative) verbs are ungrammatical as a result predicate. I argue that only gradable predicates (i.e., adjectives) can participate in resultatives.

5.1 The Temporal Iconicity Constraint and resultatives

Following Li (1993), I suggest that the fact that the result precedes the verb in resultative predication provides evidence that the result is an adjective in Hocąk.[9] Specifically, I argue that since the result precedes the matrix verb in resultatives, Li's (1993) Temporal Iconicity Constraint would be violated if the result were a verb. Rather, since the result must precede the verb in Hocąk resultatives, the result must not be a verb. Instead, I claim that the result is an adjective.

Li (1993: 499) proposes his constraint in order to account for the restrictions on the order of verbs in V-V resultative compounds in Chinese and Japanese. The first V (V-cause) always encodes the event, while the second V (V-result) indicates the result of the event.

Li shows that V-cause must temporally and morphologically precede V-result. Li formalizes this constraint as in (32).

(32) *Temporal Iconicity Constraint (TIC)*:
Let A and B be two subevents (activities, states, changes of states, etc.) and let A′ and B′ be two verbal constituents denoting A and B, respectively;

[8] Baker (2003) has previously argued that a main characteristic of adjectives is that they can occur as secondary resultative predicates.

[9] Thanks to Yafei Li (personal communication) for bringing this diagnostic to my attention.

> then the temporal relation between A and B must be directly reflected in the surface linear order of A′ and B′ unless A′ is an argument of B′ or vice versa.

For example, Li notes that in both Chinese and Japanese, V-cause is the first verb of the compound. Consider the Chinese example in (33) and the Japanese example in (34).

(33) *Táotao tiào-fán-le* (*Yōuyou le*).
Taotao jump-bored-ASP Youyou LE
'Taotao jumped and as a result he/(Youyou) got bored.' (Li 1993: 480 (1b))

(34) *John-ga Mary-o karakai-akiru-ta.*
John-NOM Mary-ACC tease-bored-PAST
'John teased Mary and as a result John got bored.' (Li 1993: 481 (2b))

What is important to note here is that V-cause always precedes V-result. In (33), the V-cause *tiào* 'jump' necessarily precedes V-result *fán* 'bored'. Without the parentheses in (33), Taotao's jumping causes Youyou to become bored. With the parentheses in (33), Taotao's jumping makes himself become bored. In (34), the V-cause *karakai* 'tease' must appear to the left of the V-result *akiru* 'bored'. A further piece of evidence for the TIC comes from serial-verbs in Sranan and Ịjọ. Sranan is syntactically a head-initial language, whereas Ịjọ is head-final. Both examples in (35) illustrate that the verb phrase that denotes getting ahold of the instrument linearly precedes the central action. That is, 'take the knife' in Sranan comes before 'cut the bread', and the same pattern is seen in Ịjọ with 'basket take' preceding 'yam cover'.

(35) a. Sranan; SVO
Mi e teki a nefi koti a brede.
I ASP take the knife cut the bread
'I cut the bread with the knife.'

b. Ịjọ; SOV
Áràú su-ye ákì buru teri-mí.
she basket take yam cover-PAST
'She covered a yam with a basket.' (Li 1993: 500, (38))

We find similar evidence from manner-of-directed motion serial verbs in Hocąk. These serial verbs consist of a manner-of-motion verb (e.g., *nųųwąk* 'run') and a directional motion verb (e.g., *hii* 'arrive'). In Hocąk, the order of these two

verbs cannot be reversed. Example (36) shows that the linear order of *nųųwąk* 'run' and *hii* 'arrive' must be *nųųwąk-hii*. The verb *hii* 'arrive' must always be the second verb. This directly follows from the TIC: a running event must logically precede the arriving event.

(36) a. *Matejaga Teejop eeja nųųwąk* ***hii****.*
Mateja-ga Teejop eeja nųųwąk ∅*-hii*
Mateja-PROP Madison there run 3S-arrive
'Mateja ran to Madison.'

b. * *Matejaga Teejop eeja* ***hii*** *nųųwąk.*
Mateja-ga Teejop eeja hii ∅*-nųųwąk*
Mateja-PROP Madison there arrive 3S-run
(Intended: 'Mateja ran to Madison.')

Despite the strong predictions that the TIC makes, it is not intended to account for all resultative constructions. According to Li's proposal, the TIC applies only if two conditions are met: one, the constituents involved are both verbal, and two, the verbal constituents must not be in a predicate-argument relation (e.g., causatives). Here I am only concerned with the first condition, as this second condition does not apply to Hocąk resultatives. Li presents an example from German to illustrate the first constraint, as in (37).

(37) *Er will das Eisen flachschlagen.*
he wants the iron flat.pound
'He wants to pound the iron flat.' (Li 1993: 501 (41))

The result encoded by *flach* 'flat' linearly precedes the activity *schlagen* 'pound'. Since *flach* 'flat' is an adjective, Li claims that the TIC does not apply. Rather the head-final structure of German determines the order of *flach* 'flat' and *schlagen* 'pound'.

In summary, while the TIC applies to verbal constituents, the TIC has nothing to say about when adjectives form similar events with verbs.

Let us return to the Hocąk data. We see that the result precedes the matrix verb, as in (38a). That is, *paras* 'flat' linearly precedes *gistak* 'hit'. In fact it is ungrammatical for the result to be postverbal, as shown in (38b).

(38) a. *Meredithga mąąsra* ***paras*** *gistakšąną.*
Meredith-ga mąąs-ra paras ∅*-gistak-šąną*
Meredith-PROP metal-DEF flat 3S/O-hit-DECL
'Meredith hit the metal flat.'

b. * *Meredithga mąąsra gistakšąną,* ***paras.***
Meredith-ga mąąs-ra ∅-gistak-šąną paras
Meredith-PROP metal-DEF 3S/O-hit-DECL flat
(Intended: 'Meredith hit the metal flat.')

Accordingly, if apparent adjectives in Hocąk are stative verbs, then the grammaticality of examples like (38a) is surprising. We expect (38a) to be ungrammatical, given the TIC. Since the TIC does not rule out examples like (38a), we can conclude that the result is not a verb. This is similar to the German example in (37). Moreover, the fact that the order that the TIC predicts, as in (38b), is ungrammatical also leads to the conclusion that the result is not a verb.[10] I take this as evidence that the result is an AP.

5.2 Barring verbs as the result

In this section, I show that adjectives can appear in resultative secondary predication, while verbs cannot. In order to account for the contrast, I argue that we need to slightly refine the structure of the result phrase: the result phrase in Hocąk is an AP that contains a degree phrase. Following Corver (1997), I assume that only gradable adjectives have a degree argument, and that degree heads need to bind such a degree argument. I show that non-gradable adjectives are incompatible with resultatives in Hocąk. Thus, if verbs do not have a degree argument to be discharged, the structure will be ruled out as an instance of vacuous quantification. Compare (39) that has *žiipįk* 'short' as the result with (40) that uses the verb *šiibre* 'fall'.

(39) *Matejaga peešjįra* ***žiipįk*** *rucgisšąną.*
Mateja-ga peešjį-ra žiipįk ∅-rucgis-šąną
Mateja-PROP hair-DEF short 3S/O-cut-DECL
'Mateja cut the hair short.'

(40) * *Matejaga peešjįra* ***šiibre*** *rucgisšąną.*
Mateja-ga peešjį-ra šiibre ∅-rucgis-šąną
Mateja-PROP hair-DEF fall 3S/O-cut-DECL
(Intended: 'Mateja cut the hair (so that) it falls.')

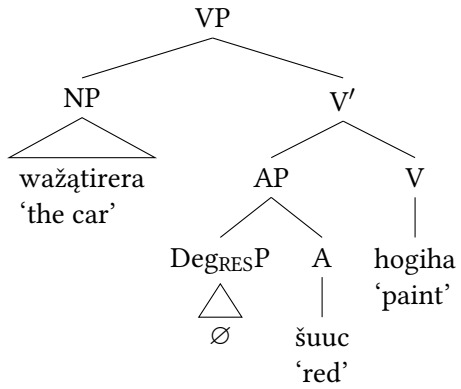
[10] More needs to be said as to why the result cannot be postverbal. Johnson & Rosen (2014) propose that constituents are moved to a postverbal position via an EPP feature that can only attract DPs. I leave a full explanation of this issue open for now.

The ungrammaticality of a verb like *šiibre* 'fall' in a resultative construction (40) indicates that this predicate is somehow fundamentally different than the one in (39). If we take a closer look at Hocąk, we notice that verbs are not the only elements that cannot be a secondary resultative predicate. While I argue that only adjectives can be resultative predicates in Hocąk, not all adjectives are available in this position. Crucially, non-gradable adjectives cannot appear as a result predicate. The example in (41) illustrates this for the non-gradable adjective *t'ee* 'dead', which is ungrammatical as a result predicate. Note that the English equivalent is grammatical, as indicated by the translation in (41).

(41) * *Bryanga caara t'ee guucšąną.*
Bryan-ga caa-ra t'ee ∅-guuc-šąną
Bryan-PROP deer-DEF dead 3S/O-shoot-DECL
(Intended: 'Bryan shot the deer dead.')

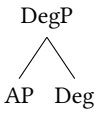
To account for the restriction seen in (41), I claim that the resultative predicate in Hocąk takes a DegP in its specifier, as shown in (42). I label this degree phrase "Deg_{RES}P."[11]

(42)



[11] Corver (1997) argues that DegP dominates the AP (as in (i)). Differently than Corver, the structure in (42) follows Jackendoff (1977) and Bhatt & Pancheva (2004), among others, and places DegP in Spec,AP. Nothing crucially hinges on the placement of the degree phrase, however.

(i)



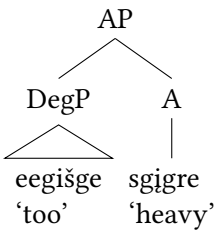
Hocąk resultatives are thus obtained by specifying the eventuality of the result to the highest degree. This is consistent with Wechsler's (2005) proposal on the constraints on the result predicate. Wechsler asserts that the result must express a gradable property with a maximum degree, when the object NP is the argument of the matrix verb. I assume that only gradable adjectives can take DegPs in their specifiers, while non-gradable adjectives lack this ability. The degree head is an operator and thus has to bind a variable. If gradable adjectives have a degree argument (or grade-role) in its argument structure, then the degree head will be able to bind it. On the other hand, if non-gradable adjectives lack this degree argument, then the structure will be ruled out since all operators have to bind a variable. Consider the contrast between the gradable adjective *sgįgre* 'heavy' in (43a) and the non-gradable adjective *t'ee* 'dead' in (43b) with the degree element *eegišge* 'too'.

(43) a. *Henryga eegišge sgįgre.*
Henry-ga eegišge ∅-sgįgre
Henry-PROP too 3s-heavy
'Henry is too heavy.'

b. * *Caara eegišge t'ee*
Caa-ra eegišge ∅-t'ee
deer-DEF too 3s-dead
(Intended: 'The deer is too dead.')

I propose that *eegišge* realizes a Deg head; thus, an example like in (43a) has the structure in (44).[12]

(44)



I attribute the ungrammaticality of non-gradable adjectives with *eegišge* 'too' to the hypothesis that the degree head associated with eegišge 'too' must bind

[12] As noted in footnote 11, it could also be the case that DegP dominates the AP. I suggest that *eegišge* 'too' would be in the specifier of a head-final and phonologically null Deg. See Rosen (2015) for more information.

the degree argument of a lexical item. Since non-gradable adjectives in Hocąk do not have degree arguments as part of their lexical entry (cf. Higginbotham 1985; Corver 1997), the degree head does not have a degree argument to bind. This results in ungrammaticality. Under the present analysis, since the result takes a degree phrase in its specifier, it is expected that a non-gradable adjective is not allowed as a result predicate. In the case of the resultative in (41), *t'ee* 'dead' is ill-formed because the degree operator in Deg does not have a variable in its scope that it can bind.[13] Let us return to the fact that verbs are ungrammatical as resultative predicates. Following Higginbotham (1985) and Corver (1997), I assume that verbs do not have a grade-role; rather they have an event-role. This is evidenced by the ungrammaticality of the verb *šiibre* 'fall' with *eegišge* 'too' in (45).[14]

(45) * *Hunterga* *eegišge šiibre.*
Hunter-ga *eegišge ∅-šiibre*
Hunter-PROP too 3S-fall.STAT
(Intended: 'Hunter fell too much/a lot.')

I argue that resultative examples like (40) with verbs are ungrammatical since a degree head can measure the state of the adjective, but it cannot link the event of a verb. In other words, example (40) is ruled out because there is a mismatch between the selectional restrictions of DegP and a verb phrase. This explains why verbs are barred from resultatives in Hocąk. In this subsection, we see that verbs cannot appear as the result predicate in Hocąk. The reason that verbs cannot appear as the result, I claim, is that the result predicate takes a special degree phrase that I labeled "Deg_{RES}P" in its specifier. A straightforward explanation arises if we assume that degree phrases in Hocąk must bind a degree argument. Since I am assuming that verbs lack a degree argument, verbs are not allowed as a result predicate. Thus, I contend that the result predicate in Hocąk is an AP.

5.3 Implications: Status of adjectives

I have presented evidence that the result predicate in Hocąk resultatives projects as an AP (an adjective). This puts resultatives in Hocąk in line with resultative constructions cross-linguistically that use APs as result predicates (cf. English

[13] I assume that color adjectives, such as *šuuc* 'red', are gradable (Kennedy & McNally 2010).

[14] The contrast between (43) and (45) illustrates another way in which stative verbs and adjectives differ. In this paper, I am only concerned with how they differ with respect to resultatives. Rosen (2014; 2015) presents more diagnostics for the existence of adjectives in Hocąk.

resultatives). Moreover, these data indicate that Hocąk has the lexical category adjective. This is a significant result since Hocąk has been previously described as only having nouns and verbs (see §2.3). The previous traditional literature (e.g., Helmbrecht 2006) has focused primarily on the morphological similarities between stative verbs and adjectives. The data from resultatives have shown that these similarities can be misleading. Rather adjectives surface in at least one environment in Hocąk; namely, resultatives (see Rosen 2014; 2015 for further discussion of these issues in Hocąk, and Baker 2003 and Dixon 2004 cross-linguistically).

6 Conclusion

This paper has offered a description and an analysis of Hocąk resultatives. I have shown that the result predicate must not be a clause and must be in a VP-internal position. I have argued that Hocąk resultatives project a phrasal AP as the complement of the verb in a Larsonian "VP-shell" (Larson 1988). This proposal is supported by the fact that resultatives in Hocąk have many of the properties that have been attributed to resultatives cross-linguistically, such as in English. In particular, resultatives in Hocąk obey the DOR, and the resultative phrase is adjectival. I conclude with the hope that this paper will continue to improve our understanding of resultatives and the structure of predication in Hocąk and Siouan languages.

Acknowledgments

I am extremely grateful to Cecil Garvin for teaching me about his language. Thanks also to Meredith Johnson, Yafei Li, Becky Shields, Catherine Rudin, Rand Valentine, and an anonymous reviewer for valuable discussion and comments, and to Iren Hartmann for allowing me to use her Lexique Pro dictionary. Lastly, material in this paper was presented at SSILA 2014. I would like to thank the audience there.

Abbreviations

1, 2, 3	first, second, third person	O	object agreement
		PL	plural
COMP	complementizer	PROP	proper noun
DECL	declarative	REFL	reflexive
DEF	definite	S	subject agreement
INDEF	indefinite		

References

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.

Bhatt, Rajesh & Roumyana Pancheva. 2004. Late merger of degree clauses. *Linguistic Inquiry* 35. 1–45.

Bowers, John. 1997. A binary analysis of resultatives. *Texas Linguistic Forum* 38. 43–58.

Carrier, Jill & Janet Randall. 1992. The argument structure and syntactic structure of resultatives. *Linguistic Inquiry* 23. 173–234.

Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Corver, Norbert. 1997. The internal syntax of the Dutch extended adjectival projection. *Natural Language and Linguistic Theory* 15. 289–368.

Dixon, Robert Malcolm Ward. 2004. Adjective classes in typological perspective. In Robert Malcolm Ward Dixon & Alexandra Y. Aikhenvald (eds.), *Adjective classes: A cross-linguistic typology*, 1–49. Oxford: Oxford University Press.

Helmbrecht, Johannes. 2006. Are there adjectives in Hocąk (Winnebago)? In Ximena Lois & Valentina Vapnarsky (eds.), *Lexical categories and root classes in Amerindian languages*, 289–316. Bern: Peter Lang International Academic Publishers.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16. 547–94.

Hoekstra, Teun. 1988. Small clause results. *Lingua* (74). 101–139.

Jackendoff, Ray. 1977. *X-bar syntax*. Cambridge, MA: MIT Press.

Johnson, Meredith. 2013. *Verb-phrase ellipsis and* v*: Evidence from Hocąk*. Paper presented at the 87th annual meeting of the Linguistic Society of America, Boston.

Johnson, Meredith & Bryan Rosen. 2014. *Rightward movement and postverbal constituents: New evidence from Hocąk*. Paper presented at the 50th meeting of the Chicago Linguistics Society, University of Chicago.

Kennedy, Chris & Louise McNally. 2010. Color, context and compositionality. *Synthese* 174. 79–98.

Larson, Richard K. 1988. On the double-object construction. *Linguistic Inquiry* 19. 335–392.

Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.

Li, Yafei. 1993. Structural head and aspectuality. *Language* 69. 480–504.

Li, Yafei. 1999. Cross-componential causativity. *Natural Language and Linguistic Theory* 17. 445–497.

Lipkind, William. 1945. *Winnebago grammar*. New York: King's Crown Press.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70. 369–415.

Perlmutter, David. 1978. Impersonal passives and the unaccusative hypothesis. In *Proceedings of the 4th meeting of the Berkeley Linguistics Society*, 157–190. Berkeley: University of California –Berkeley.

Rosen, Bryan. 2014. *On the existence of adjectives in Hocąk*. Paper presented at the 88th annual meeting of the Linguistic Society of America, Minneapolis.

Rosen, Bryan. 2015. *The syntax of adjectives in Hocąk*. Madison: University of Wisconsin (Doctoral dissertation).

Song, Hongkyu. 2005. *Causatives and resultatives in Korean*. Madison: University of Wisconsin (Doctoral dissertation).

Susman, Amelia L. 1943. *The accentual system of Winnebago*. New York: Columbia University (Doctoral dissertation).

Washio, Ryuichi. 1997. Resultatives, compositionality and language variation. *Journal of East Asian Linguistics* 6. 1–49.

Wechsler, Stephen. 2005. Resultatives under the "event-argument homomorphism" model of telicity. In Nomi Erteschik-Shir & Tova Rapoport (eds.), *The syntax of aspect*, 255–273. Oxford: Oxford University Press.

Williams, Alexander. 2008. Word order in resultatives. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 507–515. Somerville, MA: Cascadilla Press.

Williamson, Janis Shirley. 1984. *Studies in Lakhota grammar*. San Diego: University of California – San Diego (Doctoral dissertation).

Woolford, Ellen. 2010. Active-stative agreement in Choctaw and Lakota. *Revista Virtual de Estudos da Linguagem* 8. 6–46.

Chapter 14

Evidence for a VP constituent in Hocąk

Meredith Johnson

Bryan Rosen

Mateja Schuck

Since at least Williamson 1984, there has been a debate over the configurationality of Siouan languages (Boyle 2007; Graczyk 1991; West 2003; Van Valin 1985; 1987). In this paper, we argue that a nonconfigurational approach does not account for the asymmetries between subjects and objects in Hocąk. We propose that Hocąk is a configurational language in that the language has a verb phrase (VP): the object and the verb form a constituent to the exclusion of the subject. This structure captures the differences between subjects and objects with respect to locative scope, quantifier scope, verb-phrase ellipsis, and resultatives.

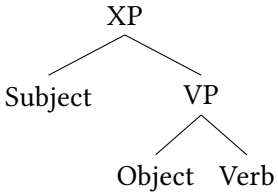
1 Introduction

Since at least Williamson 1984, there has been a debate over the configurationality of Siouan languages (Boyle 2007; Graczyk 1991; West 2003; Van Valin 1985; 1987). The purpose of this paper is to weigh in on this issue with evidence (based on original fieldwork) from Hocąk. By providing novel data from locative scope, quantifier scope, verb-phrase ellipsis, and resultatives, we argue that Hocąk has a verb phrase (VP). This adds empirical support for previous studies that have argued that Siouan languages have a verb phrase (e.g., Boyle 2007; Graczyk 1991; West 2003).

The crucial observation that we make is in this paper is that there exist a number of subject-object asymmetries. To account for these data, we propose a syntax for Hocąk that consists minimally of the structure shown in (1).

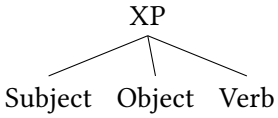
Meredith Johnson, Bryan Rosen & Mateja Schuck. 2016. Evidence for a VP constituent in Hocąk. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 339–363. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.177 DOI:10.17169/langsci.b94.177

(1)



By contrast, we argue that a flat, nonconfigurational structure such as the one in (2) cannot adequately account for the data (cf. Van Valin 1985; 1987; Williamson 1984).

(2)



This paper is organized as follows. In §2, we outline previous analyses that argue in favor of a flat structure for various Siouan languages, and then discuss how the Hocąk data compare. §3 reviews arguments for a VP in other Siouan languages, and shows that similar arguments can be made for Hocąk. In §4, we provide four new arguments in favor a of VP analysis of Hocąk. §5 concludes the paper.

2 Arguments in favor of a flat structure

In this section, we provide background on the nature of configurationality in the context of Hocąk (and other Siouan languages). §2.1 outlines the previous nonconfigurational accounts (Hale 1983 and Jelinek 1984) that stand in contrast to the configurational account that we propose in this paper. In §2.2, we review the previous arguments for a flat VP structure in Siouan languages. Then in §2.3, we show that Hocąk displays all three of the prototypical characteristics of being a nonconfigurational language.

2.1 Nonconfigurationality and pronominal arguments: Hale (1983) and Jelinek (1984)

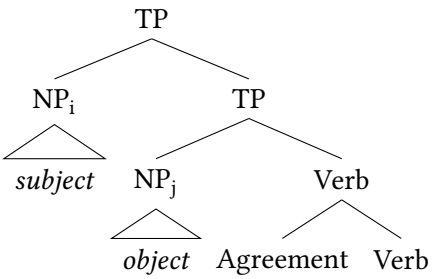
Since Hale (1983), nonconfigurational languages have been typologically characterized by the three traits given in (3):

(3) Properties of nonconfigurational languages
   i. Free word order
   ii. Extensive null anaphora
   iii. Presence of discontinuous constituents

Hale's approach makes use of two levels of representation: *lexical structure* (LS) and *phrase structure* (PS). Hale argues that all languages are configurational at LS; that is, the subject asymmetrically c-commands the object. However, this asymmetry is not realized at the level of PS in nonconfigurational languages: the phrase structure is flat. This is the definition of configurationality that is most adopted by Siouanists. For example, Boyle (2007) claims that Hidatsa is a configurational language on the grounds that there are subject-object asymmetries that are indicative of a VP constituent. (See also Van Valin 1985; 1987; Williamson 1984, and West 2003.)

Another formal account of nonconfigurationality is Jelinek's (1984) *Pronominal Argument Hypothesis* (PAH). According to the PAH, person markers are the actual arguments of the verb, while the overt NPs are adjuncts adjoined high in the clause, as in (4). We use "TP" (Tense Phrase) for the phrase that represents the sentence level.

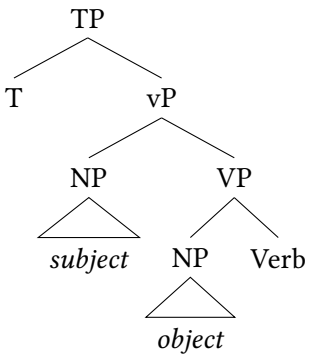
(4)



The overt NPs, when present, are coindexed with the person markers. Since adjuncts are known to have freer distribution of word order than arguments, the "free" word order in nonconfigurational languages is accounted for. Adjuncts are also never obligatory, explaining the possibility of *pro*-drop of all NPs in nonconfigurational langugages. Lastly, this proposal accounts for the presence of apparent discontinuous constituents in nonconfigurational languages. Jelinek proposes that more than one adjunct NP can be coindexed with a given person marker. Thus, what appear to be discontinuous NPs are actually two separate NPs that correspond to the same argument.

In contrast, a configurational language is one that that does show subject-object asymmetries and has a VP constituent, as depicted in (5) below.

(5)



Example (5) shows that the subject and object are not in adjunct positions: they do not adjoin to the TP (or Sentence). Following Chomsky (1995), we assume that the subject is base-generated in a position outside of the VP, which we label "vP." The object merges as an argument of the verb inside the VP. Thus, by "VP" we refer to the constituent that contains the object, the verb, and perhaps other modifier material. Crucially, the subject is not considered part of the VP.

2.2 Previous analyses: Williamson (1984), Van Valin (1985, 1987)

In this section, we discuss arguments in favor of a nonconfigurational analysis of Siouan languages that have been put forth in previous works.

Williamson (1984) argues that Lakota is nonconfigurational because it lacks the subject-object asymmetries traditionally associated with the Empty Category Principle (ECP). Long distance *wh*-extraction of the subject over an overt complementizer is possible in Lakota; that is, the language does not display *that*-trace effects. Long distance extraction out of *wh*-islands from subject position is also allowed in Lakota. Examples (6)–(8) below illustrate these facts:

(6) *Mary tuwa wąyąke* ***ki*** *ilukcha he?*
Mary who see COMP you.think Q
'Who do you think that Mary saw?' (Williamson 1984: 281, (64a))

(7) *Tuwa hel nažį he* ***ki*** *ilukcha he?*
who there stand DUR COMP you-think Q
'Who do you think that was standing there?' (Williamson 1984: 281, (65a))

(8) *Toհą tuwa u pi **ki** slolyaya he?*
when who come PL COMP you.know Q
'Who do you know when is coming?' (Williamson 1984: 281, (66a))

In a language with subject-object asymmetries, long-distance *wh*-extraction of the subject should not be possible, as doing so would constitute a violation of the ECP (as evidenced by the ungrammaticality of the English translations in (7)-(8)). Because Lakota appears to allow long-distance *wh*-extraction from subject position, Williamson argues that the language has no subject-object asymmetry and thus lacks a VP constituent.

Van Valin (1985; 1987) also argues for a nonconfigurational analysis of Lakota on the basis of the lack of Weak Crossover and Binding Condition C effects. First, let us consider the diagnostic from Weak Crossover (WCO). A WCO violation occurs when a pronoun is coreferential with the *wh*-trace in subject or object position and neither one c-commands the other (Sportiche 1985). (9) illustrates an English example of WCO: the *wh*-word *who* undergoes movement from an object position (represented by a trace, "t") to the left edge of the clause. *Who* must "cross over" the co-indexed pronoun *his*. Since *who* and *his* cannot refer to the same person, the sentence is ungrammatical.

(9) *Who$_i$ does his$_i$ mother love t$_i$?

Thus in a language with a VP node, a coreferential reading between the *wh*-word and possessive pronoun in the sentence in (10) below would be expected to be unavailable.

(10) *∅-tha-khóla-ku ki tuwá wąyąka he?*
3-POSS-friend-POSS the who 3SG.see.3SG Q
'Who$_i$ did his$_i$ friend see?' (Van Valin 1987: 379)

Because the construction in (9) does not cause a WCO violation in Lakota, Van Valin argues that no subject-object asymmetry exists in the language, and thus it does not possess a VP.

Van Valin additionally cites the lack of Binding Condition C (BCC) violations in Lakota as evidence that the language lacks a subject-object asymmetry. This is due to the fact that binding conditions crucially rely on a c-command relationship between anaphors, pronouns and r-expressions. Van Valin argues that since there appear to be no BCC violations in Lakota, the subject must not c-command the object. This falls out of an analysis where both NPs are attached at the TP (or sentence) level. We return to BCC violations in the next section.

2.3 Hocąk data

Hale (1983) and Jelinek (1984) identify three properties that they claim are common to all nonconfigurational languages: free word order, extensive null anaphora, and discontinuous constituents. Below, we show that Hocąk does display each of the three classic signs of nonconfigurationality put forth by Hale and Jelinek, as well as a number of additional characteristics of nonconfigurational languages proposed by Baker (1996).

First, NP arguments may appear in a variety of orders. This is expected in an analysis under which there is a flat structure and all NPs are adjuncts adjoined at the TP (or sentence) level. SOV word order is the most common in Hocąk, as in (11). Any variation in word order has discourse-informational effects, as hinted at by the English translations given in the examples below. As shown in (12a), a participant displaced to the left serves a topic or focus function, whereas participants displaced to the right are interpreted as anti-topics (e.g., "backgrounded" or discourse-old), as shown in (12b)–(12e).

(11) *Hinųkra wažątirera ruwį.*
hinųk-ra wažątire-ra ∅-ruwį
lady-DEF car-DEF 3S/O-buy
'The lady bought the car.'

(12) a. *Wažątirera, hinųkra ruwį.*
wažątire-ra hinųk-ra ∅-ruwį
car-DEF lady-DEF 3S/O-buy
'The car, the lady bought it.'

b. *Wažątirera ruwį, hinųkra.*
wažątire-ra ∅-ruwį hinųk-ra
car-DEF 3S/O-buy lady-DEF
'Someone bought the car, (it was) the lady.'

c. *Hinųkra ruwį, wažątirera.*
hinųkra ∅-ruwį, wažątire-ra
lady-DEF 3S/O-buy car-DEF
'the lady bought something, (it was) the car.'

d. *Ruwį, wažątirera, hinųkra.*
∅-ruwį wažątire-ra hinųk-ra
3S/O-buy car-DEF lady-DEF
'Someone bought something, (it was) the car, the lady.'

e. *Ruwį, hinųkra, važątirera.*
∅-ruwį hinųk-ra važątire-ra
3s/O-buy lady-DEF car-DEF
'Someone bought something, (it was) the lady, the car.'

It is also possible for NP arguments to have freedom of placement among each other. The default order of arguments in a ditransitive construction is Agent > Indirect Object > Direct Object; however, their order can vary. This is shown below in (13), where the subject NP *hinųkhižą* 'a woman' can appear in several different positions.

(13) *(Hinųkhižą,) hocįcįhižą (hinųkhižą,) wiiwagaxhižą (hinųkhižą,)*
hinųk-hižą hocįcį-hižą hinųk-hižą wiiwagax-hižą hinųk-hižą
woman-INDEF boy-INDEF woman-INDEF pencil-INDEF woman-INDEF
hok'ų.
∅-hokų
3s/O-give
'A woman gave a boy a pencil.'

Second, NPs corresponding to arguments can be freely omitted in Hocąk. Examples of this are shown below in (14), where the agent and patient/theme arguments are omitted:

(14) a. *Wijųkra šųųkra hoxataprookeeja haja.*
wijųk-ra šųųk-ra hoxatap-rook-eeja ∅-haja
cat-DEF dog-DEF woods-inside-there 3s/O-see
'The cat saw the dog in the woods.'

b. *Hoxataprookeeja haja.*
hoxatap-rook-eeja ∅-haja
woods-inside-there 3s/O-see
'It (the cat) saw it (the dog) in the woods.'

