Proceedings of the 41st Siouan and Caddoan Languages Conference

Held virtually via Zoom

May 20-22, 2021

Editors

Ryan M. Kasak
Catherine Rudin
Armik Mirzayan
Samiron Dutta

Kasak, Ryan M., Catherine Rudin, Armik Mirzayan, & Samiron Dutta (eds.). 2022. *Proceedings of the 41st Siouan and Caddoan Languages Conference.* Norman, OK: Siouan and Caddoan Languages Conference.

This title can be downloaded at:
http://www.siouan.org/proceedings/sclc41/

ISSN: 2641-9904 (digital)

Typesetting: Catherine Rudin, Julie Marsault, & Ryan Kasak
Reviewing & Proofreading: John Boyle, Samiron Dutta, Dylan Herrick, Armik Mirzayan, Ryan Kasak, & Catherine Rudin,
Font: Libertinus Serif
Typesetting software: XeLaTeX & Biber

First printing, 12/31/2022

Contents

Acknowledgments

The editors wish to acknowledge Edwin Ko's contribution to the Proceedings of 41st Siouan and Caddoan Languages Conference by organizing this year's conference for the second year in a row. This conference was held online via Zoom once again. The conference itself could not have been taken place without Edwin's efforts. We also wish to acknowledge the University of California, Berkeley for enabling us to use Zoom through an account afforded to Edwin. Without the involvement of either of the aforementioned parties, the conference and therefore these proceedings would not have been possible.

The editors also wish to acknowledge the passing of several major figures in the field of Siouan and Caddoan language research. Dr. Raymond J. DeMallie passed away on April 25, 2021,[1] and Dr. Douglas Parks passed away on May 20, 2021.[2] Both of these men spearheaded the establishment of the famous American Indian Studies Research Institute (AISRI) at Indiana University in 1984, an institution that has trained numerous scholars to work with speakers of languages of the Plains. Their invaluable work led to the creation of dictionaries, pedagogical materials, and text and audio corpora that are assisting members of different Indigenous communities to this day. Much of their work with Arikara, Assiniboine, Lakota, and Pawnee speakers and their communities continues to be of critical importance to the evolving understanding of these languages.

[1]His obituary, written by Anya Peterson Royce with Jason Baird Jackson, can be found here: https://www.dignitymemorial.com/obituaries/bloomington-in/raymond-demallie-10168067.

[2]His obituary, written by Indrek Park, can be found here: https://www.dignitymemorial.com/obituaries/bloomington-in/douglas-parks-10202283.

Preface

We are very pleased to put out yet another volume of the Proceedings of the Siouan and Caddoan Languages Conference. It has been a real privilege to put out novel research on Siouan and Caddoan languages since the resumption of proceedings volumes in 2018. Since moving online due to the ongoing COVID-19 pandemic, the SCLC has seen a noted increase in attendance. At SCLC 42 at the University of Virginia in 2022, the conference resumed in-person presentations, but with the addition of remote presentations and attendees. This hybrid style has augmented the reach of the conference and made the research being done on Siouan and Caddoan languages more accessible.

SCLC 41 featured two special panels: one on relative clauses in Siouan and another on topic and focus marking in Siouan. The Siouan panel featured presentations on Hidatsa, Hoocąk, Lakhota, and Omaha-Ponca. These presentations looked at both interally headed and externally headed clause formation across these different languages. The topic and focus panel investigated information structure in Mandan and Hoocąk. These presentations explored both the prosodic and morphological indicators for topichood in their respective languages.

This conference was a great success, with seventeen separate presentations across three days and two different discussion sessions after each special panel. Presenters came from a variety of backgrounds, from Indigenous scholars representing their home communities to independent researchers. Students at all levels likewise delivered presentations at this conference: undergraduates, master's students, and doctoral students.

Submissions to the proceedings increased over the previous two years. This change coincides with other major contributions to scholarship in Siouan and Caddoan languages, including by not restricted to Julie Marsault's Ph.D. disseration, *Valency-changing operations in Umó^{n}ho^{n}: Affixation, incorporation, and syntactic constructions* and Sean Panick's M.A. thesis, *r-Nasalization in Hoocąk: A diachronic and synchronic perspective.*

We look forward forward to future conferences and to the advancement of knowledge about Siouan and Caddoan languages. We again thank Edwin Ko for adroitly handling back-to-back hosting duties over the course of this pandemic, and we furthermore thank all those who were proofreaders and reviewers for this volume. *Ahó!*

Ryan M. Kasak
Norman, Oklahoma

Formal and colloquial speech in Stoney Nakoda: Initial observations*

Lloyd Buddy Wesley
Cultural Advisor & Tribal Historian
Stoney Nakoda Nation (Chiniki Band)

Corey Telfer
The Language Conservancy
University of Calgary

Abstract: This paper gives an overview of a few notable distinctions made between the colloquial and formal registers of Stoney Nakoda, also known as southern Stoney, a Dakotan language spoken in southern Alberta. Like many Indigenous languages spoken in small communities, Stoney Nakoda does not frequently make use of special politeness words in order to show respect, courtesy or formality, and instead alternative constructions or pronunciations can be used to convey these attitudes.