Sentence (14b) is grammatical and can (under the right discourse context) have the equivalent meaning to (14a); however, it is missing the agent and patient/theme NPs *wijųkra* and *šųųkra*. This is also expected under Hale's (1983), Jelinek's (1984), and Baker's (1996) analyses: NPs have adjunct status and thus are not obligatory.

Hocąk also displays discontinuous constituents. Demonstratives and quantifiers may be separated from the head noun, as shown in (15) with *že'e* 'that':

(15) a. *Wijukra šųųk že'e haja.*
wijuk-ra šųųk že'e ∅-haja
cat-DEF dog that 3S/O-see
'The cat saw that dog.'

b. *Že'e wijukra šųųk haja.*
že'e wijuk-ra šųųk ∅-haja
that cat-DDEF dog 3S/O-see
'The cat saw that dog.'

Discontinuous constituents are expected under the analyses of nonconfigurationality by Hale (1983) and Jelinek (1984), due to the fact that NPs have the status of adjuncts. Hale and Jelinek propose that multiple adjuncts can be associated with the same argument in a given sentence. Thus, the demonstrative and head noun in (15b) are actually two separate NPs that both correspond to the object.

In addition to Hale's (1983) classic characteristics of nonconfigurationality, Hocąk displays four additional traits of nonconfigurational languages discussed by Baker (1996). First, Hocąk does not display BCC effects within clauses. As discussed in the previous section, this lack of BCC effects is expected when there is no asymmetry between the subject and the object In (16) below, coreference between the subject 'he' and the possessor 'Bryan' is grammatical.

(16) *(Ee) Bryanga hi'ųni hiira homąkįnį.*
Ee Bryan-ga hi'ųni hii-ra ∅-homąkįnį
he Bryan-PROP mother POSS-DEF 3S/O-visit
'He$_i$ visited Bryan$_i$'s mom.'

However, as Baker shows to be true in other nonconfigurational languages, Hocąk does display BCC effects across clauses. In (17), coreference between the matrix subject 'she' and the embedded object 'Meredith' is impossible.

(17) *(Ee) Hunterga Meredithga hajara hiraperesšąną.*
Ee Hunter-ga Meredith-ga ∅-haja-ra ∅-hiraperes-šąną
she Hunter-PROP Meredith-PROP 3S/O-see-COMP 3S-know-DECL
'She$_{*i/j}$ knows that Hunter saw Meredith$_i$.'

Second, Hocąk lacks NP anaphors, which are also argued by Baker (1996) to be nonexistent in nonconfigurational languages. Instead, reflexive and reciprocal meanings are expressed morphologically on the verb, as seen in (18):

(18) *Meredithga anąga Hunterga hokikijįire.*
Meredith-ga anąga Hunter-ga <kiki>hojį-ire
Meredith-PROP and Hunter-PROP <REFL>hit-3S.PL
'Meredith and Hunter hit each other.'

Third, according to Baker (1996), nonconfigurational languages should lack both universal quantifiers that are grammatically singular and negative quantifiers. Hocąk does not have a universal quantifier that is grammatically singular. In (19) below, both *hanąąc* 'all/every' and *hižąkišąną* 'each' trigger plural agreement on the verb.

(19) a. *Bryanga waisgap sguu xuwuxuwura hanąą waruucšąną.*
Bryan-ga waisgap sguu xuwuxuwu-ra hanąą wa-∅-ruuc-šąną
Bryan-PROP cookie-DEF all 3O.PL-3S-eat-DECL
'Bryan ate every cookie/all of the cookies.'

b. *Hocįcįra hižąkišąną waisgap sguu xuwuxuwuhižą ruucire.*
hocįcį-ra hižąkišąną waisgap sguu xuwuxuwu-hižą ruuc-ire
boy-DEF each cookie-INDEF eat-3S.PL
'Each boy ate a cookie.'

Hocąk also does not possess negative quantifiers: instead, the equivalents to 'nothing' and 'nobody' are expressed through a combination of clausal negation and indefinite pronouns. This is shown in (20a) and (20b), respectively.

(20) a. *Wawaahiwira hąąke ważą hiiranį.*
wa<ha>hohi-wi-ra hąąke ważą hii-ire-nį
3O.PL<1S>beat-1/2PL-COMP NEG thing do-3S.PL-NEG
'When we beat them, they didn't score at all.' (Hartmann 2012)

b. *Hąąkižą nįįtašjak taaxura karasgepnį.*
hąąke-hižą nįįtašjak taaxu-ra ∅-kara-rasgep-nį
NEG-INDEF coffee-DEF 3S-own-drink.up-NEG
'Nobody finished his coffee.'

Finally, Hocąk lacks WCO effects. In (21) below, a coreferential reading between the possessive pronoun and the object *wh*-word is grammatical.

(21) a. *Hi'ųni hiira peežega haja?*
hi'ųni hii-ra peežega ∅-haja
mother 3POSS-DEF who 3S/O-see

b. *Peežega hi'ųni hiira haja?*
peežega hi'ųni hii-ra ∅-haja
who mother 3POSS-DEF 3S/O-see
'Who$_i$ did his$_i$ mother see?'

Recall from the previous subsection that Van Valin (1985; 1987) uses the lack of BCC and WCO effects in Lakota to argue for a nonconfigurational syntax. While Hocąk also lacks BCC and WCO effects, we argue that this does not constitute conclusive evidence of the lack of a VP constituent in the language. In the remainder of the paper, we provide other arguments that strongly favor a VP analysis for Hocąk. We leave an explanation for the lack of BCC and WCO effects in Hocąk for future research.

3 Arguments in favor of a VP

3.1 Previous analyses: Boyle (2007), Graczyk (1991), West (2003)

In the previous section, we presented arguments in favor of a nonconfigurational, VP-less analysis in several Siouan languages. In this section, we present arguments in favor of a configurational analysis of Siouan languages (that is, arguments in favor of a VP analysis). The first piece of evidence comes from word order restrictions. Recall that one of Hale's (1983) and Jelinek's (1984) typifying characteristics of nonconfigurational languages is free word order. Across Siouan languages, neutral word order is SOV. Several Siouanists have argued that other word orders have discourse-informational effects, and thus that word order is not actually free in these languages. For example, West (2003) shows that in Assiniboine sentences with OSV word order, the fronted object has a preferred focus reading; otherwise, the first argument is interpreted as the subject. This is shown below in (22).

(22) *Škóšobena wãží hokšína že yúda.*
banana a boy DET *ate*
'The boy ate a banana (not the apple).' (preferred translation) or
'A banana ate the boy' (West 2003: 49)

The same is true of Hidatsa. Boyle (2007) shows that unmarked word order is SOV, with exceptions occurring in topicalization or focus constructions. This is shown below in (23) with neutral SOV word order and (24) OSV order:

(23) *Buushígesh washúgash éegaac.*
puušíhke-š mašúka-š éekaa-c
cat-DET.D dog-DET.D see-DECL
'The cat sees the dog.' (Boyle 2007: 214)

(24) *Masúgash buushígesh éegaac.*
masúka-š puusíhke-š ékaa-c
dog-DET.D cat-DET.DDET see-DECL
'The cat sees the dog.' (Boyle 2007: 214)

Graczyk (1991) observes that SOV is neutral word order for Crow as well, and that other word orders have discourse-informational effects. This is shown below, where (25) has neutral word order, and (26) has OVS word order:

(25) *Shikáak-kaatee-sh ashé hii-ák.*
boy-DIMIN-DET home reach-SS
'The little boy reached home.' (Graczyk 1991: 101)

(26) *Iaxp-úua ítchi-kiss-uua-sh kootáa híi-k hinne talée-sh.*
their.feather-PL good-sport-PL-DET entirely reach-DECL this oil-DET
'It entirely covered their beautiful feathers, this oil.' (Graczyk 1991: 103)

In (26), OSV word order is used to deemphasize the discourse-old subject *talee* 'oil', and emphasize the object *iaxp* 'their feather'. Based on these word order restrictions, West, Boyle and Graczyk all argue that Assiniboine, Hidatsa and Crow are configurational.

The second piece of evidence that has been previously used to show the presence of a VP in Siouan languages comes from enclitics. West (2003) and Boyle (2007) use the scope of enclitics to argue for a VP constituent. Boyle (2007) demonstrates that the Hidatsa habitual enclitic *-ʔii* takes scope over both verbs in the example in (27) below:

(27) *Doosha wiriʔéeraga adáʔa kʰúuiidoogʔ*
tooš^{h}*a wiri-éeraka atá-a kʰúu-ʔii-took*
how sun-DEM appear-CONT come.up-HAB.SG.SPEC
'How does the Sun always appear and come up? (he wondered)' (Boyle 2007: 223)

The situation is the same in Assiniboine. In (28) below, the aspectual clitic *s'a* scopes over both verbs, not just to the one to which it is attached:

(28) *Wiyą́-bi žé-na woyúta spãyą́-bi hikná hayábi gaǧéǧe-bi s'a.*
woman-PL the-PL food cook-PL CONJ clothes sew-PL HAB
'The women usually cooked the food and sewed the clothes.' (West 2003: 39)

The sentence in (28) cannot mean 'the women cooked the food and usually sewed the clothes' (West 2003). If Assiniboine had no VP, this reading should not be possible: the clitic should only be able to scope over the verb it is attached to. Both Boyle (2007) and West (2003) argue that the clitics head a functional projection that c-commands the coordinated elements, which are VPs. Thus, enclitic scope provides evidence in support of the existence of a VP in Hidatsa and Assiniboine.

It has been argued for other Siouan languages (Boyle 2007, West 2003) that coordination itself targets VPs, since coordination can target a constituent that includes the object and verb. In contrast, coordination can never target the subject and verb to the exclusion of the object. Boyle (2007) shows that in Hidatsa, the subject of the second clause must be the same as the subject of the first clause in (29):

(29) *Alex wía ikáaa réec.*
Alex wía ikáa-a rée-c
Alex woman see-CONT leave-DECL
'Alex saw the woman and (Alex/*the woman) left.' (Boyle 2007: 217)

West (2003) provides similar data from Assiniboine to support a configurational analysis, as shown in (30) below:

(30) *Wíyã že* [*wicá še wayága*] *hĩkná* [*céya*].
woman DET man the see CONJ cry
'The woman saw the man and cried.'
*'The woman saw the man and he cried' (West 2003: 34)

As in Hidatsa, the subject of the second conjoined verb *céya* 'cry' in (30) can only be *wíyã* 'the woman'. In a nonconfigurational language, either NP should be able to be the subject of the second verb; thus Boyle and West argue that Hidatsa and Assiniboine are configurational and have a VP constituent.

3.2 Hocąk data

In the previous subsection, we presented previous arguments for a configurational analysis of several Siouan languages. In this section, we show that the

tests used by Boyle (2007) for Hidatsa, Graczyk (1991) for Crow, and West (2003) for Assiniboine yield the same results when applied to Hocąk.

First, word order is crucial to disambiguate subjects and objects in Hocąk. In (31) below, the first argument is interpreted as the subject:

(31) *Wijukra šųųkra haja.*
wijuk-ra šųųk-ra ∅-haja
cat-DEF dog-DEF 3S/O-see
'The cat saw the dog.'
≠ 'The dog saw the cat.'

A reading in which the dog saw the cat is also possible for (31), but only when the first argument is followed by an intonational pause.

As shown in the previous section, Boyle (2007) and West (2003) provided evidence from enclitic scope to show that Hidatsa and Assiniboine have a VP constituent. The same proves true in Hocąk. In (32)-(34) below, the enclitics *gįnį* 'already', *ege* 'might' and *žeeži* 'hopefully' take scope over both coordinated verbs in the (b) examples, even though they are only attached to the second verb.

(32) a. *Hunterga toora tuuc wahiigįnį.*
Hunter-ga too-ra tuuc wa-∅-hii=gįnį
Hunter-PROP potato-DEF be.cooked 3O.PL-3S-CAUS=already
'Hunter already cooked the potatoes.'

b. *Hunterga toora tuuc wahii anąga*
Hunter-ga too-ra tuuc wa-∅-hii anąga
Hunter-PROP potato-DEF be.cooked 3O.PL-3S-CAUS and
warucgįnį.
wa-∅-ruuc=gįnį
3O.PL-3S-eat=already
'Hunter already cooked the potatoes and ate them.'

(33) a. *Matejaga tookewehiege.*
Mateja-ga ∅-tookewehi=ege
Mateja-PROP 3S/O-be.hungry=might
'Mateja might (very well) get hungry.'

b. *Matejaga tookewehi anąga kerege.*
Mateja-ga ∅-tookewehi anąga ∅-kere=ege
Mateja-PROP 3S-be.hungry and 3S-leave=might
'Mateja might (very well) get hungry and leave.'

(34) a. *Bryanga nįįtašjak taaxu ruwįžeeži.*
Bryan-ga nįįtašjak taaxu ∅-ruwį=žeeži
Bryan-PROP coffee 3S/O-buy=wish
'Hopefully Bryan will buy coffee.'

b. *Bryanga nįįtašjak taaxu ruwį anąga hųųk'ųžeeži.*
Bryan-ga nįįtašjak taaxu ∅-ruwį anąga <hį>∅-hok'ų=žeeži
Bryan-PROP coffee 3S/O-buy and <1O>3S-give=wish
'Hopefully Bryan will buy coffee and give it to me.'

If Hocąk lacked a VP, this pattern would be unexpected: the clitics should only be able to scope over the verb to which they are attached. Instead, the clitics in the (b) examples above take scope over both coordinated verb phrases. This indicates that the constituent that clitics scope over is a VP, and that these enclitics attach at the VP level.

Lastly, Boyle (2007) and West (2003) showed that coordination targets VPs in Hidatsa and Assiniboine, providing further evidence for a configurational analysis of these languages. Coordination also targets VPs in Hocąk, as shown in (35) and (36) below. In these examples, the subject of the first conjunct, *wąąkwažoonįra* 'the hunter', must also be the subject of the second conjunct. Example (36) is especially revealing, as the only possible meaning is not as pragmatically plausible: it would (arguably) be more likely for the bear to die in that scenario.

(35) *Wąąkwažoonįra hkukura ruxe ankaga t'eehii.*
wąąkwažoonki-ra hkukuc-ra ∅-ruxe ankaga ∅-t'ee-hii
hunter-DEF bear-DEF 3S/O-chase and 3S-die-CAUS
'The hunter chased and killed the bear.'

(36) *Wąąkwažoonįra hųųcra guuc anąga t'ee.*
wąąkwažoonį-ra hųųc-ra ∅-guuc anąga ∅-t'ee
hunter-DEF bear-DEF 3S/O-shoot and 3S-die
'The hunter shot the bear and [the hunter] died.'

If there was no subject-object asymmetry, either 'hunter' or 'bear' should be a possible subject for the second conjuncts in (35) and (36). Thus, these examples show that coordination in Hocąk targets a constituent that excludes the subject; namely, the VP.

4 New Evidence for a VP in Hocąk

4.1 Scope of Locatives

The first piece of new evidence for a VP involves the interpretation of locative adjuncts. The neutral position of locative adjuncts is shown in (37) with *hoxataprookeeja* 'in the woods' appearing between the object and the verb.

(37) *Wijukra šuukra hoxataprookeeja haja.*
wijuk-ra šuuk-ra hoxatap-rook-eeja ∅-haja
cat-DEF dog-DEF woods-inside-there 3S/O-see
'The cat saw the dog in the woods.'

The translation in (37) is ambiguous. The English sentence has three possible interpretations, as outlined in (38) below.

(38) a. The cat is in the woods, and it saw the dog. The dog is not in the woods.
b. The dog is in the woods, and the cat saw the dog. The cat is not in the woods.
c. Both the cat and the dog are in the woods, and the cat saw the dog.

In Hocąk, however, only the interpretations in (38b) and (38c) are available for (37); that is, the locative adjunct must describe the location of the object. This is true even if the locative *hoxataprookeeja* 'in the woods' is clause-initial or clause-final, as in (39a) and (39b) , respectively. These sentences cannot have the reading in (38a), where only the dog can be in the woods.

(39) a. *Hoxataprookeeja, wijukra šuukra haja.*
hoxatap-rook-eeja wijuk-ra šuuk-ra ∅-haja
woods-inside-there cat-DEF dog-DEF 3S/O-see
'In the woods, the cat saw the dog.'
b. *Wijukra šuukra haja, hoxataprookeeja.*
wijuk-ra šuuk-ra ∅-haja hoxatap-rook-eeja
cat-DEF dog-DEF 3S/O-see woods-inside-there
'The cat saw the dog in the woods.'

A nonconfigurational analysis cannot readily account for this subject-object asymmetry: if Hocąk had a flat structure, we would not expect the locative to be able to modify only the object.

Alternatively, we argue that the object NP is the unique complement to the verb. We account for the scope facts by suggesting that the locative phrase can

merge in two locations. If the locative adjoins to the VP (that is, the constituent that contains the object and the verb) then the reading in (38b) is available: the locative only has scope over the object. On the other hand, if the locative adjoins to a position above the VP, then the reading in (38c) is obtained: the locative then scopes over both arguments.

4.2 Verb-Phrase Ellipsis (VPE)

As first discussed by Johnson (2013), Hocąk displays a process of VPE in which the light verb *ųų* replaces the verb and the object, to the exclusion of the subject (40):

(40) *Cecilga wažątirehižą ruwį kjane anąga nee šge haųų kjane.*
Cecil-ga wažątire-hižą ∅-ruwį kjane anąga nee šge ha-ųų kjane
Cecil-PROP car-INDEF 3S/O-buy FUT and I also 1S-do FUT
'Cecil will buy a car, and I will too.'

The examples in (41) show that VPE also targets certain adjuncts. (41a) shows that VPE targets VPs containing temporal adjuncts. In (41b), a locative adjunct is included in the ellipsis site. (41c) exemplifies VPE with a comitative. In all of these examples, the adjunct in the antecedent VP is interpreted as being present in the ellipsis site, indicating that *ųų* targets the entire VP rather than just the object.

(41) a. *Cecilga xjanąre waši anąga Bryanga šge ųų.*
Cecil-ga xjanąre ∅-waši anąga Bryan-ga šge ∅-ųų
Cecil-PROP yesterday 3S-dance and Bryan-PROP also 3S-do
'Cecil danced yesterday, and Bryan did too.'

b. *Cecilga ciinąk eja wažątirehižą ruwį anąga Bryanga*
Cecil-ga ciinąk eja wažątire-hižą ∅-ruwį anąga Bryan-ga
Cecil-PROP city there car-INDEF 3S/O-buy and Bryan-PROP
šge ųų.
šge ųų.
also ∅-ųų 3S-do
'Cecil bought a car in the city, and Bryan did too.'

c. *Cecilga hinųkra hakižu waši anąga Bryanga šge*
Cecil-ga hinųk-ra hakižu ∅-waši anąga Bryan-ga šge
Cecil-PROP woman-DEF be.with 3S-dance and Bryan-PROP also

ųų.
∅-ųų
3s-do
'Cecil danced with the woman, and Bryan did too.'

Constructions with *ųų* cannot be analyzed as a *pro*-form, as object extraction is permitted. (42a) shows that focused elements can be extracted from the ellipsis site. Furthermore, antecedent-contained deletion (ACD) is also possible ex:jrs:42b. ACD would not be possible if *ųų* were a *pro*-form, since the head of the relative clause is the object of the elided verb phrase.

(42) a. *Meredithga waagaxra ruwį, nųnįge **wiiwagaxra** hąąke*
Meredith-ga waagax-ra ∅-ruwį nųnįge wiiwagax-ra hąąke
Meredith-PROP paper-DEF 3S/O-buy but pencil-DEF NEG
ųųnį.
∅-ųų-nį
3S-do-NEG
'Meredith bought paper but didn't (buy) pencils.'

b. *Bryanga ruwį, įaagu Meredithga ųųra.*
Bryan-ga ∅-ruwį įaagu Meredith-ga ∅-ųų-ra
Bryan-PROP 3S/O-buy what Meredith-PROP 3S-do-COMP
'Bryan bought what(ever) Meredith did.'

VPE is also permitted in embedded clauses and adjuncts, which is also inconsistent with a *pro*-form analysis. (43a) exemplifies VPE in an embedded clause, and (43b)–(43c) show that ellipsis sites are licit inside adjunct clauses.

(43) a. *Bryanga hąąke nįįtašjak taaxu ruwįnį, nųnįge*
Bryan-ga hąąke nįįtašjak taaxu ∅-ruwį-nį nųnįg
Bryan-PROP NEG coffee 3S/O-buy-NEG but
Meredithga ųųra yaaperesšąną.
Meredith-ga ∅-ųų-ra <ha>hiperes-šąną
Meredith-PROP 3S-do-COMP <1S>know-DECL
'Bryan didn't buy coffee, but I know Meredith did.'

b. *Bryanga ųų kjanegi Meredithga Hunterga (nišge)*
Bryan-ga ∅-ųų kjane-gi Meredith-ga Hunter-ga nišge
Bryan-PROP 3S-do FUT-if Meredith-PROP Hunter-PROP also

gišja hii kjane.
∅-gišja hii kjane.
3S/O-visit FUT
'Meredith will visit Hunter if Bryan will.'

c. *Bryanga hąąke ųųnįge Meredithga (nišge) hąąke*
Bryan-ga hąąke ∅-ųų-nį-ge Meredith-ga nišge hąąke
Bryan-PROP NEG 3S-do-NEG-because Meredith-PROP also NEG
Hunterga gišja hiinį.
Hunter-ga gišja hii-nį
Hunter-PROP 3S/O-visit-NEG
'Meredith didn't visit Hunter because Bryan didn't.'

The presence of VPE constitutes strong evidence for a configurational analysis of Hocąk: in a flat structure, there is no VP constituent that can be targeted by ellipsis. Since at least Ross (1969), the presence of VPE in English has been used as an argument in favor of a VP constituent that contains the verb and object to the exclusion of the subject. Hocąk also displays VPE, which leads us to conclude that Hocąk must have a VP constituent.

4.3 Quantifier scope

Another piece of evidence in favor of a configurational analysis of Hocąk comes from quantifier scope. As discussed in Johnson (2014) and Johnson & Rosen (2014), linear order determines the scope of quantified phrases in Hocąk. In a sentence with SOV word order, the subject obligatorily distributes over the object. This is shown below in (44a), where the sentence can only describe a situation in which each man caught a different fish. However, the interpretation changes with SVO word order: (44b) can only describe a situation in which each man caught the same fish. Lastly, in a sentence with OVS word order, the subject scopes over the object, as shown in (44c).

(44) a. *Wąąkra hižąkišąną hoohižą gisikire.*
wąąk-ra hižąkišąną hoo-hižą ∅-gisik-ire
man-DEF each fish-INDEF 3O-catch-3S.PL
'Each man caught a fish.' (each > a; *a > each)

b. *Wąąkra hižąkišąną gisikire, hoohižą.*
wąąk-ra hižąkišąną ∅-gisik-ire, hoo-hižą
man-DEF each 3O-catch-3S.PL fish-INDEF
'Each man caught a fish.' (a > each; *each > a)

c. *Hoohižą gisikire, wąąkra hižąkišąną.*
hoo-hižą ∅-gisik-ire, wąąk-ra hižąkišąną
fish-INDEF 3O-catch-3S.PL man-DEF each
'Each man caught a fish.' (each > a; *a > each)

These facts cannot be adequately accounted for if the subject and object are in a flat structure in Hocąk: there is no principled way that linear order could account for the interpretation of the sentences in (44). In contrast, the interpretation of basic SOV word order in (44a) is straightforwardly explained under a VP analysis: the subject is higher than the object and thus scopes over it. Furthermore, we follow Johnson (2014) and Johnson & Rosen (2014) and propose that postverbal objects (44b) and subjects (44c) obligatorily take wide scope because they undergo movement that targets a position high in the clause.

4.4 Resultatives and the Direct Object Restriction

We now turn to an argument from resultatives in Hocąk. Resultatives are complex predicates that put together a means predicate (i.e., a verb) and a result predicate, where neither is licensed by a conjunction or an adposition (Williams 2008: 507). As seen in (45), Hocąk exhibits resultatives: (45a) shows that the result *paras* 'flat' is immediately to the left the verb *gistak* 'hit', and a similar example is shown in (45b) with the result *šuuc* 'red and the verb *hogiha* 'paint'.

(45) a. *Meredithga mąąsra paras gistakšąną.*
Meredith-ga mąąs-ra paras ∅-gistak-šąną
Meredith-PROP metal-DEF flat 3S/O-hit-DECL
'Meredith hit the metal flat.'

b. *Cecilga ważątirera šuuc hogiha.*
Cecil-ga ważątire-ra šuuc ∅-hogiha
Cecil-PROP car-DEF red 3S/O-paint
'Cecil painted the car red.'

Subjects and objects behave differently in the resultative construction. First, only the object can be modified by the result. Second, only prototypical unaccusative verbs can be used in the resultative construction. We use both of these pieces of evidence to support our claim that there is a VP constituent in Hocąk.

It has previously been observed for other languages, such as English, that the resultative predicate must be linked to the "deep" object of the verb. Levin & Rappaport Hovav (1995) refer to this constraint as the *Direct Object Restriction* (henceforth, DOR). In particular, the restriction states that only the object of a

transitive verb or the subject of an unaccusative verb can be modified by the result predicate. In contrast, a result predicate cannot be linked to the subject of an unergative verb. Consider the representative English examples below in (46).

(46) a. John hammered the metal flat. (transitive)
b. The water froze solid. (unaccusative)
c. *The dog barked hoarse. (unergative; ungrammatical as resultative)

Hocąk resultatives obey the DOR. This is restriction is shown in (47) with the transitive verb *gistak* 'hit'.

(47) *Rockyga wanįra šuuc gistakšąną.*
Rocky-ga wanį-ra šuuc ∅-gistak-šąną
Rocky-PROP meat-DEF red 3S/O-hit-DECL
= 'Rocky hit the meat red.'
≠'Rocky hit the meat red and he was red as a result.'

Since *wanįra* 'the meat' is in object position, it can be modified by the result, while the subject of matrix verb Rocky cannot. Thus, (47) establishes a clear subject-object asymmetry. If Hocąk had a flat structure, we would not expect the result to only be able to modify the object. In other words, the asymmetry would be difficult to explain without the presence of a VP constituent.

Furthermore, only unaccusative (as opposed to unergative; cf. Perlmutter 1978) verbs are compatible with resultatives in Hocąk. This is demonstrated by the contrast between (48) and (49).

(48) a. *Xaigirara sgaasgap ziibre.*
xaigira-ra sgaasgap ∅-ziibre
chocolate-DEF sticky 3S-melt
'The chocolate melted sticky.'
b. *Waisgapra seep taaxu.*
waisgap-ra seep ∅-taaxu
bread-DEF black 3S-burn
'The bread burned black.'

(49) a. * *Hinukra nįįra teek nąąwą.*
hinuk-ra nįį-ra teek ∅-nąąwą
woman-DEF throat-DEF sore 3S/O-sing
(Intended: 'The woman sang her throat sore.')

b. * *Henryga waguįirera paras nąąkšąną.*
Henry-ga waguįire-ra paras ∅-nąąk-šąną
Henry-PROP shoe-DEF flat 3S/O-run-DECL
(Intended: 'Henry ran the shoe(s) flat.')

Prototypical unaccusatives, such as *ziibre* 'melt' and *taaxu* 'burn', can serve as the matrix verb of resultatives in (48). On the other hand, prototypical unergative verbs, such as *nąąwą* 'sing' and *nąąk* 'run', cannot, as in (49). Compare the Hocąk examples in (49) to the English example in (46c). (46c) is ungrammatical because there was no object present for the result predicate to modify. In contrast, while the Hocąk examples in (49) have an object, they are still ungrammatical.

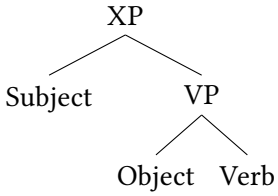
Assuming Perlmutter's (1978) unaccusative hypothesis, the single argument of an unaccusative verb is internal to the VP, whereas the argument of an unergative verb is VP-external. The contrast between (48) and (49) provides evidence that Hocąk has an unaccusative-unergative split:[1] if there were no such distinction between unaccusative and unergative verbs, (49) would be expected to be grammatical, contrary to fact. If the Hocąk VP were flat, we would not expect unergative verbs with resultatives to be ungrammatical. As a result, this shows that the VP in Hocąk is not flat: we conclude that the data in this section provides further evidence for a VP in Hocąk.

4.5 Structure of the Hocąk VP

In the sections above, we have seen that Hocąk shows subject-object asymmetries with respect to word order, the enclitic scope, and coordination. These same subject-object asymmetries have been previously documented in other Siouan languages. We also demonstrated that the facts from VPE, resultatives and the scope of adjuncts and arguments constitute additional subject-object asymmetries. The fact that we find so many asymmetries between the subject and object indicates that the subject and the object do not both form a constituent with the verb. Instead, we argue that these facts can be accounted for if the object is the complement of the verb in a VP constituent. The subject is base generated in a phrase that is external to the VP, which we tentatively label "XP." A basic transitive verb phrase is represented in (50).

[1] To the best of our knowledge, such a split has not been previously observed in Hocąk. However, see Williamson (1984) and West (2003), among others, for possible unaccusative-unergative splits in Lakota and Assiniboine, respectively.

(50)



5 Conclusion

The question of whether Siouan languages are configurational or nonconfigurational has been under debate for the past three decades. In this paper, we have presented new evidence to support a configurational analysis of Hocąk. We first showed that the tests previously used by Boyle (2007) for Hidatsa, Graczyk (1991) for Crow and West (2003) for Assiniboine to argue in favor of a VP constituent are also applicable in Hocąk. Next we presented novel evidence from locative scope verb-phrase ellipsis, quantifier scope, and resultative constructions which further support our claim that a VP constituent exists in Hocąk.

Acknowledgments

We would like to extend our deepest thanks to Cecil Garvin, our language consultant, without whom this research would never have come to be. Thanks also to Iren Hartmann for access to her Lexique Pro dictionary.

Abbreviations

The abbreviations used in the Hocąk examples are:

1, 2, 3	first, second, third person
COMP	complementizer
DECL	declarative
DEF	definite
DUR	durative
FUT	future
INDEF	indefinite
NEG	negative
O	object agreement
POSS	possessive
Q	question
PROP	proper noun
PST	past tense
isi PL	plural
REFL	reflexive
S	subject agreement
SG = singular	

The glosses for data from other languages follow the conventions of the works they are drawn from.

References

Baker, Mark C. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.

Boyle, John P. 2007. *Hidatsa morphosyntax and clause structure*. Chicago: University of Chicago (Doctoral dissertation).

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Graczyk, Randolph Jan. 1991. *Incorporation and cliticization in Crow morphosyntax*. Chicago: University of Chicago (Doctoral dissertation).

Hale, Kenneth Locke. 1983. Warlpiri and the grammar of nonconfigurational languages. *Natural Language and Linguistic Theory* 1. 5–49.

Hartmann, Iren. 2012. *Hoocąk Lexique Pro database*.

Jelinek, Eloise. 1984. Empty categories, case and configurationality. *Natural Language and Linguistic Theory* 2. 39–76.

Johnson, Meredith. 2013. *Verb-phrase ellipsis and* v*: Evidence from Hocąk*. Paper presented at the 87th annual meeting of the Linguistic Society of America, Boston.

Johnson, Meredith. 2014. Rightward movement and antecedent-contained deletion: New evidence from Hocąk. In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st west coast conference on formal linguistics*, 258–265.

Johnson, Meredith & Bryan Rosen. 2014. *Rightward movement and postverbal constituents: New evidence from Hocąk*. Paper presented at the 50th meeting of the Chicago Linguistics Society, University of Chicago.

Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.

Perlmutter, David. 1978. Impersonal passives and the unaccusative hypothesis. In *Proceedings of the 4th meeting of the Berkeley Linguistics Society*, 157–190. Berkeley: University of California –Berkeley.

Ross, John Robert. 1969. Auxiliaries as main verbs. *Studies in Philosophical Linguistics* 1. 77–102.

Sportiche, Dominique. 1985. Remarks on crossover. *Linguistic Inquiry* 16. 460–469.

Van Valin, Robert D., Jr. 1985. Case marking and the structure of the Lakhota clause. In Johanna Nichols & Anthony C. Woodbury (eds.), *Grammar inside and outside the clause: Some approaches to theory from the field*, 363–413. Cambridge: Cambridge University Press.

Van Valin, Robert D., Jr. 1987. The role of government in the grammar of head-marking languages. *International Journal of American Linguistics* 53. 371–397.

West, Shannon Louise. 2003. *Subjects and objects in Assiniboine Nakoda*. Victoria: University of Victoria (Master's thesis).

Williams, Alexander. 2008. Word order in resultatives. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 507–515. Somerville, MA: Cascadilla Press.

Williamson, Janis Shirley. 1984. *Studies in Lakhota grammar*. San Diego: University of California – San Diego (Doctoral dissertation).

Part IV

Cross-Siouan studies

Introduction to Part IV

Part IV consists of three chapters which compare some area of grammar across the Siouan family. The phenomena vary – coordination constructions, intonational and morphosyntactic variations as indices of information structure, and the expression of possession – but the approach of all three authors is similar in that they compare the facts of a number of different Siouan languages, from different branches of the Siouan family, within the framework of a broader, cross-linguistic typology.

Catherine Rudin ("Coordination and related constructions in Omaha-Ponca and in Siouan languages"), rather than demonstrating a shared characteristic of the Siouan family, shows that there is no common or typical Siouan coordination pattern. Words meaning 'and' or 'or' are not cognate among the various Siouan languages, sometimes even within one branch of Siouan. The syntactic structure of conjoining constructions varies from language to language as well, sometimes involving structures other than true coordination. This suggests that syntactic coordination was not found in Proto-Siouan, but developed later in the individual languages.

Bryan James Gordon ("Information-Structural Variations in Siouan Languages") looks at the frequently ignored area of information structure, using a corpus study of variations in constituent order, reduction, intonational contour and other markings corresponding to specific types of topic, focus and linking relations across a range of Siouan languages. The chapter includes discussion of the methodology the author developed for coding intonational and information structures.

Johannes Helmbrecht ("NP-internal possessive constructions in Hoocąk and other Siouan languages") surveys at least one language from each branch of the Siouan family and examines how they express different types of possession. Besides Ho-Chunk, the languages treated are Crow, Hidatsa, Mandan, Lakota, Osage and Biloxi. These languages have many similarities, but differ in how four basic morphosyntactic possessive types match up with semantic categories of possession (ownership, attribution of property, kinship, etc.).

Chapter 15

Coordination and related constructions in Omaha-Ponca and in Siouan languages

Catherine Rudin

Syntactic constructions expressing semantic coordination vary widely across the Siouan language family. A case study of possible coordinating conjunctions in Omaha-Ponca demonstrates that distinguishing coordination from other means of expressing 'and' relations is a non-trivial problem. A survey of words translated as 'and,' 'or,' or 'but' in Siouan languages leads to the conclusion that neither coordinating conjunctions nor the syntactic structures containing them are reconstructable across the Siouan family. It is likely that Proto-Siouan lacked syntactic coordination.

1 Introduction

All languages have ways of expressing additive, disjunctive, and adversative relations among entities or propositions. In European languages these relations are expressed by two distinct syntactic means: coordination and subordination. In Siouan languages these two types of conjunction construction are also present, but the distinction between them is less robust and less clear; coordination may not have existed at all historically. Neither coordinating conjunctions ('and,' 'or,' 'but') nor the syntactic structures containing them are reconstructable across the Siouan family.

I begin this examination of coordination in Siouan by defining coordination and discussing some of the issues involved in distinguishing coordinate from subordinate conjunction (§2). This is followed in §3 by a case study of additive coordination and coordinate-like constructions in Omaha-Ponca, the Siouan language with which I am most familiar. §4 is a survey of available data on coordination across all branches and most of the languages in the Siouan language family, with a summary table. §5 concludes the chapter with a discussion of the

Catherine Rudin. 2016. Coordination and related constructions in Omaha-Ponca and in Siouan languages. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 367–393. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.178 DOI:10.17169/langsci.b94.178

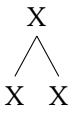
(non)universality of coordination constructions and some speculations on the history and origins of coordination in Siouan.

2 Issues in defining and identifying coordination

2.1 The syntax of coordination

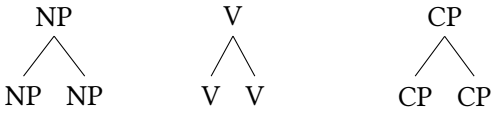
Traditionally, coordination is a structure of the type shown in (1):

(1)



In this structure two or more conjuncts of identical grammatical category together constitute a larger syntactic unit of the same category. These conjuncts might for instance be noun phrases, verbs, or clauses:

(2)



The conjuncts are sisters, of equal syntactic status, in a symmetrical constituent. Neither coordinate is subordinate to or included in the other. Equality of status is seen by coordinate NPs bearing the same case and triggering plural agreement in languages where those categories are overtly marked. In addition coordinate phrases resist extraction (Ross 1968 (1967), Coordinate Structure Constraint), and any movement out of them must be "across the board" movement out of all the conjuncts. Thus in standard English when two pronouns are coordinated, as in (3); they must both be nominative in subject position, they require a plural verb, and they cannot be separated.

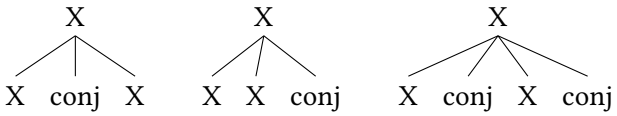
(3) a. [She and I] were chosen.
b. * [She and me] were chosen.
c. * [She and I] was chosen.
d. * [She] was chosen [and I].

This contrasts with a non-coordinate construction like that in (4), in which the two pronouns are different cases, the verb is singular, agreeing with only the first pronoun, and the subordinate portion of the construction can be moved.

(4) a. [She] [along with me] was chosen.
b. [She] was chosen [along with me].

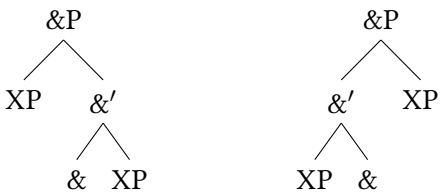
Coordinate constructions may or may not contain an overt coordinating conjunction, a word translating as 'and,' 'or,' 'but', etc. If there is one, it may occur between the conjuncts, or after the last one, or may be repeated (before or after each conjunct):

(5)



In recent theories of syntax (i.e. Minimalism), coordinate structures are instead treated as asymmetric constructions headed by the coordinator: "CoordP" or "
isi&P," or the similar "Boolean Phrase" structure argued for by Munn (1993). This type of structure is adopted partly for theory-internal reasons such as Kayne's Linear Correspondence Axiom (1994), but also for reasons having to do with intonation, ellipsis, and other phenomena which often suggest that the conjunction is more closely associated with one conjunct than with the other. See Citko (2011) for detailed discussion. Under this view coordinate structures look something like those in (6); presumably Siouan languages, being strongly head-final, would tend to have the left-branching variant shown on the right:

(6)



Issues of whether the conjunction forms a constituent with either the preceding or following X, and whether there is such a thing as a Coordination Phrase, are obviously important if one is concerned with distinguishing "true" coordination from other constructions such as comitatives which have similar meanings. Under the "
isi&P" analysis coordination has a syntactic configuration much like comitative or subordination structures, with one conjunct higher than the other, making it less straightforward to explain the distinctive behavior of coordinate structures,

as well as less clear what criteria distinguish coordinate from subordinate structures. Numerous works have wrestled with these issues theoretically and across languages, e.g. Wesche (1995) and Fabricius-Hansen & Ramm (2008). I lack data to deal with such questions in most of the Siouan languages, so the exact structure of apparently coordinate phrases is left vague in what follows. Detailed research within each language will be needed to sort it out.

It is likely that many of the structures which translate 'and/or/but' in various Siouan languages are actually not coordinate. Several other types of syntactic constructions often express semantic coordination. These include at least the following: (1) comitatives (prepositional phrases or subordinate clauses expressing 'accompaniment' or a 'with' relation); (2) adverbial clauses with temporal or other subordinate relations to a matrix clause ('when,' 'although,' 'having done X,' etc.); (3) simple listing of nouns, verbs, or clauses (that is, concatenation of separate items which do not form a larger constituent of any kind, sometimes with elements meaning 'too,' 'also,' 'furthermore,' 'however' or a phrase which sums them up ('both,' 'all'); (4) co-subordinate or clause-chaining constructions, (see e.g. Graczyk 2007; Boyle 2007).

There are a number of problematic coordination constructions in languages of the world, for instance a coordinator analyzed as a transitive verb in a Papua New Guinean language (Brown & Dryer 2009), partial/covert coordination of the *nie s Ivan* 'we with Ivan' = 'Ivan and I ' type in Slavic (e.g. McNally 1993; Larson 2014), special treatment of commonly linked items (Wälchli 2005), and overlaps with serial constructions (Carstens 2002). I do not deal with these specifically, but mention them just as a further reminder that the syntax of coordination is not necessarily a simple issue. For a useful typological overview of coordination, see Haspelmath (2007); other general treatments include Johannessen (1998) and van Oirsouw (1987).

2.2 The semantics of coordination

Coordinators join constituents with diverse semantic relations, though the semantic aspects of coordination have received less attention than its syntax. Different authors use widely varying terminology for the meanings coordination can express; see for instance Citko's (2011) discussion of Andrej Malchukov's system of classification of coordination constructions into semantic types. Among the terms used in the literature are *additive, adversative, comitative, consecutive, concessive, contrastive, correction, disjunctive, mirative* and others.

In the cursory survey of the Siouan data below I will for the most part ig-

nore issues of semantics beyond the gross level of meaning indicated by being translated in a grammar or dictionary as 'and' versus 'but' or 'or' — roughly additive, adversative, and disjunctive. From the data available it is often not clear precisely what range of meanings are covered by a given conjunction. Semantic classification of the conjunctions will require detailed investigation of usage in each individual language, and will surely interact with numerous factors, including modality, adverbial modifiers, same or different subject of conjoined clauses, and so on. I leave this entire area for future research. For the present I simply list all elements which seem to translate 'and,' 'but,' or 'or' in any of their meanings.

2.3 Identifying lexical coordinators

Another issue is that some of these lexical items, although they translate English coordinators, may in fact not be coordinators. This is yet another area which provides fertile ground for future, deeper research into each individual Siouan language. Coordinating conjunctions can be difficult to distinguish from sentence-initial or sentence-final elements (complementizers, discourse particles, switch-reference markers, and other clause-linking morphemes), and from comitative or adverbial words. Coordinators often develop historically into sentence-initial or -final elements, presumably by way of a stage involving elided conjuncts. Historical change can go the other way too: as Mithun (1988) and Stassen (2000) both point out, many languages have coordinating conjunctions which are recently and transparently derived from various sources, including comitative prepositions, adverbial particles, aspect markers, and clausal (subordinating) conjunctions. This leads to situations in which the same word is sometimes a coordinator, sometimes not, and teasing apart the two usages is tricky; such is the case for example with Bulgarian *no*, *ama*, *ami* (Fielder 2008) and Australian English *but* (Mulder & Thompson 2008). Given the slipperiness of this issue in well-studied European languages, it should be no surprise that identifying coordinators can be problematic when dealing with spoken or inconsistently written data in a language with no tradition of written prose or punctuation conventions.

3 Additive coordination in Omaha-Ponca

My interest in coordination in Siouan was sparked not by theoretical considerations but by a practical problem of language teaching. In an Omaha language class in 2002, a student's question of how to say 'and' turned out to be unex-

pectedly hard to answer, with no one word corresponding to English *and*. There are several clause connectors which are at least plausible candidates for coordinators in Omaha-Ponca, but nothing which syntactically coordinates nominal or other non-clausal phrases. To say things like 'I have a cat and two dogs' or 'That dress is black and white' our Omaha-speaking consultants rephrased with non-coordinate constructions, to the sometimes frustrated bewilderment of the English-dominant students. In this section I examine various options for expressing additive coordination ('and') in Omaha-Ponca and consider whether they are true coordination or involve some other strategy such as adverbial modifiers or subordination. This case study illustrates both the richness and complexity of the data and the difficulty of conclusively distinguishing coordination from non-coordinate structures in a Siouan language.

3.1 Coordination of clauses: *shi* and similar words

The word most commonly offered by Omaha consultants as a translation for *and* is *shi*, which often occurs as an apparent sentence conjoiner, or at least a discourse link between sentences. Koontz (1984: 52) lists *shi* along with *ki*, *go*^{n}, *go*^{n}*ki*, *o*^{n}*ska*, and *egithe* in a table of "sentence introducers" culled from James Owen Dorsey's 19th-century Omaha and Ponca materials; the same words are found in my field recordings from 100 years later. It is an open question whether these words start a new sentence or not; i.e. whether the structure is [S *shi* S] or [S][*shi* S], with [*shi* S] constituting a separate sentence.[1] Dorsey apparently considered them to be the start of a new sentence, but it is unclear why. Presumably he heard a preceding pause, or speakers when dictating to him tended to pronounce *shi* with the following sentence. But there is often a pause or break before a conjunction in English as well, sentences do begin with coordinating conjunctions (in spite of prescriptive prohibitions), and in the more recent view of coordination, the conjunction does form a tighter unit with one of the joined clauses. Even if we assume the entire string [S *shi* S] is a single sentence, it is unclear whether the two smaller sentences so joined are syntactically coordinated or one subordinate to the other. Omaha has no clear markers of subordination that I know of (e.g. no nonfinite verb forms).

The precise meaning of *shi* is another issue: Koontz states that *shi* differs from the other "introducers" in that it has a meaning of 'again' or 'marks repetition', but this meaning is not always apparent to me. *Shi* sometimes seems to indicate

[1] David Rood (pc) points out that [S shi] [S] might be a more expected split into two sentences in a verb final language, but shi is not sentence final, in written texts or spoken prosodic contours.

repetition, but not always. In the examples below,[2] *shi* (boldfaced) seems to mark not so much repetition as simple additive coordination semantics — 'and also' — or even contrast, as in (8) or (10). In some discourses *shi* strings together several sentences or clauses in a row, as in (10) and (11). Example (11) in particular is a fairly extended discourse in which nearly every sentence after the first starts with *shi*, and the discourse is a list of items, with no sense of repetition except the continued idea of praying for something. Note that *shi* cooccurs with other "sentence introducers," for example, *go^{n}ki* (in (10) and (11)), and with arguably subordinating adverbial *ki* (in (11)).[3]

(7) *Thíshti xtáwithe.* ***Shi*** *thíshti xtó^{n}thathe.* *éshti xtó^{n}tha=i* *ge* ***shi***
you 1sgA.like.2P and you 2A.like.1sgA s/he like.1P=prox ? and
wíshti xtáathe.
I.too sgA.like.3P
'I like you. And you like me. S/he likes me and I like her/him too.'

(8) *Zhi^{n}gá* *ama águdishti údon wáno^{n}ʔo^{n}=no^{n},* ***shi*** *águdishti*
children the some good listen.to.1plP=hab and some
wáno^{n}ʔo^{n}=bazhi=no^{n}.
listen.to.1plP=neg=hab
'Some of the children are good; they listen to us, but some of them don't listen to us.'

(9) ***Shi*** *gó^{n}ki shaó^{n}* *ama … shaó^{n}* *xé=ta=i* *á=bi=ama.*
and then Sioux the … Sioux bury=fut=prox say=prox=quot
'And the Sioux … His wish was for the Sioux to bury him.'

(10) *Gó^{n}ki* ***shi*** *gá=t^{h}e* *o^{n}gáhi* *ki* ***shi*** *wach^{h}ígagha ama*
then and that=the 1plA.go.there when and dancers the
shó^{n}-gagha=i=t^{h}e *ki* ***shi*** *shó^{n}sho^{n}* ***shi*** *zhuáwagthe*
end-do=prox=evid when and right.away and together

[2] These examples are from my field tapes, recorded in the late 1980s and 1990s, in Macy Nebraska. I am grateful to the National Science Foundation and Wenner-Gren Foundation for support, and to the speakers quoted here, Clifford Wolfe Sr., Bertha Wolfe, Mary Clay, and Coolidge Stabler, for sharing their language with me. The orthography used in this paper is the "Macy Standard" spelling used at Umo^{n}ho^{n} Nation School and the University of Nebraska.

[3] '?' in examples marks words which were unclear when transcribing field tapes or whose meaning is unknown. Since the morphological breakdown of most words is immaterial for the purposes of the paper, glosses are not necessarily morpheme-by-morpheme. Clitics are separated with an equal sign.

agthé=t^{h}a=ama.
took.1PLP.home=EVID=AUX
'We would go there but as soon as the dancers quit they took us right home.'

(11) a. *Wakó^{n}da thi^{n}k^{h}e shti btháha=ta=mi^{n}k^{h}e*
god the too 1SGA.pray=FUT=1SG.AUX
'I'm going to pray to God.'

b. ***Shi*** *gáge i^{n}dádo^{n} thé amá níkashi^{n}ga amá shti ewéwaha=t^{h}e*
and that what this the person the too 1SGA.pray.for.it=EVID
'I (will) pray for the people who had these things.'

c. ***Shi*** *umó^{n}ho^{n} ti tho^{n} shti agíwaho^{n}.*
and Omaha house the too 1SGA.pray.for.it.REFL
'And I (will) pray for my Omaha camp/village.' (i.e. for the present-day reservation)

d. ***Shi*** *t^{h}ó^{n}wo^{n}gtha dúba édi mo^{n}thí^{n} umó^{n}ho^{n} shti ewéwaha.*
and town several there 3A.walk Omaha too 1SGA.pray.for.3P
'And I (will) pray for the Omaha who are in various cities.' (i.e. off reservation)

e. *Gáge* ***shi*** *gahí nikashi^{n}ga.*
this and chief person
'And for the council.'

f. ***Shi*** *uzhó^{n}ge o^{n}gáthe dshto^{n}.*
and road/path LA.go maybe
'And for the path we will take.' (i.e. for our lives)

g. *Awó^{n}ho^{n} egó^{n} é=ta=mi^{n}k^{h}e*
1SGA.pray thus that=FUT=1SG.AUX
'I will pray for those things.'

The other "sentence introducers" listed by Dorsey and Koontz include *ki*, *go^{n}*, *gó^{n}ki*, and *kigó^{n}ki*, all meaning 'and, and then'. Their distribution is similar to that of *shi*; both in Dorsey's texts and in mine, they occur written at the beginning of sentences as well as joining two sentences or clauses, and they indicate a range of connections between those clauses, sometimes temporal and sometimes not.

3.2 Coordination of non-sentential categories: does it exist?

Shi and the other sentence conjoiner/introducers generally do not occur in conjoining contexts other than linking sentences. That is, they appear not to coordi-

nate nominals or other non-clausal categories (though see (35) below). In the case of nominals, several patterns occur, generally consisting of a string of NPs with a word meaning something like 'also' at the end, sometimes with some element between the individual NPs as well.

Koontz (1984: 201) gives the formula NP, NP *étho^{n}ba* for conjoined nominals in Dorsey. This pattern is found in modern materials as well. Example (12) is a sentence from the story *Jimmy and Blackie*, translated into Omaha as a school booklet in the 1980s, and (13) is an example from a conversation I recorded in 1990. *Etho^{n}ba* is etymologically related to the number two (*no^{n}ba*) and probably best treated as an element meaning 'both' or 'the two of them' instead of as a conjunction.

(12) *I^{n}no^{n}ha ak^{h}á, i^{n}dadi* ***étho^{n}ba*** *théthudi gthí^{n} é=shti.*
my.mother the my.father also here live they=too
'My mom and also my dad, they live here too.'

(13) *Ivan ak^{h}á Silas* ***étho^{n}ba*** *ukíkizhi.*
Ivan the Silas also brothers
'Ivan and Silas, those two were brothers.'

Ardis Eschenberg (pc) reports that the elders/language teachers at Umo^{n}ho^{n} Nation school in the early 2000s generally used NP, NP *shti* for conjoined nominals. I have found some examples of this too, but actually very few with this exact pattern. Example (14) is one. Most sentences with *shti* in my data have variations on the pattern such as *shti* after a single NP (15), or repeated *shti* (16), (17). Note that *shti* cooccurs with *shi* in (16) to coordinate three NPs: NP *shti*, NP *shti, shi* NP. In (17) the second conjunct looks like a postverbal afterthought. The word *shti* 'too, also' could perhaps be analyzed as a coordinator, but seems more likely to be an adverbial element, perhaps related to *xti* 'very'.

(14) *Ithádi, ihó^{n} ak^{h}a* ***shti*** *gínita ezhé go^{n}kí ithádi ama,*
his.father his.mother the too living ? and his.father the
ihó^{n} ak^{h}á zhúgigtha=bazhí.
his.mother the together=NEG
'His father and his mother are both alive, but his father and mother do not live together.'

(15) *Tim ak^{h}á iwíko^{n}=ta=ak^{h}a Clifford* ***shti*** *utháha*
Tim the 3A.help.1SGP=FUT=3AUX Clifford too ?

uwíko^{n}=ta-ak^{h}a
3A.help.3P=FUT=3AUX
'Tim will help me and Clifford.'

(16) *Shi níkashi^{n}ga húto^{n}ga wa'ú* ***shti*** *shaó^{n}* ***shti shi*** *wáxe dúba edí*
and person winnebago woman too sioux too and white some there
at^{h}í-ama.
3A.arrive-PL.AUX
'And a Winnebago woman, some Sioux, and some whites were also there.'

(17) *Shi wóndo^{n} ithádi* ***shti*** *hó^{n}di ugíkitha ihó^{n}*
and both his.father too last.night 3A.was.talking.to.3P his.mother
ak^{h}á ***shti****.*
the too
'And last night he was talking to both his father and his mother.'

In my elicited data conjoined nominals most often take the form NP (*ego^{n}*), NP *sheno^{n}*, with degree elements literally meaning 'so much' or 'that much' as in examples (18) through (24). The awkward literal gloss with 'as … that extent' could perhaps be better rendered 'as well as … all of those'. In any case, this seems unlikely to be a coordinate construction.

(18) *Téska tanúka* ***égo^{n}*** *wazhí^{n}ga* ***égo^{n}*** *nú* ***shéno^{n}*** *that^{h}é xtáathe.*
cow meat as chicken as potato that.extent eat 1SGA.like
'I like to eat beef and chicken and potatoes.'

(19) *Wat^{h}é zhíde* ***égo^{n}*** *hi^{n}bé ská* ***shéno^{n}*** *bthíwi^{n}.*
dress red as shoe white that.extent 1SGA.buy
I bought a red dress and white shoes.'

(20) *Wat^{h}é zhíde, hi^{n}bé ská, watháde pézhitu* ***shéno^{n}*** *abthi^{n}.*
dress red shoe white hat green that.extent 1SGA.have
'I have a red dress, white shoes, and a green hat.'

(21) *Sézi t^{h}e shé* ***shéno^{n}*** *áhige o^{n}gáthi^{n}.*
orange the apple that.extent much 1PLA.have
'We have plenty of (both) oranges and apples.'

(22) *Mary ak^{h}á* ***égo^{n}*** *wi* ***shéno^{n}*** *Macy ata o^{n}gátha.*
Mary the as I that.extent Macy to 1PLA.go.there
'Mary and I went to Macy.'

(23) *John ak^{h}á* ***égo^{n}*** *Mary ak^{h}á* ***shéno^{n}*** *Macy ata ahí=t^{h}e.*
John the as Mary the that.extent Macy to 3PLA.arrive.there=EVID
'John and Mary went to Macy.'

(24) *Tim akʰá Cliffford* ***égoⁿ*** *wi* ***shénoⁿ*** *iwíkoⁿ=ta=akʰa.*
Tim the Clifford as I that.extent help.1SGP=FUT=3AUX
'Tim will help Clifford and me.'

This NP *égoⁿ*, NP *shénoⁿ* pattern also occurs in bilingual booklets produced by the Umoⁿhoⁿ Nation school; the translations are from the booklets as well:

(25) *Jimmy akʰá* ***égoⁿ*** *Sabe akʰá* ***shénoⁿ***
Jimmy the as black the that.extent
'Jimmy and Blackie' (title of booklet)

(26) *Núzhiⁿga ga tʰoⁿ é=****egoⁿ*** *mízhiⁿga ga tʰoⁿ e=shti* ***shénoⁿ***
boy this the he=as girl this the she=too that.extent
uwáwakizhi. my.younger.siblings
'This is my little brother and sister.'

A more literal translation of (26) would be 'Like this boy, this girl also, as a group they are my little siblings.' Another pattern combines the previous two: NP *égoⁿ*, NP *shti*; (27) is an elicited example from my field tapes, (28) a spontaneously produced sentence.

(27) *Mary akʰá* ***égoⁿ*** *wí=****shti*** *Macy ata oⁿgátha.*
Mary the as I=too Macy to 1PLA.go.there
'Mary and I went to Macy.'

(28) *Ihóⁿ wiáxchi* ***égoⁿ*** *ithádi* ***shti*** *wiáxchi.*
their.mother just.one so their.father too just.one
'They have the same mother and the same father too.'

Simply juxtaposing a string of nominals is another coordination strategy, and quite a common one, though I will not give any examples. In fact, all of the nominal coordination patterns we have seen so far could be interpreted as simple listing of noun phrases, with some kind of focus element following one or more of the nominals and/or a summing-up element at the end of the nominal string. Given the lack of case marking and near-absence of number agreement in Omaha-Ponca,[4] as well as the likely status of most if not all lexical noun phrases as adjuncts in this language, the usual tests for coordinate as opposed to other structures tend not to apply, and it is difficult to distinguish for example coordinate from comitative constructions.

[4] Third person plural is not audibly marked in many verbs, and in those where it is, it is homophonous with proximate singular marking.

A final, very common way of expressing English 'and' in situations involving two participants acting together is with the verb *zhugthe* 'be with, accompany, be together'. This verb sometimes occurs following two nouns which could be seen as coordinated but are probably just listed; (29) is more literally 'Mary, John, being together they went to Macy.'

(29) *Mary ak^{h}á John Macy ata* ***zhúgthe*** *ahí.*
Mary the John Macy to together arrive.there
'Mary went to Macy with John. /Mary and John went to Macy.'

In non-elicited examples, there is almost never more than one lexical noun phrase with *zhugthe*; instead one nominal is given and the other is understood as accompanying it. In (30) only the woman is mentioned; the other participant is already present in the discourse. In (31) the unmentioned participant is the speaker, and interestingly the verb is first person singular, not plural, indicating that the construction is definitely comitative and not coordination of an overt with a null NP.[5]

(30) *Agthí (i)t^{h}ediki shi wa'ú shtewi^{n}* ***zhúgthe***
3A.came.home when and woman whatsoever together
agthí=it^{h}e.
came.home=EVID
'When he came home, he came home with a woman.' (He and some woman came home.)

(31) *Wa'ú wiwíta Tésó^{n}wi^{n}* ***zhuágithe*** *at^{h}í, she=k^{h}e*
woman my White.Buffalo together.1SUUS 1SGA.arrive this=the
o^{n}gát^{h}i.
1PLA.arrive
'My wife White Buffalo and I are both here; we came here.' (more literally, 'My wife White Buffalo, together with my own, I came here…')

There are thus several ways of expressing semantic coordination of nominals in Omaha-Ponca, but none for which a strong case can be made that it is a syntactic coordinate construction or any clear candidate for a coordinating construction. The nominal "coordination" patterns above are all basically lists of NPs with the option of adding a word or words stressing repetition or accompaniment. The picture is even more dubious for adverbs, nominal modifiers, and

[5] In playback speakers commented that the second verb, *at^{h}í*, could have been *o^{n}gát^{h}i* (first person plural), like the verb of the next clause; *zhuágithe* however would still be first person singular.

other non-clausal constituent types. Koontz (1984) does not mention conjunction of categories other than nominals. I did not think to elicit them in field work, and have not found naturally produced examples. The kind of sentences my Omaha-language-class students wanted to say, like 'I'm wearing a red and yellow shirt,' seem impossible to express without resorting to multiple clauses ('My shirt is red and it is also yellow.')

3.3 Discussion: Once more on *shi*

Having concluded that Omaha-Ponca has no clear coordinating conjunction or coordination construction for non-clausal coordination, I return briefly to my best candidate for clausal coordinating conjunction, *shi*. In §2.1 I presented a number of examples of *shi* apparently linking clauses together; however, it may actually be an adverbial of some sort, not a conjunction, in which case Omaha-Ponca would not have any true coordination, even of clauses. It often appears in positions other than clause-initial, most often preverbal, as in the following examples. Here it is clearly not conjoining anything, but does have an 'again' sense:

(32) *ó^{n}ba wéthabthi^{n} ki* ***shi*** *wat'éxe=ta=ama.*
day third at and funeral=FUT=AUX
'There'll be another funeral Wednesday.' (Wednesday again will be a funeral.)

(33) *O^{n}wó^{n}that^{h}o^{n} thísht^{h}o^{n}=i t^{h}edi tápuska ta* ***shi*** *háthe o^{n}gák^{h}i.*
1PLA.eat finish=PROX when school to and ? 1PLA.arrive.back
'After dinner we went back (again) to the school.'

(34) *ó^{n}ba wi^{n} Ishtí^{n}thi^{n}khe ak^{h}á* ***shi*** *edí=bi=ama.*
day one Monkey the and there=PL=QUOT
'One day Monkey was there (again), they say.' (traditional story opening)

However, it is possible that this is a different *shi* from the sentence-coordinating one. Further research is obviously needed. My data contain a few examples in which *shi* might be interpreted as conjoining nominal phrases, following the last in a string of NPs: NP, NP *shi*. None are very convincing, however, and *shi* in them can plausibly be taken as an adverbial expressing repetition. In (35), for instance, fighting was a regular occurrence.

(35) *Umó^{n}ho^{n} k^{h}e shaó^{n} k^{h}e* ***shi*** *wó^{n}do^{n} kikína=no^{n}=i*
Omaha the Sioux the and both 3A.REFL.fight=HAB=PROX
'The Omaha and the Sioux tribes used to fight each other.'