Keywords: Stoney, Nakoda, sociolinguistics, politeness

1. Introduction

1.1. Background

This paper focuses on the southern dialect of Stoney, also known as Stoney Nakoda, spoken primarily in the three communities that make up the Stoney Nakoda First Nation: Wapta Mnotha (Big Horn), Mînî Thnî (Morley), and Gahna (Eden Valley). These communities are located along the Rocky Mountains in southern Alberta. They refer to themselves as *Îyârhe Nakoda* 'People of the Mountains' or *Îethkabi*, and the language is typically referred to as *Îethka Îabi*.[1] There are also two northern dialects of Stoney spoken in central Alberta: *Isga* is spoken at Wakamne (Alexis Nakota Sioux Nation) and *Iithga* is spoken at Wihnemne (Paul First Nation), northwest and west of Edmonton respectively. The first author of this paper is from Mînî Thnî (Morley), which is the largest Stoney community.

Stoney is a Siouan language belonging to the Mississippi Valley, or Central Siouan, branch of the Siouan family. It belongs to the Dakotan group of languages within that branch, which formed a dialect continuum across the northern plains of North America, ranging from the north central United States to northwestern Alberta. Stoney shares some clear similarities with Assiniboine, which goes by several other names, including Nakona and Nakhóta, and is spoken in neighbouring Montana and Saskatchewan. Despite these similarities, there are still enough differences that Stoney and Assiniboine are classified as separate languages, however further research is needed to clarify their relationship. Stoney is the northwestern-most language of the

*We would like to acknowledge the Stoney Education Authority and The Language Conservancy for making this work possible. The second author would also like to honour the memory of the first author, the late Lloyd Buddy Wesley, traditionally named Siyaga Ayitho (Swims for a Grebe), who passed away not long after this paper was presented.

[1]The endonym *Îethka* is sometimes spelled with a *y*: *Îyethkabi*, *Îyethka Îabi*.

Siouan language family, which includes languages spoken as far south as the Gulf of Mexico (Biloxi) and coastal South Carolina (Woccon).

1.2. Politeness in Indigenous languages

Students interested in learning Indigenous languages often ask how to translate common indicators of politeness found in English and other European languages, such as *please*, *thank-you*, *you're welcome*, *ma'am*, *sir*, etc. In the course of teaching the Stoney Nakoda language, we have found that even though there are not many specific politeness terms, there are other ways of showing additional respect or being more formal.

In this paper we will focus primarily on morphosyntactic variation, but we will also consider the kinship lexicon, which includes shortened familiar versions for nearly every kinship term. There is of course extensive lexical and morphophonological variation beyond this, however these patterns are often limited to certain age categories or family dialects and therefore fall outside the scope of this paper. The morphosyntactic variation we document here involves the following types of grammatical structures: declaratives, imperatives, interrogatives, conjunctions, and relative clauses.

2. Kinship terms

Like other Dakotan cultures, the Stoney Nakoda kinship system is classified as a variation of the system used by the Haudenosaunee[2] people, where the ego's parents' same-gender siblings are referred to with terms similar to those used to address the parents themselves. Similarly, cousins are classified according to whether they are descended from a parental sibling of the same gender as the parent (parallel-cousins) or the opposite gender of the parent (cross-cousins). Parallel-cousins use the same kinship terms as siblings while cross-cousins have their own separate terms, often translated simply as *cousin* by Stoney Nakoda speakers.

As with other Indigenous North American cultures, awareness of kin relationships is very important in Stoney Nakoda culture, to the point where it can be considered rude to fail to address a relative using the correct kinship term. For this reason, familiarity with kinship terminology is very important to speakers and will be an important element of maintaining and revitalizing the Stoney Nakoda language. Virtually every kinship term has a standard version, used when speaking about a relative to someone else or when speaking to someone formally, and a shortened familiar version, that can typically only be used to address that relative directly. In addition, a few kinship terms have special alternative versions that are only used by younger children and are never used between adults. All of these diminutive kinship terms may sometimes be seen as 'slang' and as less noteworthy than their more formal versions, however we feel that

[2]This is still known as the *Iroquois System* in kinship scholarship, however *Haudenosaunee* is the autonym and generally preferred name for these people. This system is said to follow a bifurcate merging pattern, since the terms for aunts and uncles are bifurcated (split) based on the gender of the parent, but are also merged with the terms for the parents themselves (e.g. the mother and mother's sisters are addressed in a similar way, and likewise the father and father's brothers share the same or similar kinship terms). This means that the parents' same-gender siblings are addressed more like additional parents than as aunts and uncles, and their children are likewise addressed as siblings.

it is important to document them since they are very common in daily language use and are sometimes unpredictable from their longer forms.