This section thus concludes rather inconclusively: Omaha-Ponca apparently has no coordination of non-clausal constituents, and may or may not have true coordination of clauses.

4 Siouan languages: An overview

At this point we leave the details of Omaha-Ponca and turn to a shallow but broad survey of the Siouan family. In spite of limited data on many members of the family and the challenges of interpretation and analysis, there is quite a lot we can say about coordination in Siouan languages. In several of the languages coordination has been described in some detail. Nearly all of the languages have recorded equivalents of 'and,' and many have equivalents for 'or' or 'but,' though their morpho-syntactic status is often unclear. In many of the languages coordination of clauses is different than coordination of noun phrases or other categories, as we saw in Omaha-Ponca. Perhaps the most interesting result of a survey of Siouan coordination is the lack of unity within the family. No coordinators are reconstructable, there are no widespread cognates, and strategies for expressing coordination differ from language to language. It appears likely that Proto-Siouan had no true coordination. In this section I briefly describe the data from each sub-branch of Siouan (starting with Dhegiha because it is most familiar to me; information on Omaha-Ponca is repeated in brief form for completeness). No examples are given in this section and no attempt is made to justify the lexical items given as (possible) coordinators; instead, anything mentioned in sources is listed.

4.1 Dhegiha

Omaha-Ponca (data from Dorsey n.d.(a) Dorsey n.d.(b), Koontz 1984, Rudin 2003 and my own fieldwork)[6] has several ways of expressing 'and'. As discussed above, different conjunctions are used to coordinate clauses and NPs. Clauses may be conjoined with *ki*, *go^n*, *shi* 'again, and then,' *go^{n}ki*, *kigo^{n}ki* 'and then'. Dorsey considers *ki* to be Ponca and *go^n* to be Omaha; both of these are said to join "substantive clauses". *Go^n* is likely the same as subordinating *(e)go^n* 'having (done),' related to postposition *ego^n* 'like, as'. *Go^{n}ki* and *kigo^{n}ki* are pretty clearly combinations of these two conjunctions. *Shi* is perhaps the best candidate for a

[6] These sources use several different orthographies. In the interest of consistency I have spelled all Omaha-Ponca words in the modern "Macy Standard" spelling.

true coordinator, although it, like the others listed here, occurs most often sentence initially (conjoining the sentence to the preceding discourse semantically if not syntactically). NPs are occasionally joined by *go^{n}*; this may actually be the postposition mentioned above. More commonly two NPs are followed by *edo^{n}ba/étho^{n}ba* 'also, both;' literally 'the two of them'. A string of three or more NPs may be followed by *edabe* 'also'. Two or more NPs can be followed by *shti* 'too'. Although Dorsey does not list it, one of the most common strategies for coordinating NPs in my data is *ego^{n} … sheno^{n}* 'both … and;' literally 'as … that-extent', Ardis Eschenberg (p.c.) finds *ego^{n} … tho^{n}zho^{n}* used in the same way. The most common translation of 'and' with NPs is clearly not syntactic coordination: a comitative construction with the verb *zhugthe* 'be with'. Simple juxtaposition (listing) of conjuncts with no conjunction is common for both S and NP coordination. 'Or' and 'but' in Omaha-Ponca are formed with the 'and' conjunctions for joining clauses, and to the best of my knowledge do not exist at all for NPs. Dorsey lists *go^{n} … ite ki* 'either … or', *sho^{n} do^{n}ste* 'either-or, perhaps', and *do^{n}ste* at end of clause 'or' (the latter two in Dorsey's slip file). 'But' is commonly expressed by *shi* 'and' connecting two clauses, the second of which is negative or contrasts in some way.

Osage (data from Quintero 2004) coordinates NPs using *éeðǫǫpa* 'the two of them' following two or more NPs. (Compare Omaha-Ponca *etho^{n}ba*.) Verb agreement suggests that this is a true coordination structure; however, it is possible that the two NPs are appositive and the plural verb actually agrees with pronominal *ée-*. Another possible NP coordinator is *ški* 'also'. Clauses are coordinated by juxtaposition without a conjunction: "There is no Osage equivalent to the English use of *and* to conjoin sentences; rather, the elements are strung together with no intervening forms of any kind" (455). Quintero gives no information on 'or' or 'but'.

Kaw (Kansa) (data from Cumberland & Rankin 2012; Justin McBride p.c.; Robert Rankin p.c.) has an 'and' coordinator, *ši*, which is used in a variety of syntactic environments (postverbal, preverbal, postnominal, clause-initial) and apparently can conjoin both clauses and nominals. McBride states that it usually seems to be used adverbially ('again') or adjectivally ('another'), but can also symmetrically coordinate clauses. Numerous conjunctions with meanings like 'and, then, so' exist, but all seem to be subordinating rather than coordinating. The conjunction *dą* 'and, then, so' occurs between clauses and in other coordinating situations; Rankin, in a 2012 email, states that "Kaw … seems to allow the conjunction *dą* (often reduced to d-schwa …) in exactly the same places English would allow 'and"'; he suggests this is a result of adopting Spanish coordination

structures. Further evidence of Spanish influence is the clearly borrowed coordinator *pero* 'but'. I have no information on 'or' in Kaw.

Quapaw (data from Rankin 2002; 2005) probably has conjunctions similar to those in the other Dhegiha languages, but I have very little information. Rankin's grammar and dictionary list *şi* 'and' (cf. Omaha-Ponca *shi*, Kaw *ši*), but give no indication of how it is used.

4.2 Winnebago-Chiwere

Ho-Chunk (data from Helmbrecht 2004; confirmed by Iren Hartmann p.c.) has three apparently straightforward coordinating conjunctions, which Helmbrecht labels as follows: *ánąga* 'and' (coordinate); *nįįgéšge* (*nįgeešge*) 'or' (disjunction); *nųnįge* 'but' (adversative). The 'and' and 'or' words are used to conjoin all types of syntactic constituents: NP, VP, S, "obliques" (adjunct phrases), and AdvP. The conjunctions are placed between the coordinated phrases, or in the case of three coordinated NPs, preceding the last NP (X Y *ánąga* Z 'X, Y and Z'). Helmbrecht argues that *ánąga* conjunction is true coordination: the resulting constituent requires plural agreement, and an overt pronoun is needed to conjoin a 1st or 2nd person. Ho-Chunk also has a comitative construction with the verb *hakižu* 'to be together,' as well as some other, presumably subordinating conjunctions: *nąga, hireanąga* 'along with' conjoins animate subjects or objects, and clauses can be conjoined with *'eegi* 'and then' or *šge/hišge* 'also, even' (placed after 2nd conjunct). Helmbrecht also discusses negation of one or both conjuncts; a special conjunction *hąké*, used at the beginning of S or NP, expresses 'and not/but not'.

Chiwere (data from Goodtracks 1992 – present; Greer 2016 (this volume); Bryan Gordon, p.c.) has several ways of expressing 'and'. These include words meaning 'with' (*tógre, insú^{n}, inú^{n}ki*), 'also' (*heda^{n}, -da^{n}, na, -ku*), 'again' (*šige*), and a set of discourse connectives in the form of clefts, with copula *aré: aréda, edá, arédare, édare, hédare*. In addition, a string of nominals can be followed by *inu^{n}ki* or *bróge*. Gordon also lists 'bracketing' conjunctions: *šu^{n}, gasú^{n}, nahéšu^{n}*, and a number of subordinating connectives. 'But' is *núna*.

4.3 Dakotan

There is information available on several of the Dakotan languages and dialects; some sources include data from more than one dialect. I have found no information on Stoney.

Assiniboine (data from West 2003; Cumberland 2005; Levin 1964) has two main 'and' coordinators, *hīk* and *hīkná*, but sources differ somewhat in their

descriptions of how these are used. West argues explicitly that *hĩkná* conjoins VP or V, not clauses; i.e. it occurs in the context VP *hĩkná* VP or V *hĩkná* V. She analyzes it as head of a CoordP with the first conjunct VP/V as complement and the second one as specifier (pp. 32-38). Clauses are joined by *hĩk* repeated after each clause: S *hĩk* S *hĩk*. Cumberland, on the other hand, shows all categories joined by non-repeating *hĩk*: NP *hĩk* NP, V *hĩk* V, VP *hĩk* VP. Levin (1964), cited in Stassen (2000: 36) discusses a third coordinator, *ka*, which conjoins NP. There is also a comitative construction with *kici* 'with' at the end of a string of NPs. I have no information on 'but' or 'or' in Assiniboine.

Lakota[7] (data from Rood & Taylor 1996; Ingham 2003; Ullrich 2016; Boas & Deloria 1941) has several 'and' conjunctions: *na, naháŋ* 'and also'; *čha, čhaŋkhé* 'and so'; *yuŋkȟáŋ* 'and then'; *na* can coordinate nouns or clauses, while the others appear to coordinate only clauses. Lakota also has a word meaning 'or': *naíŋš*, and several expressing contrastive coordination 'but': *éyaš, k'éyaš, tkȟá, khéš, škȟá. éyaš* is also listed as an interjection meaning 'well, but'. Numerous other conjunctions are listed, including *ho, honá* 'furthermore', *nakúŋ* 'also', *hé uŋ* 'therefore', *tkȟáš* 'but indeed', and others. Ulrich gives examples of an apparent comitative, *kičhí*, as well. It is not entirely clear to me whether the 'and/or/but' conjunctions are all coordinators or whether some (or all) are subordinating conjunctions. Rood and Taylor define "conjunction" as connecting two sentences, but at least the 'and' and 'or' words can also conjoin "parts of a sentence, such as nominals or verbs". The position of all the conjunctions is between conjuncts in their examples, but they state there are "two possible positions: in the second slot from the beginning or in the last slot in the sentence." David Rood (p.c.) points out that obligatory ablaut before *na* and *naíŋš* suggests a strong bond between the conjunction and the preceding verb.

Dakota (data from Riggs 1851; Boas & Deloria 1941) has unsurprisingly some similar conjunctions to Lakota, though some also differ. Several words translate 'and': *k'a, čha, uŋkháŋ, nakúŋ. Uŋkháŋ* conjoins clauses with different subjects, while *k'a* conjoins nouns and clauses with same subject; no details are given of the usage of the other 'and' words. 'But' is *tukhá*, and 'or' is *k'a iš*. Boas & Deloria give forms from several dialects; alongside the Lakota forms in the previous paragraph they also list Dakota forms, usually labelled as "Yankton" and/or "Santee" dialect, including *k'a* 'and, *uŋkháŋ* 'and then'.

[7] In general I have used the orthography of the source in this paper. However, in the case of Lakota, I have standardized all the disparate orthographies of the various sources to the modern standard spelling system used by the Lakota Language Consortium.

4.4 Missouri Valley

Crow has very different strategies for conjoining clauses and nominals (data from Graczyk 2007). For coordinate nominals, the conjunctions are *-dak* 'and' and *-xxo* 'or'. Both are suffixes (or enclitics), but at different levels: *-dak* suffixes to NP, while *-xxo* suffixes to N′. Both conjunctions are repeated after each conjunct; *-dak* may and *-xxo* must be omitted after the final conjunct. There is also a comitative construction involving the transitive verb *áxpa* 'be with' (also 'marry') with same-subject marking or an incorporation structure. Clauses in Crow are linked by switch-reference marking rather than conjunction. Graczyk analyzes apparently coordinate clauses as 'co-subordination' or clause-chaining: a string of clauses with switch-reference markers but no sentence final clitic, except for the last clause, which determines the speech-act type of the entire string (eg. declarative). The adversative 'but' relation between clauses is marked with *-htaa* (suffix on clause) or *hehtaa* (sentence connector).

Hidatsa (data from Boyle 2005; 2007; 2011) has significantly changed its coordination constructions in quite recent times. Boyle points out that Crow and Hidatsa share some cognate morphology in the area of conjunctions (e.g. Hidatsa *-k* is cognate with Crow *-dak*), but Hidatsa has innovated a semantic distinction involving specificity and inclusiveness of NPs. In the area of clausal/verbal coordination, Hidatsa's former switch-reference markers have evolved into English-like coordinators (Boyle 2011). At present, the following morphemes express 'and': *hii* coordinates S's; *-k* coordinates NP (with a nonspecific reading when suffixed to both NPs and a specific reading when suffixed only to the first NP); *-šek* coordinates NPs with a non-specific reading; *-a* coordinates V in serial verb construction; *-ak* (the old Same Subject marker) coordinates V or VP. There is apparently no 'but' coordinator; adversative meaning "is shown with juxtaposition with one element being negated" (John Boyle p.c.).

Mandan (data from Clarkson 2012; Randolph Graczyk p.c.) links clauses via a switch reference system similar to that of Crow. The morpheme *ni* is used both as a same-subject marker for clauses and as a NP coordinator. NP coordination is accomplished with a coordinator following each NP; coordinating conjunctions used in this way include *eheni, -kini, -hini, -kiri*, all meaning 'and'. In modern usage two new coordinators appear, not found in older texts: *hi(i)* with NPs and *ush* with clauses. Both occur between conjuncts rather than after each conjunct. Clarkson claims that coordination is much more common in recent texts than in those from the early 20th century, suggesting that Mandan syntax, like that of Hidatsa, has been restructured under pressure from English. I have no information about alternative or adversative coordination in Mandan.

4.5 Southeastern Siouan

Biloxi (data from Zenes 2009; based on Dorsey & Swanton 1912) has an NP coordinator *yą* 'and' which suffixes either to each NP or just the last one; it is also possible for NPs simply to be listed. Clauses are coordinated by simple juxtaposition. Zenes treats the latter two constructions (concatenated NPs and S's) as CoordP with a zero coordinator. Coordination of a series of object NPs is expressed by coordinating clauses with the same verb repeated ('I planted onions, I planted potatoes, I planted turnips'). Disjunction of NPs is expressed by *ha* 'or' following the second NP. Zenes gives no information about 'or' with sentences or clauses. Biloxi also has a comitative construction with *nǫpa* following the second NP.

Ofo (Dorsey & Swanton 1912; Robert Rankin, p.c.) apparently coordinates clauses only by juxtaposition with no conjunction. I have no further information about Ofo coordination, and none at all about Tutelo.

4.6 Summary

The known possibly-coordinating conjunctions of the Siouan languages are summarized in Tables 1 and 2 To give some sense of their syntax, the conjunctions are shown with the type of constituents they conjoin when this is known; for instance S ***ki*** S means *ki* can occur between two clauses; NP NP ***shti*** means *shti* occurs at the end of a string of NPs.

A partial list of comitative ('with') subordinators is given in Table 3. Presumably the other Siouan languages also have comitative constructions; I list here only those which were mentioned in one of my sources as a common way to express 'and' coordination.

5 Conclusion

What can we learn from the array of facts above? The most striking conclusion that emerges from the data is the lack of unity among the Siouan languages. Even within subfamilies, the Siouan languages are quite diverse in their treatment of coordination. We can identify several areas of disagreement: (1) The languages differ in the types of constituents that can be coordinated, some having only clausal coordination, while others can coordinate NPs and other types of constituents as well, and some may have no true coordination at all, but use various types of subordination, co-subordination, or simple concatenation to express the relations English expresses with 'and'/'or'/'but'. (2) They differ in the constituent

Table 1: Coordinating(?) conjunctions

Language	Additive ***and***	Disjunctive ***or***	Adversative ***but***
Omaha-Ponca	S ***ki*** S; NP ***ki*** NP S ***go^n*** S; NP ***go^n*** NP S ***shi*** S NP NP ***edo^{n}ba***/ ***étho^{n}ba*** NP NP NP ***edabe*** NP ***ego^n*** NP ***sheno^n*** NP NP ***shti***	***go^n*** S ***ite ki*** 'either ... or' S ***dshto^n shi*** S ***dshto^n*** 'maybe and maybe' S ***sho^n*** S ***do^{n}ste*** 'either-or, perhaps' S ***do^{n}ste*** 'or'	S ***shi*** S-NEG
Osage	NP NP ***éeðǫǫpa***		
Kaw	S ***ši*** S S ***dą*** S		***pero***
Quapaw	***çi***		
Ho-Chunk	S ***ánąga*** S also conjoins NP, VP, AdvP, oblique	S ***nįį́géšge*** S also conjoins NP, VP, AdvP, oblique	***nųnįge***
Chiwere	***šige*** ***heda^n***, ***-da^n*** NP NP ***inu^{n}ki*** NP NP ***bróge***		***núna***
Assiniboine	V ***hīkná*** V, VP ***hīkná*** VP S ***hīk*** S; also with NP, V, VP, etc. S ***hīk*** S ***hīk***; also with NP, V, VP, etc. NP ***ka*** NP		
Lakota	S ***na*** S, NP ***na*** NP, V ***na*** V S ***yuŋkȟáŋ*** S S ***čha*** S S ***čhaŋkhé*** S	S ***naíŋš*** S	S ***éyaš*** S, NP ***éyaš*** NP, V ***éyaš*** V ***k'éyaš*** ***tkȟá*** ***khéš*** ***škȟá***
Dakota	S ***k'a*** S, NP ***k'a*** NP S ***uŋkȟáŋ*** S ***nakúŋ*** ***čha***	NP ***naíŋš*** NP ***k'a iš***	***tukhá***

Table 2: Coordinating(?) conjunctions continued

Language	Additive ***and***	Disjunctive ***or***	Adversative ***but***
Crow	NP ***dak*** NP ***dak***	N′ ***xxo*** N′ ***xxo***	
Hidatsa	S ***hii*** S NP-***k***; NP-***k*** NP-***k*** NP -***šek*** NP V-a V (serial verb) V-***ak*** V; VP-***ak*** VP		juxtaposition with negation
Mandan	S-***ni*** S ***ush*** S ***ush*** S NP ***eheni*** NP (***eheni***) NP-***kini*** NP-***hini*** NP-***kiri*** NP(-***kiri***) NP ***hii*** NP		
Biloxi	NP NP ***yą***; NP ***yą*** NP ***yą***	NP NP ***ha***	
Ofo	—		
Tutelo	—		

Table 3: Comitative words

Omaha-Ponca	*zhugthe*
Chiwere	*tógre*, *inú*^{n}, *inú*^{n}*ki*
Assiniboine	*kici*
Lakota	*kičhi*
Crow	*áxpa*
Biloxi	*nǫpa*

order within coordination constructions, with the conjunction following the first conjunct (XP & XP), the second conjunct (XP XP &) or each of the conjuncts (XP & XP &), and may also differ in whether the conjunction forms a constituent with a following or preceding conjunct ((XP &) XP); (XP (&XP)). The hierarchical structure of each of these configurations has not been studied in most of the languages. Given the generally head-final nature of phrase structure in Siouan languages, if the conjunction heads a coordination phrase it is expected that the complement of the "&" head would be to its left; an XP occurring to the right could be a specifier, which we would expect to be less closely associated with the conjunction than the complement. (3) They differ in the lexical items expressing additive, disjunctive, and adversative coordination. Some of the words or suffixes for 'and'/'or'/'but' are cognate among subfamilies — for instance, most of the Dhegiha branch have [ši] or something similar, and the Dakotan branch share something like [na]. But no coordinators appear to be cognate across the family. (4) Finally, the languages differ also in the expression of comitative and other "semantically coordinated" phrases.

In short, there does not seem to be a "typical Siouan" coordination pattern, nor does it look like we can reconstruct proto-Siouan coordinators. Clearly there has been innovation in at least some of the languages – perhaps all – and at least in one or two cases there has been borrowing of coordinators and/or coordination patterns from European languages, suggesting quite recent change in this semantic field. In at least some languages the most common way to conjoin NPs is with a comitative, not a coordinate construction. (This is my impression in Omaha-Ponca, and Cumberland (p.c.) has the same impression in Assiniboine, for example.) Is it possible there was no morphosyntactic coordination in proto-Siouan?

In fact, this is not as unlikely as it might first appear. Mithun (1988) suggests overt coordination tends to come with literacy: in spoken language simple concatenation tends to be common, while in writing, where intonational cues are lacking and one cannot assume the same degree of common knowledge with one's audience, explicit morphosyntactic coordination is more useful. It is certainly not the case that unwritten languages never have true coordination, but as a statistical tendency it makes some sense. Many languages, Mithun says, seem to have developed coordinating conjunctions after exposure to written languages or after developing an indigenous tradition of writing. Since Siouan languages were, until recently, not written, perhaps lack of an inherited coordination construction and associated morphology is not surprising. The borrowing or innovation of coordinators as speakers became literate in English or other European

languages (as well as perhaps in the Native languages) seems logical under this view.

In spite of the lack of overt morphological or lexical coordinators in some languages, Mithun considers coordination as a syntactic and semantic structure to be universal. Stassen (2000), on the other hand, claims coordination, or at least nominal coordination, is not universal. He divides languages into two types: "WITH-languages," which have only a comitative (NP with NP) or subordinating strategy for conjoining NPs , and "AND-Languages," which also have a coordinate strategy. Stassen acknowledges that Native American languages tend to be problematic and difficult to classify into his two categories. This preliminary study of the Siouan family certainly bears out the elusiveness of coordination constructions in these languages.

Acknowledgment

Earlier versions of much of this material were presented at the Comparative Siouan Grammar Workshop / Siouan and Caddoan Languages Conference, Lincoln, NE 2009, and at the 2015 SSILA meeting in Portland. I would like to thank the participants at both events, as well as the two reviewers for this volume, for helpful comments and data. As was so often true, at so many meetings, Bob Rankin's insightful comments at and after the 2009 workshop were especially valuable, and data he gathered was a major source of information on Dhegiha languages. I am grateful to have known Professor Rankin and hope that this paper in some small way contributes to his legacy.

Abbreviations

1, 2, 3	first, second, third person	NEG	negative
A	agent	P	patient
AUX	auxiliary	PL	plural
EVID	evidential	PROX	proximate
FUT	future	QUOT	quotative
HAB	habitual	REFL	reflexive
		SUUS	suus (reflexive possessive)

References

Boas, Franz & Ella Deloria. 1941. *Dakota grammar* (Memoirs of the National Academy of Sciences 23). Washington: Government Printing Office.

Boyle, John P. 2005. The semantics of Hidatsa NP coordination. In Adam Makkai, William J. Sullivan & Arle R. Lommel (eds.), *Interconnections* (LACUS Forum 31), 105–111. Houston.

Boyle, John P. 2007. *Hidatsa morphosyntax and clause structure*. Chicago: University of Chicago (Doctoral dissertation).

Boyle, John P. 2011. Switch-reference in Hidatsa: The evolution of a grammatical system. *International Journal of American Linguistics* 77(3). 413–428.

Brown, Lea & Matthew S. Dryer. 2009. The verbs for 'and' in Walman, a Torricelli language of Papua New Guinea. *Language* 84(3). 528–565.

Carstens, Vicki. 2002. Antisymmetry and word order in serial constructions. *Language* 78(1). 3–50.

Citko, Barbara. 2011. *Symmetry in syntax: Merge, move, and labels*. Cambridge: Cambridge University Press.

Clarkson, John. 2012. *NP coordination in Mandan*. Paper presented at the Siouan and Caddoan Languages Conference, White Cloud, KS.

Cumberland, Linda A. 2005. *A grammar of Assiniboine: A Siouan language of the northern Plains*. Bloomington: Indiana University (Doctoral dissertation).

Cumberland, Linda A. & Robert L. Rankin. 2012. *Kaá^{n}ze íe wayáje: An annotated dictionary of Kaw (Kanza)*. Kaw City, OK: Kaw Nation.

Dorsey, James Owen. n.d.(a). *A grammar and dictionary of the Ponka language*. Map-folder 217, MS 4800, National Anthropological Archives. Washington: Smithsonian Institution.

Dorsey, James Owen. n.d.(b). *Ȼegiha slip file (a dictionary of Omaha and Ponca)*. Box 13 items 133 and 144, Box 14 item 158, Boxes 17–19 item 186, Boxes 20–30 item 188, MS 4800, National Anthropological Archives, MS 4800. Washington: Smithsonian Institution. http://omahalanguage.unl.edu/images.php.

Dorsey, James Owen & John Reed Swanton. 1912. *A dictionary of the Biloxi and Ofo languages, accompanied with 31 Biloxi texts and numerous Biloxi phrases* (Bureau of American Ethnology Bulletin 47). Washington: Government Printing Office.

Fabricius-Hansen, Cathrine & Wiebke Ramm (eds.). 2008. *'subordination' versus 'coordination' in sentence and text: A cross-linguistic perspective*. Amsterdam: John Benjamins.

Fielder, Grace E. 2008. Bulgarian adversative connectives: Conjunctions or discourse markers? In Ritva Laury (ed.), *Crosslinguistic studies of clause combining*, 79–98. Amsterdam: John Benjamins.

Goodtracks, Jimm G. (ed.). 1992 – present. *Baxoje, Jiwere-Nyut^aji, Ma^unke Ich^e Wawagaxe: Ioway, Otoe-Missouria language to English dictionary*. Boulder, CO: Center for the Study of Native Languages of the Plains & Southwest. http://iowayotoelang.nativeweb.org.

Graczyk, Randolph Jan. 2007. *A grammar of Crow*. Lincoln: University of Nebraska Press.

Greer, Jill D. 2016. Baxoje-Jiwere grammar sketch. In Catherine Rudin & Bryan James Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 367–393. Berlin: Language Science Press.

Haspelmath, Martin. 2007. Coordination. In Timothy Shopen (ed.), *Complex constructions*, 2nd edn., vol. 2 (Language Typology and Syntactic Description), 1–51. Cambridge: Cambridge University Press.

Helmbrecht, Johannes. 2004. *Coordination in Hocąk (Winnebago)*. Paper presented at Sprachwissenschaftliches Doktorandenkolloquium, Universität Erfurt.

Ingham, Bruce. 2003. *Lakota* (Languages of the World / Materials 426). München: Lincom.

Johannessen, Janne Bondi. 1998. *Coordination*. Oxford: Oxford University Press.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Koontz, John E. 1984. *Preliminary sketch of the Omaha-Ponka language*. Boulder: University of Colorado (Doctoral dissertation).

Larson, Bradley. 2014. Russian comitatives and the ambiguity of adjunction. *Journal of Slavic Linguistics* 22(1). 11–50.

Levin, Norman Balfour. 1964. *The Assiniboine language*. Berlin: de Gruyter Mouton.

McNally, Louise. 1993. Comitative coordination: A case study in group formation. *Natural Language and Linguistic Theory* 11(2). 347–379.

Mithun, Marianne. 1988. The grammaticization of coordination. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 331–359. Amsterdam: John Benjamins.

Mulder, Jean & Sandra A. Thompson. 2008. The grammaticization of *but* as a final particle in English conversation. In Ritva Laury (ed.), *Crosslinguistic studies of clause combining*, 179–204. Amsterdam: John Benjamins.

Munn, Alan. 1993. *Topics in the syntax and semantics of coordinate structures*. College Park: University of Maryland (Doctoral dissertation).

Quintero, Carolyn. 2004. *Osage grammar* (Studies in the Anthropology of North American Indians). Lincoln: University of Nebraska Press.
Rankin, Robert L. 2002. *Quapaw dictionary*. Unpublished manuscript. Lawrence: University of Kansas.
Rankin, Robert L. 2005. Quapaw. In Heather Hardy & Janine Scancarelli (eds.), *The native languages of the southeastern United States*, 454–498. Lincoln: University of Nebraska Press.
Riggs, Rev. Stephen Return. 1851. *Grammar and dictionary of the Dakota language*. Vol. 4 (Smithsonian Contributions to Knowledge 40). Washington: Smithsonian Institution.
Rood, David S. & Allan R. Taylor. 1996. Sketch of Lakhota, a Siouan language. In William C. Sturtevant & Ives Goddard (eds.), *Languages* (Handbook of North American Indians 17), 440–482. Washington: Smithsonian Institution.
Ross, John Robert. 1968 (1967). *Constraints on variables in syntax*. Bloomington: Indiana University Linguistics Club.
Rudin, Catherine. 2003. *Phrasal conjunction in Omaha-Ponca*. Paper presented at the annual meeting of the Society for the Study of Indigenous Languages of the Americas, Atlanta.
Stassen, Leon. 2000. AND-languages and WITH-languages. *Linguistic Typology* 4(1). 1–54.
Ullrich, Jan F. 2016. *Lakota language handbook: A pedagogical grammar for self-study and practice*. Bloomington, IN: Lakota Language Consortium.
van Oirsouw, Robert R. 1987. *The syntax of coordination* (Croom Helm Linguistics Series). London: Routledge.
Wälchli, Bernhard. 2005. *Co-compounds and natural coordination*. Oxford: Oxford University Press.
Wesche, Birgit. 1995. *Symmetric coordination: An alternative theory of phrase structure*. Tübingen: Max Niemeyer Verlag.
West, Shannon Louise. 2003. *Subjects and objects in Assiniboine Nakoda*. Victoria: University of Victoria (Master's thesis).
Zenes, Bozena. 2009. *Coordination of noun phrases in Biloxi*. Chicago: Northeastern Illinois University.

Chapter 16

Information-structural variations in Siouan languages

Bryan James Gordon

Most previous information-structural analysis on Siouan languages has been fragmentary and based on incommensurable definitions and frameworks. A corpus study drawing on transcribed and recorded utterances from languages in all major branches of Siouan represents a first step towards generalizable and practical knowledge of the morphosyntactic and intonational indices of information-structural categories in Siouan languages. This study focuses on variations previously noted in the literature – intonational marking and demarking, postverbal arguments, reduction of referring expressions, OSV word order and switch-topic markers.

1 Introduction

The formal linguistic record offers excellent, comprehensive documentation of morphological, syntactic and phonological structures in Siouan languages. The documentary record of those forms' functions and meanings, however, is hit-and-miss – fragmented, uncomprehensive, and characterized by individual linguists' often incommensurable functional analytical frameworks. Most of us documentary linguists, in fact, have been trained to seek out, describe and privilege descriptions of context-free levels of semantic meaning, and our work has less to say about the meanings that emerge in and create our linguistic and social contexts, which may well be the most important types of meanings to community-based language-reclamation projects which try to adapt and use our work. The work I present here, while not free of its own analytical framework, attempts to rectify this situation by providing a comprehensive albeit partial description of one area of functional variation in form – information structure – based on a largely qualitative corpus study sampling data from nine Siouan languages.

Most, if not all, previous descriptions of Siouan grammars have had something to say about information structure. Rudin (1998), Koontz (2003) and Eschenberg

Bryan James Gordon. 2016. Information-structural variations in Siouan languages. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 393–423. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.179
DOI:10.17169/langsci.b94.179

(2005) provide grammar/discourse analyses of specific features of Umo^{n}ho^{n} Iye (Omaha). Graczyk (1991: 242–260) and de Reuse (1994) describe the functions of noun incorporation in Apsaalooke (Crow)[1] and Lakota, respectively. Rood offers analyses of presupposition (1977) and variation between definite articles (1985), also in Lakota. Kaufman (2008) offers some information-structural analysis of Taneksą (Biloxi). Wolvengrey (1991) describes a switch-topic marker (a "focus" marker in his terms) in Rų'eta (Mandan). Issues of "oldness", "emphasis", "topic" and "focus" surface in various formal grammars including Cumberland's (2005) Nak^{h}on'i'a (Assiniboine) and Boyle's (2007) Hidatsa grammar. Ingham's (2003) work is a notably comprehensive discourse-analytical approach to information-structural variations in Lakota. These works, however, generally focus on a single phenomenon, and do not all situate themselves as relevant to a more general field of information-structurally meaningful variation.

Here I adopt a unifying toolkit which allows us to consider, compare and extend previous findings under a common metalanguage. My hope is not to argue for or against previous analysts' theoretical or methodological goals, or my own, but to make possible a more coherent and useful conversation about Siouan information-structural variations for a variety of audiences, including but not limited to theoretical linguists and community-based language-reclamation programs.

2 Method

For this study I coded interlinear text from nine languages, and audio data from four, a corpus comprising the majority of text and audio available to me at the time of study in 2009. These languages span the four branches of the Siouan language family, including all three subbranches of the large Mississippi Valley Siouan subfamily: Rų'eta; the Southeastern Siouan language Taneksą; the Missouri Valley Siouan languages Hidatsa and Apsaalooke; and the Mississippi Valley Siouan languages Nak^{h}on'i'a, Lakota, Umo^{n}ho^{n} Iye and Pa^{n}ka Iye (Omaha and Ponca), Ho-Chunk, and Baxoje Ich^e and Jiwere Ich^e (Ioway and Otoe). Of these I was able to access audio data in Hidatsa, Umo^{n}ho^{n} Iye, Baxoje Ich^e

[1] I use autonyms where available for language names, providing the English colonial name in parentheses only on first mention. It should not be assumed that the autonyms have the same reference as English colonial names. "Mandan", for example, refers to a group of related varieties of which Rų'eta is only one. Where the English name is reasonably close in both form and reference to an appropriate autonym, as in the case of "Lakota", "Ho-Chunk" or "Hidatsa", I use the English name.

and Ho-Chunk. My sample, although representative of genetic diversity within the Siouan family, is not balanced across considerations of genre, time period or other sociolinguistic factors. These sources and the speakers who produced them are listed, with an utterance or page count, under "Primary resources" before the references at the end of the chapter. I coded each primary resource for formal variation (intonational, segmental and morphosyntactic criteria summarized in §2.2–§2.3) and information-structural function (criteria summarized in §2.1). Some of these resources I coded in their entirety; in others (e.g. Dorsey 1890) I simply sampled a few works, attempting to capture multiple genres. I also drew many examples for this paper from secondary resources, and they are cited as such when they occur.

2.1 Information-structural coding procedure

The following is a condensed version of my coding criteria. My criteria are drawn with some modification from the Givenness Hierarchy (Gundel, Hedberg & Zacharski 1993) and Ward & Birner's (2001) framework, with attention towards commensurability with other frameworks. Because commensurability is one of my objectives, I do not use the terms "given," "old," "topic" or "focus" without specifying modifiers and definitions. I am the sole coder, so I have no intercoder reliability measure for these criteria, but it may be noted that I was applying Gundel's criteria in the Minnesota Cognitive Status Research Group[2] at the time this study was conceived, and that we achieved 85% intercoder reliability. All errors and misapplications are my own.

(1) a. Code a form as a LINK if its referent stands in a poset relation with a salient or inferrable alternative. See Ward & Birner (2001: 121) for examples, and cf. the categories of CONTRAST and RESTRICTION in Erteschik-Shir (2007). Forms coded as links are underlined.

b. Code a form as RECOVERABLE if its referent is an ATTENTION-CENTRAL entity or an INFERRABLE predicate. These categories are operationalized as in Gundel et al. (1993; 2010), but I replace Gundel's term IN FOCUS with ATTENTION-CENTRAL to avoid confusion with relational focus. A referent is attention-central if its utterer can assume that her audience is consciously attending to it (cf. CONTINUING TOPIC in Erteschik-Shir 2007, and HIGHLY ACCESSIBLE in Ariel 1990). A referent is inferrable if the discourse model at time of

[2] The Minnesota Cognitive Status Research Group was funded by a National Science Foundation grant, *A cross-linguistic study of reference and cognitive status* (BCS0519890, PI Jeanette Gundel).

reference gives ample and recoverable evidence for its validity (if a predicate) or existence (if an entity). Such evidence may be logical, narrative, stereotypic or based on general cultural knowledge. Forms coded as recoverable are *italicized*.

c. Code a form as a RELATIONAL TOPIC or RELATIONAL FOCUS if its referent stands in a directed relation (i.e. a relationship of semantic scope like quantification, or pragmatic/framing scope in the sense of Goffman 1974, e.g. stage topics, scene-modifiers, activating topics, extragrammatical mentions, unactivated definites, deixis, "aboutness" and "predication-of"). Relational topics are forms whose referents take scope, and relational foci are forms whose referents are under scope. Forms coded as relational topics are followed by a right angle bracket >, while those coded as relational foci are followed by a left angle bracket <.

Example (2) diagrammatically represents the coding marks as they are presented in all examples taken from primary sources:

(2) link *recoverable* relational topic > relational focus <

2.2 Intonation-structural coding procedure

Armik Mirzayan (2011) has developed a ToBI coding protocol (cf. Pierrehumbert 1980; Beckman & Pierrehumbert 1986; and Pierrehumbert & Hirschberg 1990) for Lakota, but at the time of this study no such protocol was available, so I developed one myself. This section principally includes the basic description and justification of the protocol I developed; precise coding criteria have been omitted for length considerations and are available on request.

I was able to establish consistent criteria for identifying the standard array of prosodic phrase levels – the ACCENT PHRASE (AccP), the INTERMEDIATE PHRASE (IntP) and the INTONATIONAL PHRASE (IP) – in all four of the sampled languages. Accent phrases in all four languages maximally consist of a low-high-low contour in which either the high point or the high-low fall is accorded greatest prosodic prominence. Such contours are represented in ToBI notation as LH*L, and all four languages have them, although the first L tone in particular is often absent. Intermediate phrases in all four languages consist of one or more accent phrases followed by a phrase accent, either !H or L; and intonational phrases in all four languages consist of one or more intermediate phrases followed by a boundary

tone, either !H% or L%. The exclamation point before H phrase accents and boundary tones indicates downstep: I found no evidence of upstepped H phrase accents or boundary tones in any of the four languages.

My protocol was designed to cover four languages (Umo^{n}ho^{n} Iye, Ho-Chunk, Baxoje Ich^e and Hidatsa) rather than one, and so is shallower and less detailed than Mirzayan's. The simple apparatus sketched here is designed to capture high-level intonational variations across Siouan languages, and should not be taken as evidence that Siouan intonational structures are simple. Besides Mirzayan (2011) see also Larson (2009) on Umo^{n}ho^{n} Iye for richer descriptions of intonational variation in particular Siouan languages.

2.3 Selection of formal variants of interest

In information structure and intonational structure I have begun with high-level, a priori coding categories and attempted to apply them to the entire corpus. In deciding which formal (morphosyntactic and segmental as well as intonational) variations to correlate to information structure, however, I have made no such attempt. Instead, I have specifically looked at some of the formal variations in previous descriptions of Siouan languages: postverbal arguments (§3.2); degrees of reduction of noun phrases or referring expressions, from zero reference to "determiner drop" and noun incorporation (§3.3); OSV word order (§3.5); switch-topic markers (§3.6); and intonational processes of "marking" and "demarking" which surfaced during my ToBI coding (§3.4 and §3.1, respectively).

3 Findings

3.1 Deaccenting

I use the term DEACCENTING to describe intonational variations in which a given word or string of words may be realized either with or without the H* pitch-accent head of an AccP. The variant without the pitch accent is the DEACCENTED variant. This term implies that the accented variant is canonical. Bolinger (1986: 100) makes this explicit: "[A] neutral sentence accents all content words, … and a non-neutral or marked sentence would be one in which one or more words have been deaccented."

Every audio resource I coded has examples of deaccenting, and all examples of deaccenting signal a recoverable referent. In example (3), Umo^{n}ho^{n} Iye speaker Clifford Wolfe, Sr., deaccents the strings *k^{h}i égithe* 'and so it happened that' and

ahí 'arrived there'. Both are absorbed into the LH*L contour of the AccP headed by *weáhidexti* 'truly far':

(3) *Shkó^{n}-t^{h}e waséko^{n}, k^{h}i égithe weáhidexti ahí.*[3]

shkó^{n}-t^{h}e >	waséko^{n} <	*k^{h}i*	*égithe*	weáhidexti
[L H* L]	[L !H* L] L	[L		H* L
movement-the fast		and it.happened.that		ahead.really

ahí <
] L L%
arrive.there.PROX

'She was moving fast, so it happened that she got pretty far ahead.'

In example (3) and in other audio examples, I provide a ToBI line under the information-structural coding and above the gloss line. Here, AccP's are represented by square brackets. IntP's are recognisable by their pitch accents outside square brackets, and the IP by its boundary tone.

For this example, I also provide a Praat (Boersma & Weenink 1992) screenshot, but omit it from subsequent examples due to space considerations. In the screenshot, the blue line is the pitch contour in Hz, and the green line the intensity contour in dB.[4] The ToBI line in my Praat annotations is simplified relative to the ToBI line in the examples presented in this paper. Glosses in the Praat annotations are provided at the IntP level, which is a technique I use to simulate the intermediate-level prosodic chunking present in the audio.

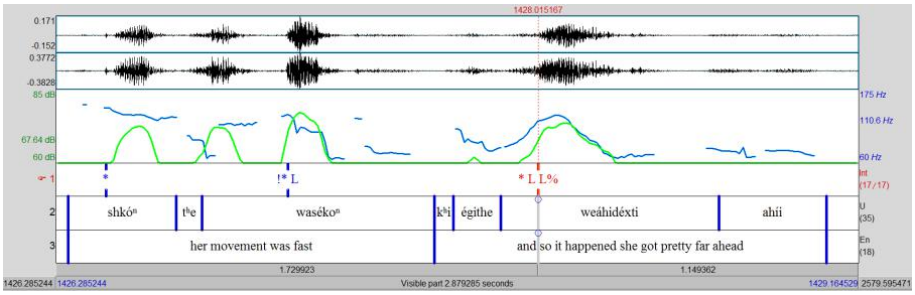


Figure 1: Praat screenshot corresponding to example (3)

In example (3), two sections have been deaccented and absorbed into the L tones that bound an adjacent H* phrase accent. Both have been coded as recov-

[3] Rudin field tapes

[4] It should further be noted that the blue line, being a computed estimate of pitch, is not sufficient evidence for any ToBI coding criterion, but must be accompanied by impressionistic auditing.

erable: *k^{h}i égithe* because it is a frequent discourse marker that is likely always recoverable in narrative. The predicate *ahíi* is inferrable given immediately preceding references to motion, speed and distance.

We may distinguish between "true" deaccenting, as in example (3), where the deaccented part of the pitch contour is either flat or very slowly declining, and "weak" deaccenting, as in example (4), where Mr. Wolfe uses a compressed version of the LH*L contour. We are looking at *níkshi^{n}ga*, whose ToBI line is annotated with (LM*L). This is because the pitch range of this contour is compressed within a range smaller than the pitch range of the L tone in the adjacent AccP, so that it is realized as something more like (LM*L) than a full [LH*L] AccP.

(4) *Hi^{n}béazhi-t^{h}e égithe, táxti sí-t^{h}e, thiwágazui níkshi^{n}ga.*[5]

hi^{n}bé-ázhi-t^{h}e	*égithe* >	táxti	sí-t^{h}e	thiwágazui
[L H* L	] !H	[L H* L]	[L !H* L]	[L H* L
moccasin-not-that	it-happened-that	deer	foot-that	notice-PROX
níkshi^{n}ga				
(L M* L)] L L%				
person				

'The man noticed that she was not wearing shoes, but rather had the feet of a deer.'

Mirzayan (2011: 121–123) presents similar compressed AccP's in Lakota. This phenomenon may suggest a continuous process at work, in which the degree of accenting in AccP's is continuously gradient. Such a process would challenge the usefulness of a discrete annotation system like ToBI for capturing deaccenting phenomena in Siouan languages. If this scale from "true" to "weak" deaccenting corresponds to variation in information-structural function, however, my coding criteria have not captured it. Instead, all types of deaccenting I have described signal recoverability of referents. This finding is entirely dependent on my methodology, however, and the area of gradient deaccenting and potential information-structural corollaries needs to be investigated further.

In short, all the audio resources I coded display deaccented forms, and all such forms have recoverable referents.

3.2 Postverbal arguments

Since all Siouan languages canonically position verbs after their arguments, the syntactic status of postverbal arguments is controversial. All Siouan languages at

[5] Rudin field tapes

very least permit postsentential referring expressions to clarify one or more verbal arguments, but grammarians have varied in accepting postverbal arguments as elements of the same sentence (cf. Rudin 1998; Mithun 1999; Ingham 2003: 76; Cumberland 2005: 421; Boyle 2007: 292–293; Gordon 2008). My attention to deaccenting phenomena in this study may shed light on this question and enabling language workers to distinguish between those postverbal arguments which are clearly "outside" the sentences and those which are more likely "inside".

Accented postverbal arguments (those which are pronounced with the [LH*L] contour of an AccP) typically have nonrecoverable referents. Instead, they clarify uncertain information, providing new information about both entities and predicates. (See Graczyk 1991: 103 for a relevant discussion on Apsaalooke afterthought.) Accented postverbal arguments are usually separated from the verb not only by an AccP boundary but by a prosodic break (i.e. an IntP phrase accent or IP boundary tone). In example (5), Baxoje Ich^e speaker ThigréPi concludes an IntP on *ráye u^{n}k^ú^{n}ñe^su^{n}* 'they gave me a name', as seen in the L phrase accent, here evidenced by early and sustained low pitch. Then he adds another IntP in pronouncing the postverbal argument *Baxóje ráye* 'a Baxoje name':[6]

(5) *Woxáñi migragáñi^su^{n}, ráye, ráye u^{n}k^ú^{n}ñe^su^{n}, Baxóje ráye.*[7]
woxáñi migragáñi^su^{n} < ráye < *ráye* u^{n}k^ú^{n}ñe^su^{n} <
[L H* L L] [L H* L L] [L H* L L]
cherished me.their.named.now name name me.give.PL.now
Baxóje < *ráye*
[L !H* L L] L%
Ioway name
'They named me as a cherished one, a name, they gave me a name, a Baxoje name.'

Deaccented postverbal arguments, in contrast – like *níkshi^{n}ga* in example (4) – are not separated by prosodic breaks, and signal recoverable referents. Example (6) is a useful model of the relationship between postverbal arguments and recoverability in general. Hidatsa speaker Helen Wilkinson says *kúahe* 'this' first in preverbal position when it is not recoverable, and then in postverbal position when in the second sentence (with the form *kúac*) it is recoverable:

[6] Square brackets in the ToBI line here refer to IntP boundaries rather than AccP boundaries as previously.
[7] ThigréPi in Goodtracks (2004)

(6) *Kúahe aku-iháaraci. Maapúkšaruxpáaka wáahaapʰaak ráahaa'he ita-arukiwé' kúac.*[8]
kúahe > aku-iháaraci < Maapúkšaruxpáaka wáahaapʰaak ráahaa'he
Here.this different.sorta Snake.People raid go.CAUS.that
ita-arukiwé' < *kúac*
their.story.tell here.this
'This is a different story [from the story of the Bird Woman]. This is the story of the Shoshone raid.'

There are some examples of multiple postverbal arguments. In example (7), I present two. In example (7a), Helen Wilkinson first gives us the recoverable postverbal argument *maa-iháa'š* 'the enemy' before then further specifying *Waapúkšaruxpáaka'š* 'the Snake People'. Although the two phrases share the same entity referent, the function of *Waapúkšaruxpáaka'š* is to supply a predicate which is not inferrable at this point in the discourse model, so *Waapúkšaruxpáaka'š* is coded nonrecoverable (one of the few unseparated by commas which I coded nonrecoverable). In example (7b), however, the telling of the Baxoje legend of Béñeiŋe recorded by Gordon Marsh (1936) displays two postverbal arguments. Here the two are not coreferent, nor are they in the same argument relation to the verb, and they are both recoverable.

(7) a. *Hii šee awá ihtúutiru ú'šiak káawarec maa-iháa'š Waapúkšaruxpáaka'š.*[9]
hii šée awá > ihtúutiru ú'šiak káawarec < *maa-iháa'š*
and that ground hill.base.at arrive.ss be.there.PL.NE enemy.PL.DEF.the
Waapúkšaruxpáaka'š <
Snake.People.PL.DEF.the
'And the enemy, the Shoshone, were on that ground, having gotten to the base of the hill.'

b. *Nahésge, igwáhuŋa súŋe Béñeiŋe.*[10]
nahésge > igwáhuŋa < *súŋe Béñeiŋe*
be.if his.know horse Béñeiŋe
'And so the horse$_i$ recognized his$_i$ [owner] Béñeiŋe.'

Without audio we can't be sure whether the apparent double postverbal arguments are really all part of the same utterance, but it makes sense to speculate

[8] Parks, Jones & Hollow (1978: 18: 1)
[9] Parks, Jones & Hollow (1978: 22: 26)
[10] Marsh (1936)

that, of the two, Ms. Wilkinson's example (7a) is more likely to involve a major prosodic break than the Baxoje Ich^e example (7b).

In written transcripts for which no audio is available, I found the transcriber's comma a somewhat useful index of prosody – and recoverability. Many of the postverbal arguments I coded – like Ms. Wilkinson's in example (6) – were not separated by commas from preceding material. Although transcriber's commas may somewhat reliably correlate with IntP or IP boundaries as annotated in ToBI, we cannot be too sure of how strong this correlation is. Still, nearly all of the postverbal arguments not separated by commas had recoverable referents. Presence of a comma, on the other hand, does not seem to reliably predict recoverability of the referent.

Not all languages (or speakers) are alike with respect to the frequency of postverbal arguments. I found many examples of postverbal reference to recoverable entities in Taneksą, Rų'eta, Hidatsa, and Umo^{n}ho^{n} Iye and Pa^{n}ka Iye texts, and very low rates in Ho-Chunk and Apsaalooke texts – but still nearly all of the postverbal arguments in the Ho-Chunk and Apsaalooke texts I coded had recoverable referents. All the languages I looked at allow at least some use of postverbal arguments to refer to recoverable entities. In some languages, like Umo^{n}ho^{n} Iye and Pa^{n}ka Iye, this use may be (or have been) obligatory. In Gordon (2008), I sampled 51 continuing topics in Umo^{n}ho^{n} Iye and Pa^{n}ka Iye and found 42 of them were referred to postverbally. Of the 9 preverbal, most were within repeated collocations. This suggests that in some languages recoverable referents not only *can* be referred to postverbally, but *must* in most contexts.

To summarize, all the languages in this study make use of postverbal arguments, although they vary in frequency. Even without audio data to distinguish between deaccented and accented postverbal arguments, the strong tendency is observable for postverbal arguments to refer to recoverable entities. The index of the transcriber's comma may serve as circumstantial evidence. Its absence indirectly indexes recoverability of postverbal referents while also directly, if weakly, indexing deaccenting. Forms with nonrecoverable referents are often found after commas, on the other hand. Where audio data is available, the case is clearer: deaccented postverbal arguments, like other deaccented material, have recoverable referents, and are intrasentential (part of the same sentence as the verb they follow) by intonational if not syntactic criteria.

3.3 Reduced nominal referring expressions

3.3.1 Determiner drop in languages with two indefinite articles

Of the languages represented in this study, four – Lakota, Nak^{h}on'i'a, Apsaalooke and Ho-Chunk – have two indefinite articles, one specific and one nonspecific. The texts I coded in these four languages display determiner use in all nominal expressions which are specific (i.e. which refer to entities and not just to predicates or types) but nonrecoverable, allowing bare nominal expressions only for nonspecific or recoverable referents.[11] In example (8) Apsaalooke speaker Francis Stewart refers twice to a specific quantity of water, but only the second reference is recoverable. This exemplifies how, among specific referents, recoverability makes the difference between determiner use and bare expressions:

(8) *Hinné wíliash kala iháatak huuk. Bilé xatáakelak.*[12]
hinné wilíash kala > iháatak < huuk *bilé* xatáakelak <
this water.the now strange hearsay water move-DS
'This water (they say) was getting weird now. The water was moving.'

This finding at first glance contradicts Cumberland (2005), who found bare specific NP's acceptable in elicitation with Nak^{h}on'i'a speakers. However, my finding is based on text analysis as opposed to elicitation – a social informational setting which presents extraordinary information-structural pragmatic conditions even when information structure is deliberately controlled (as it usually is not). It may also be noted that Cumberland's definition of "specific" is not information-structural.

3.3.2 Determiner drop in languages with one or no indefinite article

Three other language groups – Umo^{n}ho^{n} Iye and Pa^{n}ka Iye, Rų'eta and Hidatsa – have either one or no indefinite article. These languages tend to fit a weaker version of the generalization in §3.3.1, with the exception that speakers do sometimes refer to specific indefinite referents (nonidentifiable entity referents) using bare nominal expressions. In example (9), Hidatsa speaker Annie Eagle makes two specific indefinite references – 'a sinew' and 'a fire' – but only uses a determiner on the first. This may be because sinews are more canonically objectlike and therefore have higher specificity potential than fires.

[11] I do not consider generic reference in this study.
[12] Wallace (1993: 194: 178)

(9) *Macúawa rúhcak wiráa' úawa.*[13]
macúawa rúhcak wiráa' úawa <
sinew.a take.ss fire make.fire.TEMP
'He [the Thunderbird] took a sinew and made a fire.'

The extreme case is Rų'eta, which has no indefinite articles at all. Specific indefinites, thus, are usually bare in Rų'eta, as in example (10) from speaker Stephen Bird:

(10) *Dá·ha·mį·mį· mą́nąrok pą́·xe híro·mąko'š.*[14]
dá·ha·mį·mį· > mą́nąrok > pą́·xe > híro·mąko'š <
go.while.PROG.PROG tree.in potato arrive.there.NARR.PST.DECL.M
'While he was going along, he came to some wild potatoes in the woods.'

A possible explanation for the difference between the languages in this section and those in §3.3.1 follows: Languages with both a specific and a nonspecific indefinite article may tend to make the use of the specific indefinite article obligatory while leaving the use of the nonspecific indefinite article optional. Languages which lack this distinction, on the other hand, do not have the opportunity to develop obligatory use of their one indefinite article – and those with no indefinite article at all, like Rų'eta, must allow bare expressions to refer to specific indefinite referents.

3.3.3 Determiner drop in all sampled languages

One generalization may be made which holds of all nine languages in this study, including those with incipient article classes (Taneksą and Baxoje Ich^e and Jiwere Ich^e): The majority of bare nominal expressions have referents which are either nonspecific or recoverable, regardless of the particular distribution or grammatical constraints involved in each individual language. This is, surprisingly, true even of languages like Baxoje Ich^e which regularly use bare nominal referring expressions for specific and nonrecoverable referents. In example (11) speaker Wa^{n}^sígeChéMi uses a bare nominal referring expression for a recoverable referent. Had 'my grandmother' not been recoverable (e.g. had she been new to the narrative), a determined form like *hi^{n}kúñi nahé* would have been more likely (Jimm GoodTracks, p.c.).

[13] Parks, Jones & Hollow (1978: 15–16: 76–77)
[14] Carter (1991: 29–30: 8)

(11) *Ídare hi^{n}kúñi wárudhàshgu^{n} warúje.*[15]
ídare > *hi^{n}kúñi* wárudhàshgu^{n} < *warúje*
then my.grandmother some.take.INFER food
'Then it seems my grandmother took some of the food.'

This generalization is weak, and may say nothing special about Siouan languages, but it raises interesting theoretical questions. Why are recoverable referents, and referents which lack a specific entity, lumped together on the same end of a formal variation? I speculate that they are similar in their light processing load.

3.3.4 "Determiner-drop drop": recent rises in obligatoriness

In older Lakota narratives speakers often use bare nominal expressions to refer to recoverable referents. Ella Deloria (1932: F831) makes five references to Rabbit (and two vocatives), three of which are determined and two bare.[16] In example (12), we may observe one of the bare references:

(12) *K'eyaš Maštíŋčala tákuni yútešni čhaŋké tókȟa-wok'ušni ke'.*[17]
k'éyaš *Maštíŋčala* tákuni yútešni < čhaŋké > tókȟa-wok'ušni <
but Rabbit nothing ate.not so how-something.give.not
ke'.
HEARSAY.DECL
'But Rabbit ate nothing and so he had nothing to give [to the boy], they say.'

In more contemporary registers, however, Lakota tends not to exhibit determiner drop, requiring determiners for all specific referents. Many contemporary speakers, thus, consider the bare NP *Maštíŋčala* in example (12) to be missing its article *kiŋ* or *k'uŋ*. Hiroki Nomoto (p.c.) informs me that the same phenomenon occurs in certain obligatory-classifier languages such as Cantonese, in which

[15] Marsh (1936)

[16] I am grateful to an anonymous reviewer who suggests that in Lakota, as opposed to other languages sampled here, references made with proper names are ordinarily made using bare expressions. A close examination of Deloria (1932), however, does not categorically bear this out. Rabbit is mentioned in Deloria (1932: F831, F843, F844 and F847). Of these works F831 and F844 typically include determined expressions referring to Rabbit while F843 and F847 typically include bare expressions. Thus, the picture for Lakota is one of variation.

[17] Deloria (1932: 4–5: 12)

speakers vary on whether they drop determiners (classifiers) or even accept determiner drop as grammatical for highly presupposeable referents. Similarly, at the Title VII Umo^{n}ho^{n} Language and Cultural Center we often "reinsert" dropped determiners into transcripts and materials, and this is described as a correction. I believe that this move may emerge in part from the influence of the "rule"-based project of documentary linguistics upon community-based programs. Documetary linguistics has set in motion rapid codificational change to community language ideologies, and has privileged questions of speaker skill and grammaticality (heard as "correctness" by most audiences) over descriptions of legitimate variation. Determiner drop has not, however, disappeared from fluent spoken language, and may be viewed as itself correct.

It's possible that the shift towards obligatoriness in Lakota will be completed in Umo^{n}ho^{n} Iye too, and that in both cases it will have been led by stylistic variation driven by the metalinguistic notion of correctness. Interestingly, the Nak^{h}on'i'a speakers who worked with Cumberland viewed determiner use metalinguistically as optional in general, and accepted constructed examples with determiner drop regardless of recoverability (Cumberland 2005: 345), but Nak^{h}on'i'a texts do not differ appreciably from Lakota or Umo^{n}ho^{n} Iye texts in the use of bare nominal expressions for recoverable referents. In all cases, a broader spectrum of genres and registers will need to be analysed before drawing conclusions about the state of determiner drop in Siouan languages as a whole.

3.3.5 Determiner drop and noun incorporation as continuum

Noun incorporation occurs primarily in Missouri Valley Siouan, and is most extensive in Apsaalooke. Graczyk considers a variety of its functions, information-structural and otherwise, and synthesizes from the literature his claim that incorporated objects tend "to be non-referential, non-individuated, non-specific, non-autonomous, non-countable, and the object-verb compound typically expresses unitary, habitual, characteristic, typical, institutionalized activities" (1991: 244). Autonomy, individuation and countability are characteristic of entity referents, so what Graczyk says here is essentially that incorporated nouns tend to refer nonspecifically, and that the noun-verb compound itself has a relatively unitary concept structure. This second claim touches on an aspect of information structure which I have not covered in this study.

As for the first claim, this study supports it as a statistical generalization, although not as a categorical one. Apsaalooke does display instances of incorporated nouns with specific referents, as in example (13) (uncoded):

(13) a. *biíttaashteelitdialaalak*[18]
biíttaashteelitdialaalak
me/my.shirt.sorta.make.you.if
'if you make my shirt' *or* 'if you make me a shirt'

b. *Basahpawaannáastawiilakoosh ítchikissuuk.*[19]
basahpawaannáastawiilakoosh ítchikissuuk
me/my.moccasin.bead.you.string.me.you.give.the good.sport.PL.DECL
'The moccasins that you beaded for me are pretty.'

Example (13a) is ambiguous between two readings. In one, the first person pronominal *b-* is interpreted as a possessive prefix on the incorporated noun *iíttaashtee* 'shirt', and in the other it is interpreted as a pronominal argument on the predicate *iíttaashteelitdia* 'make-shirt'. In example (13b), the possessed phrase *(b)asahpa* '(my) moccasins' is incorporated in the predicate *waannáasta* 'you string beads' which itself has the incorporated *waan* 'bead' in it. This double-incorporation is essentially an example of the body-part incorporation in example (13a), except that the possessor is another incorporated object.

The specific referents of the incorporated nouns *iíttaashtee* and *asahpa* in example (13) are both recoverable, and thus fit my observations about determiner drop in §3.3.3: Like determiner drop, incorporation may be conditioned similarly by both nonspecificity and recoverability. Even in the case of nonspecific referents, as in example (14) (uncoded), recoverable references to types are made by incorporated nouns (*hunnáappaxbialaalak*) while less recoverable first references are made by bare nouns (*hulé*):

(14) *Dáassuua ashkawúuan hulé dappaxíssah, hunnáappaxbialaalak awéeleen díah.*[20]
dáassuua ashkawúuan hulé dappaxíssah
your.house inside.at bone split.not.COMMAND
hunnáappaxbialaalak awéeleen díah
bone.you.split.MOD.you.if outside.at do.COMMAND
'Don't split bones inside. If you want to split bones, do it outside!'

Although Lakota lacks full incorporation of the sort seen in Missouri Valley Siouan, de Reuse (1994) considers noun incorporation and determiner drop as

[18] Graczyk (1991: 247)
[19] Graczyk (1991: 257)
[20] Graczyk (1991: 250), cited as Old Coyote (1985: 13)

related phenomena in his study. He finds cases of recoverable referents with determiner drop ("noun stripping" in his terms[21]), like the second instance of *ištá* in example (15) (uncoded):

(15) *Čhaŋké maǧážu mní ištá kiŋ owíčhakičaštaŋ hiŋ na úŋ ištá wičhákičiyužaža haŋ čhaŋké tuŋwáŋ pi skhé'.*[22]
čhaŋké maǧážu mní ištá kiŋ owíčhakičaštaŋ hiŋ na úŋ
and.so rain water eye the in.them.BEN.INSTR.pour CONT and INSTR
ištá wičhákičiyužaža haŋ čhaŋké tuŋwáŋ pi skhé'
eye them.textscben.instr.wash CONT and.so see PL INFER.DECL
'So he poured rain water in their eyes, and washed their eyes with it until they were able to see.'

De Reuse (1994) considers "syntactic compounding" as an intermediate phenomenon between "noun stripping" (determiner drop) and full noun incorporation. He cites the example in (16) (uncoded), from Ella Deloria. Although de Reuse's example is out of context, it appears that 'the child' is likely recoverable and certainly specific:[23]

(16) *Hokší okìle pi škhé.*[24]
hokší okìle pi škhé
child look.for PL INFER.DECL
'They looked for the child.'

As Graczyk (1991) noted, linguists have often drawn associations between noun incorporation and nonspecificity. Recoverability has perhaps been less frequently associated, but as de Reuse (1994) suggested, and as my study corroborates, it is useful to look at how recoverability functions alongside nonspecificity to encourage not only noun incorporation, but a continuum of related variations with noun incorporation on one extreme and determiner drop near the other.