The first-person possessive forms of the kinship terms are used both to address relatives and to refer to them in conversation, so they are generally the most common forms heard in the language and are therefore our focus here. In addition, terms for siblings and cross-cousins are gendered with regard to both of the related individuals. We use a following F and M to indicate the gender of the possessor, which is also the gender of the speaker in all of the examples given below.[3]

Table 1: Some Stoney Nakoda kinship terms and their diminutive variants

	Standard form	**Familiar form**	**Children's form**
my grandmother	*îkusín/îkúsin*	—	*chízi*
my grandfather	*mîtûgásin*	—	*tadá*
my mother	*înã́*	—	*nãnâ*
my father	*adé*	—	*téde*
my daughter	*mîchṹksi*	*mîchṹs*	—
my son	*mîchîksi*	*mîchîs*	—
my older sister (F)	*mîchṹn*	—	*chṹchû*
my older sister (M)	*mîtâgén/mîtãgen*	*tãge*	—
my older brother (F)	*mîtîmnón*	*mîtîmno*	—
my older brother (M)	*mîchîn*	*chî*	*chîchî*
my younger sister (F)	*mîtâgán*	*mîtã́*	—
my younger sister (M)	*mîtãksin/mîtâksín*	*tãksi*	—
my younger brother	*mîthûgan*	*mîthṹn/mîthû*	—
my female cross-cousin (F)	*mîsepásin/mîsèpasín*	*pásin*[4]	—
my female cross-cousin (M)	*mãhâgásin*	*(mâhãga)*[5]	—
my male cross-cousin (F)	*mîsichésin*	*(mîsíche)*[6]	—
my male cross-cousin (M)	*mîtãsin*	*mîtãs*	—

[3]This paper makes use of the Stoney Nakoda orthography, which has the following notable features: *â*, *î*, and *û* are nasal vowels, *ch* and *j* are aspirated and voiced affricates respectively, *r* is a voiced pharyngeal approximant, *rh* is a voiceless pharyngeal fricative, *s* and *z* are voiceless and voiced alveolopalatal fricatives, *th* and *th* are voiceless and voiced dental fricatives, and *y* is a palatal approximant. The stops *p*, *t*, *k* are always aspirated prevocalically, but are typically unaspirated when following another consonant. The stops *b*, *d*, *g* are typically voiced but may rarely be voiceless unaspirated (this difference is not used to distinguish any minimal pairs). Adjacent vowels are typically interrupted by a glottal stop in slow speech, but this may be replaced by an appropriate glide or even slight diphthongization in faster speech. Stress is phonemic in Stoney, and primary stress tends to fall on the penultimate syllable in longer words. Stress is not marked in the practical Stoney orthography, but is shown in this paper for clarity.

[4]This term can also be used between unrelated female friends.

[5]This kinship term means *my sister-in-law (of a male)* and does not typically refer to a cousin.

[6]This kinship term means *my brother-in-law (of a female)* and does not typically refer to a cousin.

Note that the familiar and children's forms are used to indicate intimacy and are typically only used between immediate family members. Likewise, usage of the familiar forms for cross-cousins can depend on the closeness of the relationship, and therefore they tend to be used to indicate strong family ties or in more informal situations.

3. Declaratives

3.1. Gendered enclitics in Stoney Nakoda

Stoney Nakoda has a set of gendered endings that are used in a number of different grammatical constructions. These are classified as enclitics rather than suffixes in part because they always occur at the very end of a word or phrase, unlike most other suffixes. They are not distinguished orthographically from suffixes, and they typically do not receive stress or interact with the prosodic system at all (they appear to be largely extrametrical). We will document a number of these enclitics in what follows, as they play an important role in expressing formal and informal speech in Stoney Nakoda. Unfortunately our investigation into these enclitics is still preliminary, and although we have included the most common gendered enclitics, there are undoubtedly others that have not been documented yet.