[21] As a reviewer points out, de Reuse defines *noun stripping* more or less as a lexical process, and if I were to follow his approach I might consider the absence of determiners to be point made moot by definition. I see no contradiction, however, in subsuming a lexicalized *noun stripping* under more general considerations of determiner use which can be directly observed from the data without making inferences about the lexical status of noun-verb collocations.

[22] de Reuse (1994: 232)

[23] De Reuse uses the grave accent, as on *okìle*, to indicate that the word does not receive primary stress, i.e. that it is part of the same AccP as the preceding word (or, it might be argued, a nested, subordinate AccP within the main AccP, cf. example (4)). "Syntactic compounding", then, necessarily includes intonational structure alongside the compositional, ordered rules de Reuse considers in his analysis.

[24] Deloria (1932: 48: 4), cited in de Reuse (1994)

3.3.6 Zero reference (argument drop)

All Siouan languages make use of zero reference in all argument positions, in all persons. The referents of such zero expressions are recoverable. Rų'eta speaker Stephen Bird says utterance (17) at a point in the narrative where both Trickster and some potatoes are recoverable, and so he refers to both with no nominal expression:

(17) *Ó'haranį ké'nį dutó·mąko'š.*[25]
ó'haranį > ké'nį dutó·mąko'š <
so.and dig.and eat.NPST.DECL.M
'And so he digs and eats them.'

3.4 Intonational bounding of links, relational topics and relational foci

We have seen the importance of AccP boundaries in previous sections: Words which do not project one have recoverable referents. In this section we will see how the phrase accents which bound IntP are put to use in demarcating informationally prominent material. IntP boundaries tend to coincide with strings coded as having referents which are either links, relational topics or relational foci. The converse is not true: Links, relational topics and relational foci do not in general tend to require an IntP boundary in the texts I have coded. Recoverable referents, on the other hand, tend *not* to be associated with forms specially demarcated by IntP (or AccP) boundaries. Examples in this section follow the ToBI presentation I used in §3.1, with the exception that square brackets here represent IntP boundaries instead of AccP boundaries, and thus include rather than exclude phrase accents.

Pierrehumbert & Hirschberg (1990) generalize that H phrase accents are projected on relatively "forward-looking" material. If this holds of Siouan languages, then the threshold of "forward-looking enough" must be higher in Siouan languages than in English, where a L phrase accent on a stage topic might sound a bit odd. I found stage topics are referred to by IntP's with both L and !H phrase accents, but L phrase accents were more common in this study. A Hidatsa example from the Water Buster Account is given in (18):

[25] Carter (1991: 33: 28)

(18) *Še'erúhaak waapixupá rúupatook kiráahuac.*[26]

še'erúhaak	waapixupá	rúupatook >	kiráahuac <
[L H* L L]	[L H* L !H]	[L H* L L]	[L H* L L] L%
then	Sunday	two	we.came.for.them

'Then, two weeks later, we came for them.'

The first and third IntP in example (18) provide complete stage topics, but the second ends with a hesitation not resolved until the third. The second IntP is the only one with a !H phrase accent. This weakly supports Pierrehumbert & Hirschberg's generalization in that it is the most "forward-looking" of the three. The fourth IntP, like nearly all the other relational foci I coded, receives a L phrase accent.

In example (19), Ho-Chunk speaker Cecil Garvin uses a !H phrase accent on both the linking *coowexjįšgera* 'just a little' and the stage topic *karacgą 'ųnąąkašgera* 'since they were drinking'. On the other hand, the last two links – *coowera* 'a little' and *hoinąk haanįsge* 'I started too' – both of which are also coded as relational foci, receive L phrase accents:

(19) *Coowexjįšgera karacgą 'ųnąąkašgera coowera hoinąk haanįsge.*[27]

Coowexjįšgera	karacgą	'ųnąąkašgera >	*coowera* <	hoinąk	haanįsge <	
[L H* L !H]	[L !H* L	L]	[L !H* L L]	[L H* L	[L%	
just.a.little	drink	they.were.since	a.little	I.start	I.too	

'Since they were drinking, just a little, I started a little too.'

Another kind of material we might even more strongly expect to take a !H phrase accent, following Pierrehumbert & Hirschberg (1990), are explicitly forward-looking references like list items and other incomplete references, e.g. *waapixupá* in example (18). But, like stage topics and links, forward-looking reference appears to make use of both !H and L phrase accents in Siouan languages. The speaker in the Hidatsa example (20), again from the Water Buster Account, makes a complete predication in his first IntP, and then elaborates it in his second. He may have used the !H phrase accent "foward-lookingly" to signal that an elaboration was planned:

[26] Lowie (1939)

[27] "Connection (humour)" in Hartmann & Marschke (2010)

(20) *Úuwaca kirakap^{h}a'áhku pirakíhtia toopatóok kirakap^{h}áapak.*[28]
úuwaca kirakap^{h}a'áhku < pirakíhtia toopatóok *kirakap^{h}áapak* <
[L H* L L !H* L !H] [L H* L L !H* L L] L%
money they.kept.collecting hundred four they.collected
'They kept raising money; they raised four hundred dollars.'

ThigréPi's forward-looking reference *ráye* 'name', on the other hand, has a L phrase accent in example (21) (repeated from example (5)). It is unclear to me whether these L-final IntP's convey more of a sense of autonomy or finality than the "incomplete" IntP in example (20).[29]

(21) *Woxáñi migragáñi^su^{n}, ráye, ráye u^{n}k^ú^{n}ñe^su^{n}, Baxóje ráye.*[30]
woxáñi migragáñi^su^{n} < ráye < *ráye* u^{n}k^ú^{n}ñe^su^{n} <
[L H* L L] [L H* L L] [L H* L L]
cherished me.their.named.now name name me.give.PL.now
Baxóje < *ráye*
[L !H* L L] L%
Ioway name
'They named me as a cherished one, a name, they gave me a name, a Baxoje name.'

Example (22), from Umo^{n}ho^{n} Iye speaker Clifford Wolfe, Sr., contains three boundaries between different coding categories. There is a boundary between the linking stage topic *shó^{n}xti* 'nevertheless' and the nonlinking activating topic *wa'ú-thi^{n}* 'the woman', then one before the relational focus *ní tható^{n}-bazhíi* 'he didn't drink water', and another before the separate relational focus *wa'ú-thi^{n} uthúhai* 'he followed the woman'. Each of these relational boundaries coincides with an IntP boundary:

(22) *Shó^{n}xti, wa'ú-thi^{n}, ní tható^{n}-bazhíi, wa'ú-thi^{n} uthúhai.*[31]
shó^{n}xti wa'ú-thi^{n} > ní tható^{n}-bazhíi < wa'ú-thi^{n} uthúhai <
[L H* L !H] [L H* L !H] [L !H* L L] [L H* L L] L%
nevertheless woman.the water drink.not.PROX woman.the follow.PROX
'Nevertheless, the woman, he didn't [stop to] drink water, he just followed the woman.'

[28] Lowie (1939)
[29] Square brackets in the ToBI line here refer to IntP boundaries rather than AccP boundaries as previously.
[30] ThigréPi in Goodtracks (2004)
[31] Rudin field tapes

The fact that *ní thatóⁿ-bazhíi* 'he didn't drink water' concludes with a phrase accent, despite not being notably "forward-looking", suggests that many IntP breaks may be arbitrary with respect to the narrow kind of information structure measured by my coding criteria, and conditioned by a more general, working-memory-related chunking process alongside other factors like weight, complexity and semantic unity. Generally, however, the recordings I have coded have the boundaries of links, relational topics and relational foci wherever a phrase accent occurs. The mapping tends to be 1–1, in that a single IntP tends to include a single relational category, but I suspect that in faster speech – of which I coded very little – speakers may stuff more than one relational category into a single IntP. When the referent of a string is a relational focus, it gets a L phrase accent, while links and relational topics tend to get !H or L phrase accents, and tend to weakly support Pierrehumbert & Hirschberg's (1990) generalization that !H phrase accents are reserved for more "forward-looking" material, albeit with a different threshold than in English.

3.5 Object-subject-verb (OSV) word order

Like postverbal arguments, OSV word order occurs in all the languages I included in this study, albeit with varying frequency. It is generally used when the object (O) is a link. Hidatsa speaker Annie Eagle says utterance (23) at a point in the narrative when the subject of the previous clause is 'our parents', so that *matawácʰo'* 'our relatives' is in a linking relation and is syntactically fronted:

(23) *Matawácʰo' maapúkšihtíawa šé'ri pʰéekšáwa.*[32]
matawácʰo' maapúkšihtíawa šé'ri pʰéekšáwa <
our.relative.PL.INDF snake.big.a that.ST eat.up.ITER.DS
'Our relatives are always being eaten by this big snake.'

Graczyk (1991: 102) presents Apsaalooke example (24) out of narrative context, so I have not coded it, but the linking relation between the two instances of *hawáte* 'one' is clear even out of context. The second instance of *hawáte* is a syntactically fronted object:

(24) *Hawáte isdáxxiia kulushkúam hawáte áxpe dappiíok.*
hawáte isdáxxiia kulushkúam hawáte áxpe dappiíok
one his.gun grab.from.DS one his.companions kill.PL.DECL
'One$_i$ of them, he$_j$ grabbed his$_i$ gun from him$_i$, and the other$_k$, his$_j$ companions killed him$_k$.'

[32] Parks, Jones & Hollow (1978: 3: 14)

Example (25), from speaker Francis Stewart, may be evidence that Apsaalooke allows OSV for other information-structural categories besides links. Although the clause *huulé kala kuluúkkuuk* has no overt subject, the position of the object before a time adverb indicates that a similar object-fronting process is at play. A linking poset relation like 'remnant of' may in fact be at play here, and the predicate 'bone' seems inferrable from the preceding word *chilishíak* 'they ate them', but the context did not meet my coding criteria for LINK, and so the form is not coded as linking:

(25) *Chilishíak huulé kala kuluúkkuuk huuk.*[33]
chilishíak huulé < kala > kuluúkkuuk < huuk
they.ate.them.ss bone now they.piled.them.up HEARSAY
'And after eating them, they piled their bones up.'

3.6 Switch-topic markers

Switch topics are nonrecoverable referents which function to shift hearers' attention away from currently recoverable referents. Although switch topics are typically new, they may be recoverable. They send hearers the signal, "enough of the old topic, pay attention to this now". This definition of "switch topic" subsumes a variety of special cases linguists often describe as "presentational focus" or "topic competition/resolution". Many Siouan languages have special morphological marks which signal some type of switch topic. Here I give a few Rų'éta examples and make some observations about Yesánq (Tutelo, not otherwise included in this study), Lakota and Nak^{h}on'i'a morphology. See also Boyle (2007: 288–293) for a syntactic analysis of the Hidatsa switch-topic marker *-ri*, which we saw in example (23). In that example, the snake, *maapúkšihtíawa šé'ri*, does not function to fill an open proposition like 'what's eating our relatives?', but rather is introduced as a new character for subsequent narrative.

Similarly, Wolvengrey (1991) analyses the Rų'éta suffix *-enq* as a "focus marker" which is used for presentation of new topics, as in example (26a), "atypical" subjects as in example (26b), and topic competition as in example (26c) from speaker Stephen Bird. The material marked by *-enq* in each of these examples functions to replace the current attention-central discourse topic with another, so I find it more useful to describe *-enq* as a switch-topic marker than as a "focus" marker. Note that in example (26c) the two characters marked with *-enq* are postverbal arguments with recoverable referents. This is a good case of how switch topics

[33] Wallace (1993: 192: 153)

are not always *new* topics, and how, though they have much in common with relational foci, they are not relational foci either.

(26) a. *Kanį mi·heną heromąko'š.*[34]
kanį mi·heną heromąko'š <
and woman.ST saw.DECL
'And he saw a woman.'

b. *Oreną tinį; napupušereka'ehe.*[35]
oreną tinį napupušereka'ehe <
fire.ST arrive.and burn.in.streaks.HEARSAY
'A prairie fire arrived and burned him in streaks.'

c. *Kašká'nįk inąk kimą·xero·mąko'š, kinąmą'kšise·ną. Káki "Mįkó'š" éhero·mąko'š, pą·xese·ną.*[36]
kašká'nįk > inąk < *kimą·xero·mąko'š,* *kinąmą'kšise·ną* káki >
be.DISJUNCT.that again ask.NARR.PST.DECL.M Coyote.the.ST be.that
"Mįkó'š" < *éhero·mąko'š* *pą·xese·ną*
no.DECL.M say.NARR.PST.DECL.M potato.the.ST
'But Coyote asked him again. And Potato said, "no".'

Lakota encodes switch-topicality grammatically. As a special case of contrast, switch topics may be marked with the *įš* set of independent personal pronouns as opposed to the non-contrastive *iyé* set. Nak^{h}on'i'a similarly uses the suffix *-įš* on switch-topical pronouns (Cumberland 2005: 129–130); and Oliverio (1996: 149) describes similar functions for Yesánq *-ma*, *-są* and *ik^{h}á-*.

Independent personal pronouns rarely referred to recoverable entities in the texts I coded. When they do, there is typically a repeated or conventional collocation at play. My study suggests that all of the nine Siouan languages in this study may observe this constraint on pronoun use. Dakotan languages further set aside a series of pronouns for use in referring to switch topics and other contrastive referents only. Other languages lack this mark. Yet despite these differences, in all nine languages the majority of independent personal pronouns in the texts I coded functioned as switch topics.

[34] Wolvengrey (1991: CWW9)
[35] Wolvengrey (1991: SA17)
[36] Carter (1991: 31: 15–16)

4 Discussion

The findings in §3 may be usefully sorted into two distinct categories. In §3.1 we saw that lack of a full [LH*L] contour signals referent recoverability. In §3.2 we saw this same phenomenon intersect with postverbal argument position in a way that usefully distinguishes between two constructions. In §3.3 we considered bare (undetermined) nominal expressions, noun incorporation and zero reference as stages on a continuum of phenomena variably constrained by recoverability and nonspecificity. I also suggested that the violent, colonial contact conditions under which recent language change is occurring may underlie a recent shift away from optionality and towards obligatoriness of determiner use and overtness of reference.

These first variations I term "prominence variations" – variations in which reduced variants are used for lighter (recoverable or nonspecific) referents. In the other category are "marking variations" – variations in which marked variants are used to refer to referents with marked information-structural functions.

In §3.4, I showed how the phrase accents which demarcate intonational intermediate phrases (IntP) tend to coincide with the boundaries of information-structural categories like LINK, RELATIONAL TOPIC and RELATIONAL FOCUS. !H phrase accents are reserved for "forward-looking" material – and by no means all of it. Specific information-structural categories are also associated with fronted objects and OSV word order, as we saw in §3.5, and with switch-topic markers like Rų'éta *-enąˀ* and Hidatsa *-ri*, as we saw in §3.6.

This distinction between prominence variations and marking variations maps roughly to Gundel's (2003; 1988) distinction between "relational givenness" and "referential givenness".

Although I have presented many of my findings as categorical generalizations when the data called for it, I caution readers away from assuming that any of these constraints really are categorically binding in all contexts in any one language, or in the family as a whole. More breadth and depth – more texts, more genres, more time periods, more languages; and more detailed, language-specific descriptions of documented variations with information-structural meaning – are required to be able to make any definitive statement on Siouan information structure, but I hope this sketch serves as a preliminary step towards imagining what such a statement might look like.

Abbreviations

BEN	benefactive	MOD	modal
CAUS	causative	NARR	narrative
CONT	continuative	NE	narrative ending
DECL	declarative	NPST	nonpast
DEF	definite	PL	plural
DS	different subject	PROG	progressive
INDF	indefinite	PROX	proximate
INFER	inferential evidentiality	PST	past
INSTR	instrumental	SS	same subject
ITER	iterative	ST	switch topic
M	masculine	TEMP	temporal progression

Primary resources

Carter (1991), Rų'eta

- Stephen Bird: Kinąmą'kšinį pą·xe (legend) – 74 utterances

Cumberland (2005), Nak^{h}on'i'a

- Bertha O'Watch: Snohéna T^{h}ąga (history) – 34 utterances
- Bertha O'Watch: Įktómi and Fox (legend) – 46 utterances
- The Red Fox (legend) – 27 utterances

Deloria (1932), Lakota

- multiple histories and legends

Dorsey (1880), Jiwere Ich^e

- The Rabbit and the Grasshopper (legend) – 1½ pages

Dorsey (1890; 1891), Umo^{n}ho^{n} Iye and Pa^{n}ka Iye

- multiple histories, legends, stories and orations – 915 pages

Dorsey & Swanton (1912), Taneksą

- multiple histories, legends and letters – 102 pages

Goodtracks (2004), Baxoje Ich^e and Jiwere Ich^e

- Tadá^{n}jeMi: Reminiscences of Grandmother (reminiscence) – 1½ pages
- ThigréPi: Being a present-day (1970's) Ioway (history/metaculture) – 1½ pages

Hartmann & Marschke (2010), Ho-Chunk

- Bill O'Brien: A bear appears (reminiscence) – 24 utterances
- Bill O'Brien: Bill O'Brien & Hollywood (reminiscence) – 31 utterances
- Bill O'Brien: The moccasin game (picture description) – 54 utterances
- Child teaching (history/metaculture) – 100 utterances
- Bill O'Brien & Chloris Lowe, Sr.: Horses (history) – 113 utterances
- Cecil Garvin: Connection (humour) – 14 utterances
- Chloris Lowe, Sr.: Buffalo hunt (history) – 18 utterances
- Ed Lonetree: Stealing watermelons (reminiscence) – 25 utterances
- Richard Mann: A warrior honor (reminiscence) – 46 utterances
- Richard Mann: Picking cherries (reminiscence) – 12 utterances

Ingham (2003), Lakota

- George Bushotter: Hunting eggs in the spring (reminiscence) – 1 page
- George Bushotter: How young men and women behaved towards each other among the People (reminiscence) – 2 pages
- George Bushotter: War customs (reminiscence) – 2 pages
- George Bushotter: Holy men and healers (reminiscence) – 1½ pages

Kennard (1936), Rų'eta

- untitled (legend) – uncounted utterances

Lowie (1939), Hidatsa

- First Worker Intrudes on Sun's Realm (legend) – 94 utterances
- First Worker Captures Geese But Loses Them to Spotted Tail (legend) – 68 utterances
- First Worker Captures Prairie Dogs But Loses Them to Spotted Tail (legend) – 71 utterances

- The Story of a Girl Who Became a Bear (legend) – 59 utterances
- The Water Buster Account (history) – 22 utterances

Marsh (1936), Baxoje Ich^e and Jiwere Ich^e

- MáñiHú – Twin Holy Boys (legend) – 14 pages
- Béñeiŋe (legend) – 11 pages
- Mishjiñe Aheri Warax^edhe (legend) – 5 pages
- HinágeSda^{n}: Udwá^{n}ge Mishjíŋe (legend) – 16 pages
- Wa^{n}^sígeChéMi: Hi^{n}kúñi (reminiscence) – 7 pages

Mixco (1997), Rų'eta

- Résike Wįke (legend) – 220 utterances

Parks, Jones & Hollow (1978), Hidatsa

- Annie Eagle: Packs Antelope (history) – 85 utterances
- Helen Wilkinson: The Return of Wolf Woman (history) – 152 utterances
- John Brave: Lone Man and First Creator Make the World (legend) – 87 utterances
- John Brave: Old Man Coyote and the Rock (legend) – 149 utterances

Rudin field tapes and transcripts, Umo^{n}ho^{n} Iye

- field tapes and transcripts of Catherine Rudin (elicitation), Microsoft Word .doc format with CD audio – 19 CD's

Wallace (1993), Apsaalooke

- Francis Stewart: Thunder Medicine (history) – 203 utterances

Yellow Brow & Short Bull (1980), Apsaalooke

- Yellow Brow & Short Bull: Bitáalasshia Alítchiasshiituualak Baháa Awúuasshiituualak (legend) – 41 utterances

References

Ariel, Mira. 1990. *Accessing noun-phrase antecedents* (Croom Helm Linguistics Series). London: Routledge.

Beckman, Mary & Janet B. Pierrehumbert. 1986. Japanese prosodic phrasing and intonation synthesis. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, 136–144. New York: Association for Computational Linguistics.

Boersma, Paul & David Weenink. 1992. *Praat: Doing phonetics by computer*. Software. http://www.praat.org.

Bolinger, Dwight. 1986. *Intonation and its parts: Melody in spoken English*. Stanford: Stanford University Press.

Boyle, John P. 2007. *Hidatsa morphosyntax and clause structure*. Chicago: University of Chicago (Doctoral dissertation).

Carter, Richard T., Jr. 1991. Old Man Coyote and the Wild Potato: A trickster tale. In H. Christoph Wolfart (ed.), *Linguistic studies presented to John L. Finlay* (Algonquian and Iroquoian Linguistics Memoirs 8), 27–43. Winnipeg: University of Manitoba.

Cumberland, Linda A. 2005. *A grammar of Assiniboine: A Siouan language of the northern Plains*. Bloomington: Indiana University (Doctoral dissertation).

de Reuse, Willem. 1994. Noun incorporation in Lakota (Siouan). *International Journal of American Linguistics* 60. 199–260.

Deloria, Ella. 1932. *Dakota texts* (Publications of the American Ethnological Society 14). New York: G. E. Stechert.

Dorsey, James Owen. 1880. The rabbit and the grasshopper: An Oto myth. *American Antiquarian and Oriental Journal* 3. 24–27.

Dorsey, James Owen. 1890. *The Ćegiha language, the speech of the Omaha and Ponka tribes of the Siouan linguistic family of North American Indians* (Contributions to North American Ethnology 6). Washington: Government Printing Office.

Dorsey, James Owen. 1891. *Omaha and Ponka letters* (Bureau of American Ethnology Annual Report 11). Washington: Government Printing Office.

Dorsey, James Owen & John Reed Swanton. 1912. *A dictionary of the Biloxi and Ofo languages, accompanied with 31 Biloxi texts and numerous Biloxi phrases* (Bureau of American Ethnology Bulletin 47). Washington: Government Printing Office.

Erteschik-Shir, Nomi. 2007. *Information structure*. Oxford: Oxford University Press.

Eschenberg, Ardis. 2005. *The article system of Umo^{n}ho^{n}.* Buffalo: State University of New York – Buffalo (Doctoral dissertation).

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

Goodtracks, Jimm G. 2004. *Text for Baxoje~Jiwere Ich^e – Ioway~Otoe language – CD II.* Lawrence, KS: Baxoje Jiwere Ich^e – Ioway Otoe-Missouria Language Project.

Gordon, Bryan James. 2008. Antitopic and afterthought: Givenness in the grammar in Ponca and Omaha. In Jie Zhang (ed.), *Proceedings of the 2007 Mid-America Linguistics Conference* (Kansas Working Papers in Linguistics 30), 74–87. Lawrence: University of Kansas. https://kuscholarworks.ku.edu/bitstream/handle/1808/3905/kwpl-30-Gordon.pdf.

Graczyk, Randolph Jan. 1991. *Incorporation and cliticization in Crow morphosyntax.* Chicago: University of Chicago (Doctoral dissertation).

Gundel, Jeanette Kuhn. 1988. Universals of topic-comment structure. *Studies in Syntactic Typology* 17. 209–239.

Gundel, Jeanette Kuhn. 2003. Information structure and referential givenness / newness: How much belongs in the grammar? In Stefan Müller (ed.), *Proceedings of the Head-Driven Phrase-Structure Grammar 2003 Conference*, 122–142. East Lansing, MI: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2003/gundel.pdf.

Gundel, Jeanette Kuhn, Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69. 274–307.

Gundel, Jeanette Kuhn, Mamadou Bassène, Bryan James Gordon, Amel Khalfaoui & Linda Humnick. 2010. Testing predictions of the givenness hierarchy framework: A cross-linguistic investigation. *Journal of Pragmatics* 42. 1770–1785.

Hartmann, Iren & Christian Marschke (eds.). 2010. *Texts with analysis and translation and an audio-cd of original Hocąk texts* (Hocąk Teaching Materials 2). Albany: State University of New York – Albany Press.

Ingham, Bruce. 2003. *Lakota* (Languages of the World / Materials 426). München: Lincom.

Kaufman, David V. 2008. Focality and topicality marking in Biloxi. *Kansas Working Papers in Linguistics* 30. 150–158.

Kennard, Edward A. 1936. Mandan grammar. *International Journal of American Linguistics* 9. 1–43.

Koontz, John E. 2003. *A focus marker in Omaha-Ponca (Dhegiha Siouan).* Paper presented at the Siouan and Caddoan Languages Conference.

Larson, Rory. 2009. *Pronunciation and prosody in Omaha*. Paper presented at the Siouan and Caddoan Languages Conference, University of Nebraska – Lincoln.
Lowie, Robert H. 1939. *Hidatsa texts (with grammatical notes and phonograph transcriptions by Zellig Harris and C. F. Voegelin)*. Vol. 1 (Prehistory Research Series 6). Indianapolis: Indiana Historical Society.
Marsh, Gordon H. 1936. *Materials for a study of the Ioway language*. Manuscript 4178. Philadelphia: American Philosophical Society.
Mirzayan, Armik. 2011. *Lakota intonation and prosody*. Boulder: University of Colorado (Doctoral dissertation).
Mithun, Marianne. 1999. *The languages of Native North America* (Cambridge Language Surveys). Cambridge: Cambridge University Press.
Mixco, Mauricio J. 1997. Mandan switch-reference: A preliminary view. *Anthropological Linguistics* 39. 220–298.
Oliverio, Giulia. 1996. *A grammar and dictionary of Tutelo*. Lawrence: University of Kansas (Doctoral dissertation).
Parks, Douglas R., A. Wesley Jones & Robert C. Hollow (eds.). 1978. *Earth lodge tales from the upper missouri*. Bismarck, ND: Mary College.
Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation*. Cambridge, MA: Massachusetts Institute of Technology (Doctoral dissertation).
Pierrehumbert, Janet B. & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Peter R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 271–311. Cambridge, MA: MIT Press.
Rood, David S. 1977. Some Lakhota presuppositions. *Colorado Research in Linguistics* 7. 110–124.
Rood, David S. 1985. Definiteness subcategorized in discourse: Lakhota *kˀų*. *Kansas Working Papers in Linguistics* 10. 144–183.
Rudin, Catherine. 1998. *Postverbal constituents in Omaha-Ponca*. Paper presented at the Siouan and Caddoan Languages Conference, Bloomington, IN.
Wallace, Karen K. 1993. *Verb incorporation and agreement in Crow*. Los Angeles: University of California – Los Angeles (Doctoral dissertation).
Ward, Gregory & Betty J. Birner. 2001. Discourse and information structure. In Deborah Schiffrin, Deborah Tannen & Heidi Hamilton (eds.), *Handbook of discourse analysis*, 119–137. New York: Wiley-Blackwell.
Wolvengrey, Arok. 1991. A marker of focus in Mandan discourse. In Frances Ingemann (ed.), *Proceedings of the 1990 Mid-America Linguistics Conference and*

Siouan and Caddoan languages conference, 584–598. Lawrence: University of Kansas.

Yellow Brow & Short Bull. 1980. *Bitaalásshia alítchiasshiituualak baháa awúuasshiituualak*. Mary Helen Medicine Horse (ed.). Crow Agency, MT: Bilingual Materials Development Center.

Chapter 17

NP-internal possessive constructions in Hoocąk and other Siouan languages

Johannes Helmbrecht

Languages usually have more than one construction to express a possessive relationship. Possessive constructions in an individual language usually express semantically different relations, which are traditionally subsumed under the notion of possession such as part-whole relationships, kinship relationships, prototypical ownership, and others. Hoocąk and the other Siouan languages are no exception from this many-to-many relationship between possessive constructions and semantic kinds of possession. The present paper deals with NP-internal types of possession in Siouan languages leaving aside constructions that express possession on the clause level such as benefactive applicatives, reflexive possessives and the predicative possession. The NP-internal possessive constructions will be examined according to the semantic/syntactic nature of the possessor (regarding the Animacy Hierarchy), and the semantic nature of the possessed (alienable/inalienable distinction). I will begin with an analysis of Hoocąk and will then compare the Hoocąk constructions with the corresponding ones in some other Siouan languages. At least one language of each sub-branch of Siouan will be discussed. It will be shown that the choice of different NP-internal possessive constructions depends on both semantic scales (the Animacy Hierarchy and the alienable/inalienable distinction), but in each Siouan language in very individual ways.

1 The structure of NP-internal possessive constructions

It may safely be assumed that all languages have grammatical and lexical means to express a possessive relation between an entity A and an entity B. Semantically, possession is a cover term for a broad range of distinct relations, which are expressed by possessive constructions (PC) in the languages of the world. Central to the notion of grammatical possession are the relations of ownership, whole-part relations, and kinship relations. Less central to the general notion of possession are attribution of a property, spatial relations, association, and perhaps

Johannes Helmbrecht. 2016. NP-internal possessive constructions in Hoocąk and other Siouan languages. In Catherine Rudin & Bryan J. Gordon (eds.), *Advances in the study of siouan languages and linguistics*, 423–461. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b94.180 DOI:10.17169/langsci.b94.180

nominalization. All these relations may be expressed by NP-internal possessive constructions in English as exemplified in Table 1.[1]

Table 1: Semantics of possessive relation in the broad sense (cf. Dixon 2010: 262–267)

Entity A	Possessive relation	Entity B	English example
Possessor	⟷	**Possessed**	
	Ownership		my car/ Peter's house
	Whole-part		Mary's teeth/ the teeth of the bear
	Kinship		Peter's wife/ my daughter
	Attribution of property		her sadness/ his age
	Spatial relation		the front of the house/ the inside of the church
	Association		Jane's teacher/ her former school

Not all of the different kinds of relations in Table 1 can be expressed by possessive constructions in all languages, but in most cases ownership, whole-part, and kinship relations are covered by their NP-internal PCs. It is still an open question whether there is a general semantic notion of possession that covers all relations expressed by PCs. There is at least one prominent approach to possession which claims that there is a semantic prototype with a core and a periphery (cf. Seiler's prototype approach (Seiler 1983; 2001); and a critical examination of it in Helmbrecht 2003). Others reject this idea (cf. for instance Heine 1997; Dixon 2010: 263 and others).

Languages usually have more than one syntactic construction expressing possessive relations on the clause level as well as on the NP level. In Tables 2 and 3, there are examples of different possessive constructions from Hoocąk,[2] English and German for illustrative purposes.

The present paper deals only with PCs on the NP level. Languages often possess more than just one NP-internal PC, as is the case for instance in English. English has the *of*-construction and the genitive *=s* construction to express possession NP-internally; similarly for German. If there are two or more NP-internal PCs in a language, the choice of these constructions often depends on the semantic and syntactic category of the possessor and/or the semantic type of the possessed entity.

[1] See Dixon (2010: 262–267) for a more detailed discussion of these relations.

[2] Hoocąk, formerly also known as *Winnebago*, is a Siouan language still spoken in Wisconsin. Hoocąk together with Otoe, Ioway, and Missouria forms the Winnebago-Chiwere sub-branch of Mississippi-Valley Siouan. For the widely accepted classification of Siouan languages see, for instance, Rood (1979), Mithun (1999: 501), and Parks & Rankin (2001).

Table 2: A brief typology of possessive constructions (part 1, clausal)

level	construction type	examples
Clause	predicative possession	English *I **have** a blue car.* *The blue car **belongs** to me.*
	external possession[a] (possessor raising)	Hoocąk (BO979) *Huuporo=ra **hį**-teek-ire …* knee=DEF **1E.U**-hurt-SBJ.3.PL 'When my knees hurt, … '
	dative of interest	German *Sie schneidet **ihm** die Haare* she cuts **him.DAT** the hair 'She cuts his hair.'
	beneficiary	Hoocąk (Helmbrecht 2003: 28) *Wažątírera hįįgí'eenąk.* *wažątíre=ra hi<**hį**-**gí**>'e=nąk* car=DEF <**1E.U-APPL.BEN**>find=DECL 'He found the car for me.'/ 'He found my car.'
	-possessor polysemy	
	possessive	Hoocąk (Helmbrecht & Lehmann 2010) *Hinįk=ra nąąk<**kara**>t'ųp='anąga* son=DEF <**POSS.RFL**>embrace(SBJ.3SG&.OBJ3SG)=and 'He (i.e. the father) embraced his son, and …'
	reflexive	

[a] See, for instance, Payne & Barshi (1999) on types of external possession.

Table 3: A brief typology of possessive constructions (part 2, non-clausal)

level	construction type	examples
NP	juxtaposition (no marking at all)	Hoocąk (Helmbrecht 2003: 13) *Peter=gá šųųk=rá* ‘Peter’s dog’ Peter=PROP dog=DEF
	genitive attribute (genitive case marker on possessor)	English *Peter**’s** dog*
	prepositional attribute	German *der Hund **von** Peter* DEF dog **of** Peter ‘Peter’s dog’
	pronominal index on possessed noun (possessor marking on possessed)	Mam (England 1983: 142) ***t**-kamb’ meenb’a* 3SG-prize orphan ‘the orphan’s prize’
	mixed strategy (genitive case marking plus pronominal index)	Turkish (Kornfilt 1990: 633) *Ayşe-**nin** araba-**sı*** Ayşe-GEN car-3SG ‘Ayşe’s car’
	nominalized predicative possession	Hoocąk (Helmbrecht 2003: 19) hicųwį́ waháara *hicųwį́ **wa-háa=ra*** aunt obj.3PL-have.kin(1E.U)=DEF ‘my aunts’
Word	nominal compounds	German *das Regierungsauto* *das Regierung-s-auto* DEF government-LINKER-car ‘the car of the government’

With regard to the possessor, the choice of the PC may depend on the specific NP type of the possessor. For instance, if the possessor is expressed by means of a possessive pronoun a different construction may be used than with a possessor expressed by a lexical noun phrase. If the possessor is a proper name or kinship term this may determine the selection of a specific PC, too. Animacy proper of the possessor, i.e. possessor NPs with a human, animate or inanimate common noun, may play a role as well. The implicational scale that brings together these different NP-types that may be relevant for the choice of different NP-internal PCs in Siouan languages as well is called Animacy Hierarchy (AH). The AH is a scale that describes many different grammatical phenomena cross-linguistically. The AH is usually considered as: 1/2 > 3 > proper noun/kin term > human common noun > animate common noun > inanimate common noun (cf. for instance Dixon 1979; Comrie 1981; Croft 2003).

With regard to the possessed, it can be observed that the choice of the NP-internal PC depends on the semantic class of the possessed, i.e. these languages often have two sets of nouns, so that set_1 nouns designating the possessed entity require one type of the PC, and set_2 nouns designating the possessed entity the other. This classification of nouns with regard to PCs is better known under the heading alienable versus inalienable distinction. Alienable nouns usually designate entities that can be owned in the prototypical sense implying that the possessor has full control over these possessed entities; for instance the possessor can sell them, give them away, and so on. The class of inalienable nouns is much more heterogeneous with regard to the semantics; inalienable nouns designate entities that bear a close association to the possessor implying that the possessor has no or only a limited control over them. Often, kinship terms, body-part terms, and other relational nouns (local/spatial nouns) belong to this class. With regard to the formal marking of the respective PCs, the following possibilities can be distinguished (cf. Dixon 2010: 286–290):

i. the alienable PC is similar to that for inalienable possession with an added grammatical element;

ii. the grammatical marking for alienable possession is longer than that for inalienable possession;

iii. the alienable PC requires a classifier, the inalienable construction does not;

iv. overt marking only in an alienable PC

The possibilities i-iv cover the cross-linguistic observation that inalienable PCs tend to be shorter and morphological less complex than alienable PCs. In other words, the PC for alienable possession is always more marked than the PC for inalienable possession. In what follows it will be shown that this observation also holds in general for the different NP-internal PCs in Siouan.

2 Methodical remarks

The goal of this study is to search for all different NP-internal PCs in selected Siouan languages and to describe the conditioning factors for their choice. The guiding hypothesis is that the syntactic/semantic properties of the possessor (Animacy Hierarchy) and the semantic properties of the possessed (alienable vs. inalienable) is a fruitful notional frame for the discovery and the description of the splits in the expression of possession; cf. Table 3.

Typological studies on possession show that the properties of the possessive relation itself such as actual possession vs. possession in the past, temporary possession vs. permanent possession, close possession vs. loose possession and so on, may trigger a constructional split too, in some languages (cf. Dixon 2010: 274–276). This is, as far as I can see, not the case in Siouan languages. Therefore, these semantic parameters won't play a role in the rest of the paper.

Table 4: Semantic/syntactic parameters for constructional splits in NP-internal possessive constructions.

Semantic-syntactic properties of the possessor (Animacy Hierarchy)	Semantic properties of the possessive relation	Semantic properties of the possessed (alienable - inalienable)
pronoun (SAP/3rd person)	temporal/ closeness	A) ownership
proper name	temporary/ permanent	B) whole-part relation
kinship term	close/ loose	C) kinship relation
common noun	general type of possession	D) attribution
[human]		E) orientation/ location
[animate]		F) association
[inanimate]		G) nominalization

The properties of the possessor and the possessed as summarized in Table 4 serve as a kind of questionnaire or guideline for the search for constructional splits in the various grammatical descriptions of Siouan languages that are used

here. The data and descriptions of PCs are taken from the grammars that are available for the different Siouan languages. For Hoocąk, data from a text corpus and from fieldwork sessions will be taken.

I will exclude the question of the relation between NP-internal PCs and the clause level PCs for later research. My own experience with text data from Hoocąk makes me think that clause level PCs are often preferred over the NP-internal PCs at least in Hoocąk, but this needs to be shown in more detail.

3 NP-internal possession in Ho-ChunkHoocąk

Hoocąk and Chiwere (Missouria, Otoe, Ioway) are closely related and constitute the Winnebago-Chiwere sub-branch of the Mississippi Valley group of the Siouan languages. Hoocąk is taken as a representative of this sub-branch, then.

Hoocąk has no possessive pronouns comparable to English *my, your, his, her*, etc., no nominal case marking in general, and no genitive case marker in particular. In addition, there are no connectives, linkers or possessive markers, i.e. grammatical forms that indicate a possessive relation between two nominals. Hoocąk has in principle two different types of NP-internal PCs. The first one is a simple juxtaposition of two nouns without any special possessive marking, see example (1) below. The second type of PC is a complex construction with an inflected verb of possession, e.g. *=hii* 'have.kin' plus a definite article nominalizing the entire construction illustrated in example (1b). Without this definite article, we have a clause expressing a kind of predicative possession.

(1) Hoocąk Helmbrecht 2003: 16

a. *Petergá šųųkrá*
Peter=gá šųųk=rá
P.=PROP dog=DEF
'Peter's dog'

b. Hoocąk Helmbrecht 2003: 19
hicųwį́ wahaará
hicųwį́ wa-haa=rá
aunt OBJ.3PL-have.kin(1E.U)=DEF
'my aunts'

Both types of NP-internal PC will be discussed in more detail in the subsequent sections.

3.1 Juxtaposition

The semantic/syntactic type of the possessor does not require the choice of the juxtaposition PC in Hoocąk with one exception. If the possessor is a speech act participant or a third person, the second construction type with the nominalized possessive verb has to be chosen obligatorily (see §3.2 below).

The following series of examples demonstrates that neither the AH – except with regard the pronoun/noun distinction – nor the distinction between alienable vs. inalienable nouns have any effect on the expression of attributive possession in Hoocąk. The example in (2) is an attributive possessive relation with a proper name as possessor and a kinship term as possessed noun. The relation is inherent and inalienable. The definite article is required.

(2) Hoocąk (Helmbrecht 2003: 16)
Petergá hi'ącrá
Peter=gá hi'ąc=rá
P.=PROP father=DEF
'Peter's father'

The possessive relation in (3a) is a part-whole relationship with a human possessor and a body-part term as possessed. The possessive relation is inherent and inalienable. The same holds for the examples in (3b)-(3c). The whole PC needs to be specified by a determiner, i.e. the definite article, or a demonstrative pronoun. If there is a definite article following the possessor (cf. (3b)), then it is the possession of a specific and definite possessor. If the indefinite article follows the possessor (cf. (3c)), it is the possession of an indefinite possessor.

(3) Hoocąk (Helmbrecht 2003: 13)
a. *hinúk hišja=rá*
woman face=DEF
'the woman's face'
b. *hinúk=**rá** hišja=rá*
woman=**DEF** face=DEF
'the face of the (specific/definite) woman'
c. *hinúk=**ízą** hišja=rá*
woman=**INDEF** face=DEF
'the face of an (indefinite) woman'

The PCs in (4) and (1) (above) are alienable. Both contain alienable possessed nouns, the inanimate noun *hiráati* 'car' and the animate noun *šųųk* 'dog'. The possessor is a human being (proper name) in both cases.

(4) Hoocąk (Helmbrecht 2003: 13)
John=ga hiráati=ra
J.=PROP car=DEF
'John's car'

The possessive relation in (5) includes a body-part term as possessed noun (inseparable, inalienable) with a non-human possessor. The example in (6) represents a part-whole relation with an inanimate object as possessor and an inanimate object as possessed (separable, alienable). Both possessors in (5) and (6) can be interpreted either as specific or as generic.

(5) Hoocąk (Helmbrecht 2003: 13)
wijúk huu=rá
cat leg=DEF
'the leg(s) of the/a cat'

(6) Hoocąk (Helmbrecht 2003: 13)
važątíre hogis=rá
car circular.part=DEF
'the wheel(s) of a/the car'

Note that the constructions in (5) and (6) often resemble a nominal compound with the first noun specifying the second noun thus creating a new word and concept instead of expressing a possessive relation. For instance, the Hoocąk word *nąąhá* 'bark' is a compound of the noun *nąą* 'tree' and *haa* 'skin, pelt, hide' thus giving the new concept 'tree skin' which corresponds to 'bark' in English. This combination of two nouns is a nominal compound on phonological grounds. The vowel in the second noun is shortened, which is a normal word-internal process in Hoocąk. However, the boundary between compound and juxtaposition is often blurred and the function "the first noun specifies the second" can be found in phrasal juxtaposition as well as in nominal compounding. The expressions in (5) and (6) are certainly phrasal in nature. Both words in these expressions have their own primary accent and there are no sandhi processes between the two nouns.

The same type of construction employed for the expression of possession in the preceding examples is also used for the expression of spatial relations. There

are numerous local nouns such as *coowé* 'front part', *nąąké* 'back part', *rook* 'inside', *hihák* 'top, surface', and so on, which are used to express the specific local/spatial relation of an object vis-à-vis the spatial region of another object. The local nouns are the possessed nouns in these constructions. They designate the spatial position of the possessor. The possessor functions as the reference point (cf. Langacker 1993) of the localization, it represents the object with regard to which another one is localized, cf. the examples in (7). The clitic *=eja* 'there' is a local adverb almost obligatorily used in these constructions.

(7) Hoocąk (Helmbrecht 2003: 14)

a. *šųųkrá hirarúti coowéja ʻakšąną*
šųųk=rá **hirarúti coow=éja** ʻak=šąną
dog=DEF **car** **front=there** be.lying=DECL
'The dog is (in a lying position) in front of the car.'

b. *šųųkrá hirarúti hihákeja jeenąą*
šųųk=rá **hirarúti hihák=eja** jee=nąą
dog=DEF **car** **top=there** be.standing=DECL
'The dog is (in a standing position) on the top of the car.'

c. *šųųkrá hirarúti rookéja nąkšąną*
šųųk=rá **hirarúti rook=éja** nąk=šąną
dog=DEF **car** **inside=there** be.sitting=DECL
' The dog is (in a sitting position) inside of the car.'

The expressions in examples (2) through (7) show that the semantic nature of a lexical possessor does not trigger a shift to another construction type: this holds if the possessor is a proper name (=*ga* PROP), human noun (=DEF/=INDEF/=Ø), animate noun (=DEF/=INDEF/=Ø), or inanimate noun (=DEF/=INDEF/=Ø). In addition, the expressions in (2) through (7) show that there is no alienable-inalienable distinction: the same construction type is chosen with kinship terms, body-part terms, relational spatial nouns, as well as with alienable nouns. The possessor noun may be marked by a definite (DEF), an indefinite (INDEF) article, or by zero. If the possessor is a proper name (PROP), it will be marked by the proper name marker. The entire PC is always definite (DEF) (marked on the possessed noun) except with spatial nouns. They are usually marked by means of a local adverb clitic *=eja* 'there' which – in this respect – could also be analyzed as a general local postposition. The examples also show that this type of PC may express real ownership, part-whole relations, kinship relations, and spatial relations.

3.2 Nominalized verbal possessive constructions

The juxtaposition of two nominals is a general construction type to express possession and other binary relations in Hoocąk. There is, however, an alternative NP-internal possessive construction, which indeed exhibits a classification of nouns: inalienable nouns such as kinship terms, domestic (pet) animals, and alienable nouns. These alternative constructions are in each case a nominalized version of the possessive predication employing different possessive verbs for different types of possessed entities. The nominalized possessive clauses appear in the same syntactic position as the juxtaposed nouns, i.e. in a noun phrase position of the clause.

(8) Hoocąk (Helmbrecht 2003: 16)

a. ***Ɉohn=gá hiráati=ra*** *hacáa=nąk*
J.=PROP car=DEF see(1E.A&OBJ.3SG=DECL
'I see John's car.'

b. ***Ɉohn=gá hiráati hanį=rá***
J.=PROP car own(SBJ.3SG&OBJ.3SG)=DEF
hacáa=nąk
see(1E.A&OBJ.3SG)=DECL
'I see John's car.'

Both clauses in (8) have the same translation, but speakers indicate that they prefer the nominalized variant over the juxtaposed variant. The same constructional pairs exist for possessive constructions with kinship terms and pet animals (domestic animals). These nominalized possessive clauses represent a kind of transition from attributive to predicative possession. The general structure of these nominalized possessive clauses is given in (9).

(9) General structure of the nominalized verbal possessive construction
[(N-POSSESSOR$_i$) N-POSSESSED$_j$ PRO$_j$-PRO$_i$-Verb of possession=DET]$_{NP}$

If the possessor is a speech act participant or third person, these nominalized PCs are the only possibility. Since the possessor is often a topic (given and definite) in discourse and hence expressed pronominally as a 3SG, this type of PC prevails in discourse over the alternative juxtaposition. Note that 3SG arguments are always marked zero on the verbs. Both entities $X_{possessor}$ and $Y_{possessed}$ are cross-referenced in the verb of possession utilizing the two different series of pronominal prefixes, the actor/subject series for the possessor and the undergoer/object

series for the possessed. The verbs of possession are treated as regular transitive verbs.

If the possessor is a lexical human noun, this construction type competes with the juxtaposition type of PCs dealt with in the preceding section; cf. the following examples in (10).

(10) Hoocąk (Helmbrecht 2003: 16)

a. *Peterga hi'ąc hiirá*
Peter=ga hi'ąc **hii**=ra
P.=PROP father have.kin=DEF
'Peter's father'

b. *Peterga šųųk nįįhíra*
Peter=ga šųųk **nįįhí**=ra
P.=PROP dog have.pet=DEF
'Peter's dog'

c. *John=gá hiráati **hanį**=rá*
J.=PROP car own=DEF
'John's car'

The verbs of possession that are used in the PCs in (10) are restricted in their usage. The verb *=hii* 'X has Y as kin' can only be used with kinship terms or with terms designating close friends. This verb is homophonous with the causative auxiliary *=hii* 'to cause'. There are reasons to believe that both verbs are historically cognate, and that they should be considered as different usages of one verb rather than homonyms. The main reason for this analysis is that the causative verb *=hii* has an irregular personal inflection, and the possessive verb *=hii* shows exactly the same pattern.

The possessive verb *nįįhí* 'X has Y as pet' is used only with pet animals. Usually, pet animals are domesticated animals such as cats, and dogs, etc. The semantic boundaries of this class are not clear-cut. Historically, *nįįhí* is presumably a combination of **nį* 'to live, living thing', which does not occur independently in Hoocąk and the causative auxiliary *=hii*.[3] The verb *nįįhí* shows the same inflectional irregularities as the causative verb *=hii*.

The possessive verb *hanį* 'to own' is a regular (lexical) transitive verb designating the possession of alienable entities such as inanimate objects, artifacts, ani-

[3] **nį* is the reconstructed Proto-Mississippi-Valley Siouan form for 'live, be alive' (cf. Rankin et al. 2015). This form can be found in other verbs in Hoocąk such as *nįįhá* 'to breathe' or in *nįį'ąp* 'be alive'.

mals, and so on. Body parts belong to this group of nouns, too. It is restricted to human possessors. Part-whole relations with inanimate possessors, on the other hand, are never expressed with this construction. Cf. the summary in Table 5.

Table 5: Alienable vs. inalienable distinction in Hoocąk

inalienable/ inseparable		alienable/ separable
set_1: =híi	set_2: nįįhí	set_3: hanį́
kinship (including close social relations such as friendship)	pet animals (usually domestic animals such as dog, cat, horse, etc.)	animate and inanimate objects such as non-domestic animals, artifacts, and so on including body parts

All three verbs in Table 5 form the same type of nominalized verbal PC with pronominal and lexical human possessors. There is no difference between them with regard to structural markedness or with regard to the iconic relationship observed for the inalienable vs. alienable distinction and the size of the corresponding PCs. The paradigms for all three verbs of possession are given below; cf. Table 6, Table 7, and Table 8. The paradigms contain only constructions with a 3SG possessed noun. If the possessed nouns were plural ('aunts', 'dogs', and 'cars') the verbs of possession would be inflected for the third person plural object (*wa-* OBJ.3PL).

Table 6: Paradigm of the possessive verb *hii* 'to have kin'

possessor	possessed N *hicųwį́*	
1SG	hicųwį́ haa=rá	'my aunt' (father's sister)
2SG	hicųwį́ raa=ra/=gá	'your aunt'
3SG	hicųwį́ hii=rá	'his aunt'
1I.D	hicųwį́ hįhi=rá /=ga	'my and your aunt'
1I.PL	hicųwį́ hįhiwí=ra	'our aunt'
1E.PL	hicųwį́ haawí=ra	'our aunt'
2PL	hicųwį́ raawí=ra/=ga	'your aunt'
3PL	hicųwį́ hiíre=ra	'their aunt'

The kinship term *hicųwį́* 'aunt (father's sister)' has a variant form that is used for address purposes, *cųwį́* '(my) aunt!'. These address forms of kinship terms

Table 7: Paradigm of the possessive verb *nįįhi* ‘to have pet’

possessor	possessed N *šųųk*	
1SG	šųųk nįįháa=ra	‘my dog’
2SG	šųųk nįįná=ra/=ga	‘your dog’
3SG	šųųk nįįhí=ra	‘his dog’
1I.D	šųųk nį́įhi=ra/=ga	‘our dog’
1I.PL	šųųk nįįháwi=ra	‘our dog’
1E.PL	šųųk nį́įhiwi=ra	‘our dog’
2PL	šųųk nįįnáwira/=ga	‘your dog’
3PL	šųųk nįįhíre=ra	‘their dog’

Table 8: Paradigm of the possessive verb *hanį́* ‘to have’

possessor	possessed N *važątíre*	
1SG	važątíre haanį́=ną[a]	‘my car’
2SG	važątíre hašįnį́=ną	‘your car’
3SG	važątíre hanį́=ną	‘his car’
1I.D	važątíre hįįnį́=ną	‘our car’
1I.PL	važątíre hįįnį́wį=ną	‘our car’
1E.PL	važątíre haanį́wį=ną	‘our car’
2PL	važątíre hašįnį́wį=ną	‘your car’
3PL	važątíre hanį́įne=ra	‘their car’

[a] There are two phonological rules in Hoocąk a) that underlying /r/ becomes [n] after nasal vowels and b) that oral vowels are nasalized after nasal consonants. Sometimes rule a) is indicated orthographically by a hačekcaron <ň>.

— often simply lacking the initial syllable *hi-* — cannot occur in a possessive construction. This seems to be a general rule for obvious reasons. The usage of kinship terms as address terms usually presupposes that such a kinship relation holds between speaker and hearer.

There is another kind of variation in the paradigm of kinship possession that may be rooted in the mutual knowledge of the interlocutors. The common determiner in possessive constructions with a kinship term is the definite article *=ra*. However, in the second person singular and plural the determiner is *=ga*, a deictic element also used for the indication of proper names. Lipkind claims that *=ga* has to be used exclusively in these instances (cf. Lipkind 1945: 31), but Hoocąk speakers gave me forms that show that there is actually a choice between *=ra* and *=ga* in the second person and in the first person inclusive dual form;[4] *=ga* is ungrammatical in all other person categories. One of my most important language consultants, Phil Mike, indicated to me that this choice has to do with the mutual knowledge of the kinsman by both interlocutors. The definite article is used in the second person if the speaker does not know the kinsman (assuming that the hearer knows his or her kinsmen), but *=ga* is used when both interlocutors know the person talked about (which is more naturally the case if the speaker talks about the kinsman of the hearer). This could also explain why *=ga* is not allowed if the possessed is plural. The deictic suffix *=ga* is also used with the address forms of kinship terms indicating the first person as possessor. Lipkind (1945: 31) says that all kin terms with initial *hi-* take *haará* 'my' in the first person; the few ones without it take solely *=ga* instead; the reason is that the shorter forms are terms of address while the *hi-* forms are terms for reference. For instance, the form *cųwį́* is the address term corresponding to *hicųwį́* 'aunt (father's sister)'. Hence the 1SG possessive form is *cųwį́-gá* which translates literally 'that aunt' implying that everybody knows that she is the aunt of the speaker (EGO). It is a kind of reduced form of speaking. The address term implies that the person so addressed has the kin relation designated by the term toward the speaker. It is an effect of the Animacy Hierarchy. Shared background knowledge of the possessor plays an important role here (cf. also Heine 1997: 26f).[5]

[4] I am particularly grateful to Henning Garvin helping me to collect the relevant forms here.

[5] This can also be interpreted as an instance where the inherent relationality of kin terms leads to a structural reduction of the expression of possession confirming the prediction of the prototype approach.

4 Constructional splits in the other Siouan languages

In what follows a few other Siouan languages are examined with regard to constructional splits that have to do with the NP type of the possessor and the semantics of the possessed. I will begin with the Northwestern Siouan languages Crow, Hidatsa, and Mandan (§4.1–4.3), then I will continue with Lakota (the Dakotan sub-branch of Mississippi-Valley Siouan; §4.4) and Osage (Dhegiha sub-branch of Mississippi-Valley Siouan; §4.5), and I will close this investigation with Biloxi as a representative of the South-Eastern branch of Siouan (Ohio-Valley Siouan; §4.6).

4.1 Crow

4.1.1 The possessor

Crow has four different NP-internal PCs depending on the semantic/syntactic nature of the possessor; cf. the examples in (11) through (14).

(11) Crow (Graczyk 2007: 234)
  a. [Poss.Pro — $N_{possessed}$]
  b. *Ø-iilápxe*
     3SG.POSS-father
     'his father'

(12) Crow (Graczyk 2007: 234)
  a. [$N_{possessor}$(-DET/-Ø) Poss.Pro — $N_{possessed}$]
  b. *Charlie-sh Ø-iilápxe*
     C.-DET 3SG.POSS-father
     'Charlie's father'

(13) Crow (Graczyk 2007: 235)
  a. [Emphatic PRO-POSS.PRO-$N_{possessed}$]
  b. *bii-w-achuuké*
     1SG.EMPH-1SG.POSS-younger.brother
     'MY younger brother'

(14) Crow (Graczyk 2007: 236)
  a. [[[$N_{possessor}$] [$N_{possessor}$]] [$N_{possessed}$]]

b. *úuxbishke* *chíis-uua* *iía*
white.tailed.deer tail-PL hair
'hair from the tail of the white-tail deer'

No matter whether the possessed noun is alienable or inalienable, there has to be a possessive pronoun attached to the possessed noun indicating the possessor (cf. example (11)). The same is true if there is a lexical possessor in addition (cf. example (12)). The possessive prefix may be emphasized by means of a bound emphatic pronoun prefixed to the possessive prefix (cf. example (13)). Interestingly, there are also PCs that do not show any possessive marking and hence look like a juxtaposition expressing a whole-part relationship, cf. the example in (14). I did not find more examples like this in Graczyk's grammar, so I cannot say if this is generally an alternative possibility or required for non-human possessors.

4.1.2 The possessed

Crow has different paradigms of proper bound possessive pronouns distinguishing different sets of possessed nouns according to the alienable versus inalienable distinction. The paradigm of possessive pronouns for alienable possession is given in Table 9; the paradigm of inalienable possession is given in Table 10.

Table 9: Alienable possession in Crow (Graczyk 2007: 53)

	stem *íilaalee*	
1SG	**ba-s-**íilaalee	'my car(s)'
2SG	**dí-s-**iilaalee	'your car(s)'
3SG	**i-s-**íilaalee	'his/her car(s)'
1I.PL	**balee-is-**íilaalee	'our car(s)'
1E.PL	**ba-s-**íilaalee-**o**	'our car(s)'
2PL	**dí-s-**iilaalee-**o**	'your car(s)'
3PL	**i-s-**íilaalee-**o**	'their car(s)'

The possessive pronouns of alienable possession in Table 9 are formally invariable; they have an additional /-*s*/ thus being phonologically more marked than the prefixes of the inalienable paradigm. The 2SG possessive pronoun of the alienable paradigm shows a shift of the primary stress from the stem to the prefix, a pattern which is found also in some of the active verb paradigms. The 1I.PL prefix *balee-* is taken from the B-set pronominal paradigm for stative verbs. This

Table 10: Inalienable possession in Crow (Graczyk 2007: 52)

	stem *apá*	
1SG	**b**-apé	'my nose'
2SG	**d**-ápe	'your nose'
3SG	**Ø**-apé	'his/her nose'
1I.PL	-	-
1PL	**b**-ap-**úua**	'our noses'
2PL	**d**-áp-**uua**	'your noses'
3PL	**Ø**-ap-**úua**	'their noses'

form is added to the 3SG.POSS *is-* prefix, probably a late innovation introducing a 1PL inclusive-exclusive distinction into the alienable paradigm. This distinction is lacking in the inalienable paradigm of possessive pronouns as well as in the verbal paradigms. The suffixes in both paradigms (*-o* in the alienable possessive paradigm, *-úua* in the inalienable possessive paradigm) indicate the plurality of the possessor.

The paradigm of inalienable possession varies in form depending on the stem-initial sounds. There are three phonologically conditioned allomorphic paradigms, for stems in /d-/, /i+consonant-/, and /vowel-/. As can be seen in Table 10, the stem itself also undergoes some sound changes.

There are, however, three additional paradigms of inalienable possession: a) one that marks possession with the undergoer series of pronominal prefixes (called B-set of pronominal prefixes in Graczyk's grammar), b) one with an irregular paradigm, and c) one residual paradigm that shows stem suppletion. Graczyk (2007: 57) finds the following classification of nouns associated with these three different inalienable paradigms.

a) Inalienable possession with the B-set prefixes is used with nouns referring to internal body parts such as 'gland', 'joint', 'limb', 'hip', 'bone', 'lung', 'stomach', etc. (cf. Graczyk 2007: 57).

b) There are not enough nouns requiring the irregular paradigm for a semantic classification, but they all seem to belong to the inalienable class of nouns.

c) The nouns that require suppletive stems refer to kinship relations, clothing, and some culturally important possessions, cf. the examples in Table 11.

The first column shows the nouns in citation form, the second column in a possessive construction. The corresponding stems are clearly suppletive.

Table 11: Suppletive stems in Crow (Graczyk 2007: 58)

ihkáa	'mother'	is-ahká	'his mother
huupá	'shoe'	is-ahpá	'his shoe'
alúuta	'arrow'	is-aá	'his arrow'
buú	'song'	is-huú	'his song'

There is also a prefix *bale-* that is used if inalienable nouns are used without indicating a possessor. This form is called depossessivizer in Graczyk (2007: 53/234) and it is obligatorily used with unpossessed body-part nouns. This form is not used with kinship terms.

Table 12 summarizes the findings with regard to the alienable/inalienable distinction. Inalienable nouns are a closed class of nouns in Crow. It is clear that the semantic classification of the nouns with regard to the different PCs is not sharp. There are even body-part nouns that belong to the alienable class (set_5). Gross modo, however, the nouns in set_1 - set_4 could be subsumed under a class of inalienable nouns semantically.

4.2 Hidatsa

4.2.1 The possessor

Hidatsa and Crow are closely related and constitute the Missouri Valley subbranch of Siouan. Although they belong to the same sub-branch of Siouan, there are differences in the expression of possession. Hidatsa has different PCs depending on the syntactic/semantic type of the possessor. As in Crow, there is an obligatory marking of the possessor on the possessed noun no matter whether the possessed noun is alienable or inalienable; cf. the alienable PC in (15b). If there is an additional lexical possessor, the structure of the PC in Hidatsa is analogous to the one in Crow, cf. the alienable PC in (15a).

Table 12: Alienable vs. inalienable distinction in Crow

inalienable				alienable
set_1	set_2	set_3	set_4	set_5
phonologically conditioned inalienable paradigm	B-set prefixes	irregular paradigm	suppletive possessed forms	alienable paradigm
body parts, kinship	closed class of nouns referring to internal body parts	'chest', 'tail', 'husband'	closed class of nouns referring to objects closely associated to a person (e.g. clothing, a few kin terms, culturally important possessions)	open class of nouns not inherently possessed; exceptions are: *huli* 'bone', *íili* 'blood', *kahkahká* 'forearm' and a few others.

(15) Hidatsa (Boyle 2007: 81)

a. *macée idawashúga*
wacée ita=wašúka
man 3SG.POSS=dog
'man's dog'

b. *idawashúga*
ita=wašúka
3SG.POSS=dog
'his dog'

Boyle (2007) does not mention in his grammar of Hidatsa whether there exists a juxtaposition of possessor-possessed as another possible PC in Hidatsa. One of

the peculiarities of PCs in Hidatsa is that they can freely be modified by a definite article and/or a demonstrative pronoun. Since there are a lot of similarities between Crow and Hidatsa, the discussion of the properties of the possessed will be brief.

4.2.2 The possessed

As in Crow, there are two paradigms of possessive pronouns in Hidatsa, one indicating inalienable possession, the other alienable possession; cf. Table 13.

Table 13: Alienable and inalienable possessive pronouns in Hidatsa (Boyle 2007: 72; 80)

	inalienable possessive pronouns		alienable possessive pronouns	
1	ma- /wa-/	'my'	mada= /wa-ta=/	'my'
2	ni- /ri-/	'your	nida= /ri-ta=/	'your
3	i- /i-/	'his, her'	ida= /i-ta=/	'his, her'

The paradigm for inalienable possession shows — as with Crow set_1 nouns — phonologically conditioned allomorphy (stem-initial vowel vs. stem-initial consonant, and /r/-initial stems). It seems that there is no semantic sub-classification associated with the allomorphy in the inalienable prefixes and the corresponding irregularities. Therefore, I lumped these different formal properties of inalienable nouns together in one set_1 class of nouns in Table 14.

However, there are also differences. For instance, the 2.POSS forms do not trigger a shift in stress assignment as in Crow, and the inalienable possessive prefixes are true prefixes, whereas the corresponding alienable forms are analyzed as clitics. The alienable forms are identical to the ones for inalienable possession plus /*ta*-/ which can be found in other Siouan languages as well (cf. e.g. in Lakota alienable PCs of set_4 nouns which have a *-t^ha* prefix added to the undergoer pronominal prefix; cf. Table 16 below). There is no mention of a depossessivizer in Boyle's grammar of Hidatsa.

4.3 Mandan

Mandan is considered a proper sub-branch of Siouan neither belonging to the Missouri Valley nor the Mississippi Valley group of Siouan.

Table 14: Alienable vs. inalienable distinction in Hidatsa

inalienable set_1	alienable set_2
inalienable paradigm (including phonologically conditioned allomorphy and some irregular forms)	alienable paradigm (no allomorphy)
closed class of nouns: body parts, many kinship terms, some clothing items	open class of nouns not inherently possessed

The semantic/syntactic properties of the possessor and their possible effects on the choice of the PC are not discussed and described in Mixco's grammatical sketch (Mixco 1997). However, looking into the appended Mandan text, it seems that juxtapositions are possible in case the possessor is a lexical noun. There is at least one clear example of this construction (cf. (16)) that shows that association may be expressed by this PC.

(16) Mandan (Mixco 1997: 70: text line 24)
ʾwį=ti rųʾwąʔk=ši-s
village man=good-DEF
'the village chief'

If the possessor is a speech act participant or a third person, one of the following distinct PCs has to be used. In one construction the possessive pronominal affixes, which are in principle identical to the undergoer series of pronominal affixes (called 'stative' in Mixco 1997: 44) are attached directly to the noun stem that designates the possessed $[\text{POSS-N}_{\text{stem}}]_{\text{inalienable possession}}$. This construction is used for inalienable possession; see the relevant forms in Table 15.

The second PC inserts a prefix *ta-* between the stem and the possessive prefix $[\text{POSS-}ta\text{-N}_{\text{stem}}]_{\text{alienable possession}}$. This construction is used for alienable possession. The form *ta-* as an alienable marker is cognate to Lakota *tʰá-*, see below. The possessive prefixes are the same as in the inalienable PC, see Table 15.

There are some peculiarities with PC for inalienable possession. First, there are some kinship terms that require a prefix *ko-* for third person possessor. I suppose

Table 15: Possessor affixes in Mandan (Mixco 1997: 16f,44)

	SG	PL
1	wį-[a]	ro:-
2	rį-	rį-stem-rįt
3	i-	-kræ[b]

[a] Note that this form of the 1SG.POSS differs from the corresponding form of the undergoer series, which is *wą-*. Mixco speculates that the *wį-* form is a contraction of *wą-* + *i-* for the third person, but provides no evidence for this idea.

[b] Mixco does not give the full paradigm, neither for the stative or undergoer affixes nor for the possessive affixes. This is the reason for the question mark. In addition I did not find a single example in Mixco's sketch of Mandan that corresponds to 'their Y'. Note, however, that Kennard (1936: 8) gives the form *-kɛrɛ* for the 3PL possessive affix. The forms are identical, but the transcription is different.

this form is related historically to *ku-/tku-* in Lakota. Secondly, there are kinship terms and a few other alienable terms (old nominalized verb forms) that take the actor series of pronominal prefixes in order to express the possessor. For instance, the kinship term for 'mother' takes the usual undergoer series of prefixes for inalienable possession, but requires *ko-* for the third person possessor; cf. (17).

(17) Mandan (Mixco 1997: 45)

a. *wį-hų:-s*
1SG.POSS-mother-DEF
'my mother'

b. *rų-hų:-s*
2SG.POSS-mother-DEF
'your mother'

c. *ko-hų:-s*
3SG.POSS-mother-DEF
'his mother'

The term for 'father', on the other hand, requires the actor series of pronominal affixes in Mandan in order to express the possessor, cf. the examples in (18).

(18) Mandan (Mixco 1997: 45)

a. *wa-aʔt-s*
1SG.A-father-DEF
'my father'

b. *a-aʔt-s*
2SG.A-father-DEF
'your father'

c. *ko-aʔt-s*
3SG.A-father-DEF
'his father'

Interestingly, no mention is made of the way body parts are possessed in Mandan. A quick look into the Mandan text (cf. Mixco 1997: 66ff) reveals that body-part nouns never occur in one of the above described PCs with possessive affixes. They appear always without the *ta-* form and never carry any possessive affixes. The possessor always has to be inferred from the text.

4.4 Lakota

4.4.1 The possessor

Lakota is a language of the Mississippi Valley Siouan group, more specifically of the Dakotan sub-branch of this group. Lakota does employ possessive pronouns, which are almost entirely identical to the set of undergoer pronominal prefixes in stative/inactive verbs. If the possessor is a SAP/pronoun and the possessed noun belongs to the class of alienable nouns, the following constructions may be used. Note that the 1SG.POSS *mi-* is a special form that does not correspond to the regular 1SG form of the pronominal undergoer prefixes (*ma-*).[6]

a) Ownership [$N_{possessed\text{-}inanim}$ PRO.POSS-HAVE DET]

(19) Lakota (Buechel 1939: 98)
thípi mi-tȟáwa kiŋ
house 1SG-have DEF
'my house'

[6] Data in this section has been re-spelled in the current Lakota orthography.

b) Ownership, attribution of property [PRO-*tȟa*-$N_{possessed\text{-}inanim/abstr}$ DEF]

(20) Lakota (Buechel 1939: 98)

a. *mi-tȟá-makȟočhe kiŋ*
1SG.POSS-POSS-land DEF
'my land'

b. *nitȟóksape kiŋ*
ni-tȟá-wóksape kiŋ
2SG-POSS-wisdom DEF
'thy wisdom'

There is no information about the conditions or the differences between the two constructions; it is clear that the one in (19) contains a stative verb of possession *tȟáwa-* 'have' that is nominalized in this context inflecting for the person and number of the possessor and the number of the possessed. In Rood & Taylor (1996: 458) it is said that the stative verb of possession *itȟáwa* 'have' depends only on the category of the possessor in this PC and not on the number of the possessed. It seems that this stative verb of possession has been grammaticalized towards a marker of possession quite recently in Lakota.

The PCs in (20) contain a marker for possession *tȟá-* 'POSS' which is attached to the possessed noun and preceded by the pronominal affix of the possessor. This marker is common Siouan (cf. Rankin et al. 2015). If there are lexical nouns expressing the possessor, the following PCs are used.

c) Ownership [$N_{possessed\text{-}anim}$ $N_{possessor\text{-}PROP}$ PRO.POSS-HAVE DEF]

(21) Lakota (Buechel 1939: 91)
šúŋka wakȟáŋ David Ø-tȟáwa kiŋ
horse D. 3SG-have DEF
'David's horse'

[$N_{possessed\text{-}anim}$ $N_{possessor\text{-}PROP}$ PRO.POSS-HAVE DEF]

(22) Lakota (Buechel 1939: 91)
šúŋka wakȟáŋ Peter na Paul Ø-tȟáwa-pi kiŋ
horse P. and P. 3SG-have-PL DEF
'Peter and Paul's horses (or horse)'

d) Association [$N_{possessor\text{-}PROP}$ PRO.POSS-*tȟa*-$N_{possessed\text{-}hum}$ DEF]

(23) Lakota (Buechel 1939: 92)
Itȟáŋčhaŋ Ø-tȟa-wóilake kiŋ
Lord 3SG-POSS-servant DEF
'the Lord's servant'

[$N_{possessor\text{-}PROP}$ PRO.POSS-*tȟa*-$N_{possessed\text{-}hum}$ DEF]

(24) Lakota (Buechel 1939: 92)
Abraham Ø-tȟa-wámakȟaškaŋ-pi kiŋ
A. 3SG-POSS-animal-PL DEF
'Abraham's animals'

Again we have two different PCs in the examples (21)-(24) with a lexical possessor, one with a verb of possession that is nominalized, and the other exhibiting a morphological possessor marking on the possessed noun. These examples represent alienable possessions. It can be concluded that the syntactic status of the possessor does not play a role for the choice of the PCs.

If the relation between the possessor and the possessed is a whole-part relation, or a partitive relation, or the possessor noun is an abstract noun or a nominalization, the following constructions are used.

e) Whole-part relationships [$N_{possessor\text{-}inanim}$ $N_{possessed\text{-}anim}$ DEF] (juxtaposition)

(25) Lakota (Buechel 1939: 92)
maȟpíya zitkála-pi kiŋ
cloud bird-PL DEF
'the birds of the air'

[$N_{possessor\text{-}inanim}$ $N_{possessed\text{-}inanim}$ INDEF]

(26) Lakota (Buechel 1939: 92)
čheȟ íkȟaŋ waŋ
bucket rope INDEF
'a bucket handle, rope of a bucket'

f) Partitive

(27) Lakota (Buechel 1939: 93)
itȟáŋčhaŋpi kį óta
chiefs DEF many
'many of the chiefs'

Example (27) is not really a PC, but a regular quantified NP. The same holds for (28). It can hardly be considered a PC. It is rather a juxtaposition expressing a NP ('good works') modifying another NP ('man').

g) With an abstract possessor N

(28) Lakota (Buechel 1939: 93)
wičháša oȟ'aŋ wašté kiŋ héčha
man in.actions good DEF such
'a man of good works'

4.4.2 The possessed

There are different PCs according to the semantic type of the possessed noun; body-part terms are simply affixed by the pronominal series of undergoer prefixes. Among the body-part terms, there is a split between body parts that are "conceived as particularly subject to willpower" (Boas & Deloria 1941: 128), and the others. Buechel (1939: 100) describes this difference as "possession of one's incorporeal constituents" versus "possession of one's body and its physical parts"; compare the examples in (29) and (30).

(29) Lakota (Buechel 1939: 101)
mi-náğ̌i kiŋ 'my souls'
mi-čháže kiŋ 'my name'
mi-óȟ'aŋ kiŋ 'my occupation'
etc.

(30) Lakota (Buechel 1939: 100)
ma-čhéži kiŋ 'my tongue'
ma-íšta kiŋ 'my eye'
ma-sí kiŋ 'my foot'
etc.

Note that this distinction has become partially obsolete in contemporary Lakota. Rood & Taylor (1996: 458) note that this distinction is semantically maintained only in the Oglala variety of Lakota. There *ma-* (1SG.POSS) is used for "concrete visible possessions", and *mi-* (1SG.POSS) for "intangibles" (cf. Rood & Taylor 1996: 458). Otherwise, both forms are in free variation.

Kinship relations with a possessor of the first and second person are expressed solely by the possessive prefixes. A possessor of the third person requires an additional marker *-ku, -tku, -ču* which is suffixed to the possessed kinship term; cf. (31).

(31) Lakota (Buechel 1939: 102)

mi-tȟúŋkašila	'my grandfather'
ni-tȟúŋkašila kiŋ	'thy grandfather'
*Ø-tȟúŋkaši**tku** kiŋ*	'his/her grandfather'

Table 16 summarizes the findings. As was mentioned above, the set_1 and set_2 possessed nouns are no longer separated formally in Lakota (except for Oglala).

Table 16: Alienable vs. inalienable distinction in Lakota (Boas & Deloria 1941: 127–133)

inseparable/inalienable		separable/alienable	
set_1	set_2	set_3	set_4
body-part terms [+control] [incorporeal constituents] mouth, lips, facial expression, eye, arm, voice, hand, spirit, etc.	body-part terms [-control] [physical parts] kidney, knee, liver, lungs, blood, etc.	kinship relations ownership	distal affinal kinship terms prototypical
PC	PC	PC	PC
[PRO.POSS-noun] with a special form in the 1SG.POSS (*mi-*)	[PRO.POSS-noun]	[1./2.POSS-noun] [3.POSS-noun-*ku*] *-tku*] *-cu*]	[PRO.POSS -t^ha-noun] [noun PRO.POSS-$t^ha'wa$]

As in Hoocąk, the causative verb is used for the clause-level predicative expression of possession of a kinship term, cf. (32).

(32) Lakota (Buechel 1939: 102)

a. *tȟuŋkášila-wa-ya*
grandfather-1SG.A-have.kin
'I have (him) as grandfather.'

b. *tȟuŋkášila-uŋ-yaŋ-pi*
grandfather-1I.A-have.kin-PL
'We have (him) as grandfather.'

I found no example showing that this verb of possession could be used like the alienable verb of possession *tȟáwa* illustrated in (19) above. If this were the case, we would have a quite similar opposition of verbs of possession in Lakota as we found in Hoocąk.

In addition, it should be mentioned that Lakota allows the non-modifying auto-referential usage of the possessive pronouns, however only the expressions based on the verb of possession *tȟáwa* plus the definite article. This could be interpreted as a nominalized possessive predication; cf. (33).[7]

(33) Lakota (Buechel 1939: 22)
mitȟáwa kiŋ hé ahí ičú
mine DEF she came take
'She came and took mine'

Interestingly, this is a PC in which there is no possessed noun. All other PCs discussed so far require a possessed lexical noun.

4.5 Osage

Osage is taken as a representative of the Dhegiha sub-branch of Mississippi Valley Siouan. It was chosen because there is a recent extensive grammatical description of this language (Quintero 2004). Unfortunately, it is difficult to find the relevant data in Quintero's grammar of Osage. There is no specific chapter on possession, and there is no index in the grammar. Quintero uses the terms alienable and inalienable, but it is not made explicit which nouns are alienable and which are inalienable. However, some conclusions about this question can be drawn from the numerous examples provided by the grammar. There is a

[7] One of the reviewers mentioned that *mitȟáwa kiŋ hé* could be analyzed as a null head relative clause. This is probably the best way to treat it. It does not, however, change the argument here. The example only demonstrates that a nominal expression for the possessed is not required in this possessive construction.

special construction for PCs with possessed kinship nouns. Kinship nouns are inflected with a series of inalienable pronominal prefixes, cf. Table 17.

Table 17: Inalienable possessive prefixes for kinship terms in Osage (Quintero 2004: 481f)

Possessor	inalienable prefix paradigm	example	translation
1SG	wi-	wi-sǫ́ka	'my (male's) younger brother'
2SG	ðe-	ðe-sǫ́ka	'your (male's) younger brother'
3SG	i-	i-sǫ́ka	'his (male's) younger brother'
1PL	does not exist	-	-
2PL	?	?	
3PL	?	?	

The question marks in Table 17 indicate that Quintero did not provide the expected forms. In addition, PCs with possessed body-part nouns are not provided either.

Alienable nouns require another construction, which has the following properties. There is a pronominally inflected (bound) stem *-hta*, which marks possession.[8] The pronominal prefixes resemble the ones used for the PCs with possessed kinship terms, with one exception. There is a dual and plural form for the first person, which does not exist in the PCs with possessed kinship terms. The inflected possessive form follows the possessed noun; cf. the examples in (34) and (35). The full paradigm is given in Table 18.

(34) Osage (Quintero 2004: 298)
ówe che ***hcí*** ***ąkóhta-api*** *aðį́-ahi-a*
groceries those **house** **1PL.POSS-PL** have-arrive.there-IMP
'Bring those groceries to our house!'

(35) Osage (Quintero 2004: 299)
Máry J̌óhn-a ***hcí*** ***íhta-api***
M. J.-SYL **house** **3SG.POSS-PL**
'Mary and John's house'

[8] Again, this is the Common Siouan marker for alienable possession (cf. Rankin et al. 2015).

Table 18: Alienable possession in Osage (Quintero 2004: 297f)

	possessed	possessor	translation
1SG	hcí 'house'	wihta ? (<wi-hta)	'my house'
2SG	hcí 'house'	ðíhta (<ðí-hta)	'your house'
3SG	hcí 'house'	ihta (<i-hta)	'his/her house'
1DU	hcí 'house'	ąkóhta (< ąkó-hta)	'our house'
1PL	hcí 'house'	ąkóhtapi (<ąkó-hta-api)	'our house'
2PL	hcí 'house'	ðíhtaapi (<ðí-hta-api)	'your house'
3PL	hcí 'house'	ihta-api (<i-hta-api)	'their house'

Quintero analyzes the possessive form *-hta* as a noun or nominal element for two reasons: first, this stem is inflected by the same prefixes as the inalienable nouns (kinship terms), and secondly, if it were analyzed as a verbal stem, the possessive inflection would be quite irregular (cf. Quintero 2004: 317f).

One problem with this reasoning is that one would have to expect that the nominal stem *-hta* belongs to the group of inalienable nouns because it requires the inalienable series of prefixes. There is, however, no evidence for that. Secondly, the order of elements suggests that the *-hta* stem is of verbal origin. If it were nominal, it should precede the possessed noun. Attributive nouns always precede the head nouns; all other modifying elements follow the head noun. That the pronominal prefixes are different from the ones for stative/inactive verbs is not necessarily an argument for the non-verbal character of the stem – there are often deviations in possessive paradigms. Furthermore, this possessive form may be used autonomously without a possessed noun, cf. the example in (36). This construction is not possible in Hoocąk. The utterance in (36) would require the reflexive possessive prefix *k-/kara-* in Hoocąk.

(36) Osage (Quintero 2004: 413)
ąkóhta akxa Ø-xǫ́-api-ðe
1PL.POSS SBJ 3SG.SBJ-break-PL-DECL
'Ours is broken'

Part-whole relationships - at least with regard to inanimate parts - seem to be expressed by means of a simple juxtaposition. However, I found only one example illustrating this in Quintero's grammar, cf. example (37).

(37) Osage (Quintero 2004: 423)
oðíhtą hci hcíže áðiitą-a
car house door close-IMP
'Close the garage door!'

To summarize: there is an alienable/inalienable distinction in Osage and it seems that kinship terms belong to the inalienable set of nouns (set_1), while all other nouns belong to the alienable set of noun (set_2); cf. Table 19.

Table 19: Alienable vs. inalienable distinction in Osage

inalienable	alienable
set_1	set_2
kinship terms	all other nouns ?
PC	PC
PRO-$N_{possessed}$	($N_{possessor}$) $N_{possessed}$ PRO.POSS-*hta*

4.6 Biloxi

Biloxi was chosen as a representative of the Ohio Valley sub-branch of Siouan. The standard reference work with respect to a grammatical description is Einaudi (1976). She mentions two NP internal PC types in her grammar of Biloxi, a) a juxtaposition of two nominals to be used for all kinds of possessed nouns, and b) pronominally inflected nouns designating body parts and kinship relations (cf. Einaudi 1976: 57–68). Concerning a) the order of nouns in the juxtaposition PC is possessor precedes possessed. Concerning b) if body parts and kinship terms are possessed, the possessed nouns have to be inflected obligatorily with pronominal prefixes that are identical to the ones in verbs. This holds also for some intimate personal possessions such as 'house', 'clothing', etc. See two examples for the juxtaposed PC construction in (38) and two examples of the inflected PC construction in (39).

(38) Biloxi (Einaudi 1976: 139f)

a. *qya ti-k*
man house-DET
'the man's house'

b. *ama tupe kạ*
ground hole DET
'the ground's hole'

(39) Biloxi (Einaudi 1976: 139f)

a. *tuhe Ø-tukạni yandi*
T. 3SG-uncle DET
'Tuhe's uncle (mother's brother)'

b. *ạya Ø-anahị kạ*
man 3SG-hair DET
'people's hair'

Full paradigms of inalienable possession are given in Table 20.

Table 20: Paradigm of inalienable possession in Biloxi (Einaudi 1976: 57f/62f)

possessor	kinship term adi 'father'	body-part term cake 'hand'
1SG	nk-adi	nk-cake
2SG	iy-adi	i-cake
3SG	Ø-adi	Ø-cake
1PL	nk-ax-tu	nk-cak-tu
2PL	iy-adi-tu	i-cak-tu
3PL	ax-tu	cak-tu

I did not find any examples that illustrate how alienable nouns are possessed by SAP possessors, something like 'my horse', 'your car', etc.

5 Conclusions

There is an alienable-inalienable distinction in one way or other in all Siouan languages, even in Biloxi, as seen in Table 21, but there, the inalienable nouns (kinship, body parts) are inflected by means of the subject prefixes. As the examination of PCs in the various Siouan languages shows, there are at least four kinds of constructions that are used to express possession on the NP level. The simplest construction is juxtaposition, which is used in all sample languages except for Hidatsa, for which no data were available. Inalienable possession is expressed in

Table 21: Alienable vs. inalienable distinction in Biloxi

inalienable	alienable
set_1	set_2
kinship terms body-part terms intimate personal possessions such as 'house', clothing'	all other nouns
PC	PC
PRO-$N_{possessed}$ DET	$N_{possessor}$-$N_{possessed}$ DET

all sample languages with a series of possessive affixes directly attached to the possessed. The sole exception is Hoocąk, which has no possessive affixes. There are two principal constructions that express alienable possession in the sample Siouan languages. There is a construction that has a possessive marker attached to the stem indicating alienable possession. The same set of possessive affixes appears with these constructions. This construction is not available in Hoocąk and Biloxi. The second construction utilizes a verb of possession that is nominalized by a determiner and inflected by the same paradigm of possessive affixes. It follows the possessed noun. This construction is missing in Missouri Valley Siouan and in Biloxi. I have no clear data for Osage. The principle types of constructions that are used in Siouan languages to express possession are summarized in Table 22 together with the semantic kinds of possessed nouns.

The nominalized verbs of possession appear only in Mississippi Valley Siouan, most prominently in Hoocąk. Hoocąk is particular also with regard to the lack of the two middle construction types in Table 20; one could perhaps say that Hoocąk has not really grammaticalized a NP-internal possessive construction: juxtaposition is semantically the most abstract means, hence able to subsume all kinds of binary relations (among them also real ownership) and the verbal expression of possession is semantically the most concrete one, hence excluding many binary relations that are often expressed by means of possessive constructions (there is no possibility to express association, whole-part, attribution of property relations with these PCs).

Another interesting observation is that there is no neat classification of nouns with respect to the alienable/inalienable distinction. Alienable and inalienable

Table 22: Distribution of NP-internal possessive constructions among Siouan languages

	juxtaposition $N_{poss'or}$ $N_{poss'ed}$	less marked ⟶ POSS.PRO-$N_{poss'ed}$	POSS.PRO-POSS-$N_{poss'ed}$	more marked $N_{poss'ed}$ POSS.PRO-verb.poss-DET
Crow	1) part-whole 2) others?	1) body parts, kinship terms 2) internal body parts 3) 'chest', 'tail', 'husband' 4) closely associated with possessor, e.g. clothing items, kin terms, cultural possession	rest, plus some exceptions	Ø
Hidatsa	?	1) many kinship terms 2) body parts 3) some clothing items	rest	Ø
Mandan	1) association 2) body parts	1) kinship terms 2) ?	1) kinship terms 2) ?	?
Lakota	1) ownership 2) part-whole	1) body parts 2) internal body parts 3) kinship terms	1) kinship terms 2) ownership 3) attribution of property 4) association	1) ownership 2) kinship
Hoocąk	1) part-whole 2) body parts 3) kinship 4) local nouns	Ø	Ø	1) kinship 2) domestic/ pet animals 3) rest
Osage	part-whole	kinship terms	ownership?	
Biloxi	1) part-whole 2) ownership 3) rest	1) kinship 2) body parts 3) intimate personal belongings ('house', 'clothing')	Ø	Ø

nouns are distributed over all kinds of PCs and it seems that the often observed markedness relations between alienable and inalienable PCs do not really hold in Siouan. For instance, juxtapositions as the least marked PCs comprise real ownership (Lakota, Biloxi) as well as body parts (Mandan, Hoocąk) and kinship terms (Hoocąk). On the other hand, nominalized predicative PCs, which are the most complex PCs in this study, include not only real ownership (Lakota, Hoocąk) but also kinship terms which are inalienable nouns. The two construction types in the middle columns in Table 20 show a markedness relation between inalienable and alienable nouns that is much clearer. The PC with the possessive pronouns attached to the possessed nouns (second column from left) are chosen primarily for inalienable possession (all languages except Hoocąk) and the PC with the added possession marker (POSS) are used overwhelmingly for alienable possession such as real ownership or as a kind of rest category that always includes alienable nouns (all languages except Hoocąk). In Lakota and Mandan, however, kinship terms as possessed nouns are included, which blurs this distinction to some degree.

Abbreviations

1, 2, 3,	first, second, third person	INDEF	indefinite article
A	actor	OBJ	object
AH	Animacy Hierarchy	PC	possessive construction
APPL.BEN	benefactive applicative	PL	plural
APPL.SUPESS	locative applicative superessive	POSS PRO	possessive pronoun
		PREP	preposition
		PROP	proper name
DAT	dative	REFL.POSS	reflexive possession
DECL	declarative		
DEF	definite article	SAP	speech act participant
E	exclusive		
EMPH	emphatic	SBJ	subject
GEN	genitive	SG	singular
I	inclusive	U = undergoer.	

References

Boas, Franz & Ella Deloria. 1941. *Dakota grammar* (Memoirs of the National Academy of Sciences 23). Washington: Government Printing Office.

Boyle, John P. 2007. *Hidatsa morphosyntax and clause structure*. Chicago: University of Chicago (Doctoral dissertation).

Buechel, Eugene, S. J. 1939. *A grammar of Lakota: The language of the Teton Sioux Indians*. Saint Francis, SD: Saint Francis Mission.

Comrie, Bernard. 1981. *Language universals and linguistic typology*. Chicago: University of Chicago Press.

Croft, William. 2003. *Typology and universals*. 2nd edn. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Dixon, Robert Malcolm Ward. 1979. Ergativity. *Language* 55. 59–138.

Dixon, Robert Malcolm Ward. 2010. *Basic linguistic theory*. Vol. 2 (Grammatical topics). Oxford: Oxford University Press.

Einaudi, Paula Ferris. 1976. *A grammar of Biloxi*. New York: Garland Publishing.

England, Nora. 1983. *A grammar of Mam, a Mayan language* (Texas Linguistic Series). Austin: University of Texas Press.

Graczyk, Randolph Jan. 2007. *A grammar of Crow*. Lincoln: University of Nebraska Press.

Heine, Bernd. 1997. *Possession: Cognitive sources, forces and grammaticalization*. Cambridge: Cambridge University Press.

Helmbrecht, Johannes. 2003. *Possession in Hocąk (Winnebago): Problems for a prototype approach* (Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt (ASSidUE) 8). Erfurt: Universität Erfurt. http://www.db-thueringen.de/servlets/DerivateServlet/Derivate-1887/ASSidUE08.pdf.

Helmbrecht, Johannes & Christian Lehmann (eds.). 2010. *Elements of grammar, learner's dictionary* (Hocąk Teaching Materials 1). Albany: State University of New York – Albany Press.

Kennard, Edward A. 1936. Mandan grammar. *International Journal of American Linguistics* 9. 1–43.

Kornfilt, Jaklin. 1990. Turkish and the Turkic languages. In Bernard Comrie (ed.), *The world's major languages*, 619–645. Oxford: Oxford University Press.

Langacker, Ronald W. 1993. Reference-point constructions. *Cognitive Linguistics* 4. 1–38.

Lipkind, William. 1945. *Winnebago grammar*. New York: King's Crown Press.

Mithun, Marianne. 1999. *The languages of Native North America* (Cambridge Language Surveys). Cambridge: Cambridge University Press.

Mixco, Mauricio J. 1997. *Mandan* (Languages of the World / Materials 159). München: Lincom.

Parks, Douglas R. & Robert L. Rankin. 2001. Siouan languages. In Raymond J. DeMallie & William C. Sturtevant (eds.), *Plains*, vol. 13 (Handbook of North American Indians), 94–114. Washington: Smithsonian Institution.

Payne, Doris L. & Immanuel Barshi (eds.). 1999. *External possession* (Typological Studies in Language 39). Amsterdam: John Benjamins.

Quintero, Carolyn. 2004. *Osage grammar* (Studies in the Anthropology of North American Indians). Lincoln: University of Nebraska Press.

Rankin, Robert L., Richard T. Carter, A. Wesley Jones, John E. Koontz, David S. Rood & Iren Hartmann (eds.). 2015. *Comparative Siouan dictionary*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http : / / csd . clld . org (22 May, 2015).

Rood, David S. 1979. Siouan. In Lyle Campbell & Marianne Mithun (eds.), *The languages of Native America: Historical and comparative assessment*, 236–298. Austin: University of Texas Press.

Rood, David S. & Allan R. Taylor. 1996. Sketch of Lakhota, a Siouan language. In William C. Sturtevant & Ives Goddard (eds.), *Languages* (Handbook of North American Indians 17), 440–482. Washington: Smithsonian Institution.

Seiler, Hansjakob. 1983. *Possession as an operational dimension of language*. Tübingen: Gunter Narr Verlag.

Seiler, Hansjakob. 2001. The operational basis of possession: A dimensional approach revisited. In Irène Baron, Michael Herslund & Finn Sørensen (eds.), *Dimensions of possession*, 27–40. Amsterdam: John Benjamins.

Name index

Language index

Subject index

Advances in the study of Siouan languages and linguistics

The Siouan family comprises some twenty languages, historically spoken across a broad swath of the central North American plains and woodlands, as well as in parts of the southeastern United States. In spite of its geographical extent and diversity, and the size and importance of several Siouan-speaking tribes, this family has received relatively little attention in the linguistic literature and many of the individual Siouan languages are severely understudied. This volume aims to make work on Siouan languages more broadly available and to encourage deeper investigation of the myriad typological, theoretical, descriptive, and pedagogical issues they raise.

The 17 chapters in this volume present a broad range of current Siouan research, focusing on various Siouan languages, from a variety of linguistic perspectives: historical-genetic, philological, applied, descriptive, formal/generative, and comparative/typological. The editors' preface summarizes characteristic features of the Siouan family, including head-final and "verb-centered" syntax, a complex system of verbal affixes including applicatives and subject-possessives, head-internal relative clauses, gendered speech markers, stop-systems including ejectives, and a preference for certain prosodic and phonotactic patterns.

The volume is dedicated to the memory of Professor Robert L. Rankin, a towering figure in Siouan linguistics throughout his long career, who passed away in February of 2014.

ISBN 978-3-946234-37-1