The Stoney Nakoda dialects appear to be quite unique among the Dakotan languages in that the gender-specific enclitics are primarily used for communication within genders, rather than to indicate the gender of the speaker. It is therefore not too surprising to learn that Stoney also has 'neutral' enclitics, used in communication between speakers of different genders or in general situations.

Including a declarative enclitic is obligatory for most declarative assertions made in the language, so they occur at very high frequency. It is possible the declarative enclitics also act as evidentials to some degree by implying that the speaker has firsthand knowledge of the proposition, or at least believes the statement to be true, but more research is required in order to determine if this is indeed the case. When used without any other tense or aspect-marking suffixes, the declarative enclitics additionally convey a progressive (ongoing) or retrospective (just completed) meaning, depending in part on the semantics of the verb in question. We have developed our own terminology to describe the specific semantics and pragmatics of these enclitics: *intergender* or 'neutral' enclitics are used between people of different genders; *intrafeminine* enclitics are used between women; and *intramasculine* enclitics are used primarily between men. Note that the latter two are often referred to informally as simply 'feminine' and 'masculine', but this can be misleading as they are primarily used only when the speaker and addressees are of the same gender.

The intergender declarative *-ch* is used most frequently, as situations with listeners of multiple genders are the most common. Even though the enclitics are divided by gender, there are situations where females can use the intramasculine declarative, such as when addressing close male relatives, when speaking formally to a mixed audience, or when a female elder addresses a male elder. Also note that gender usage is based on the gender that the speaker identifies with rather than biological sex, so double-spirited, transgender, and other gender identities can be expressed to some degree through the speaker's usage of declarative enclitics. In addition, younger females may use the intramasculine declarative *-no* for a kind of 'tom-boy' effect, although this

Table 2: Stoney Nakoda declarative enclitics

	Enclitic	**Usage**	**Abbreviation**
Intergender	*-ch*	• between genders • in generic situations	IG
Intrafeminine	*-chwe/-che*	• between women	IF
Intramasculine	*-no*	• between men • when women address close male relatives • between male and female elders • when speaking formally to a mixed audience	IM

does not appear to be very common. The pronunciation of the intrafeminine declarative enclitic varies between *-chwe* and *-che* within the community, most likely due to differences between families. Here are examples of all three declarative enclitics using the common greeting typically translated as 'good day'.[7]

(1) a. *Ấba wathtéch*
ấba wathté=ch
day to.be.beautiful=IG.DECL
'Good day, Hello'[8] (between interlocutors of different genders)

b. *Ấba wathtéche*
ấba wathté=che
day to.be.beautiful=IF.DECL
'Good day, Hello' (between women)

c. *Ấba wathténo*
ấba wathté=no
day to.be.beautiful=IM.DECL
'Good day, Hello' (between men)

Using the correct declarative enclitic for a particular situation is an important aspect of Stoney Nakoda culture, and if adult learners address same-gender elders using intergender enclitics they will typically be corrected. Also, intramasculine enclitics can be used when formally addressing an audience, or even when speaking more formally in a smaller group. Understanding how to employ these enclitics properly is critical to learning to speak Stoney, and so understanding these categories is integral to understanding many of the examples that follow. Here is a summary of the gendered enclitics that have been documented at this time, along with the affirmative and negative particles, which follow the same pattern of usage:

[7]Special abbreviations are required for the gendered enclitics: IG is used for intergender, IF for intrafeminine and IM for intramasculine.

[8]This phrase appears as though it could also mean 'it is a beautiful day', however that meaning typically requires a different translation since this construction is used as a greeting. The addition of the definite article *ne* 'the' can be employed to achieve that meaning: *ấba ne wathtéch* 'it is a beautiful day, the day is beautiful'.

Table 3: Some common gendered particles and enclitics in Stoney Nakoda

	Intergender	**Intrafeminine**	**Intramasculine**
affirmative particle	*hâ*	*hấwe*	*haw*[9]
negative particle	*hiyá*	*hiyáwe*	*hiyó*
declarative	*-ch*	*-chwe/-che*	*-no*
singular imperative	*-∅*	*-we*	*-wo*
plural imperative	*-m*	*-mwe*	*-bo*
nonpast polar interrogative	*-nî*	*-nîwe*	*-nûwo*
interrogative particle	*(hî)*[10]	*hîwe*	*hű̂wo*

3.2. Formal declaratives using *hâ*

When making declarative statements speakers can optionally add the verb *hâ* 'to be in an upright position, to stand' in order to be even more formal.

(2) *Ấba wathtéya* *hâch*
ấba watht*é*-ya hâ=ch
day to.be.beautiful-ADVR to.be.upright=IG.DECL
'Good day, Hello' (very formal, between interlocutors of different genders)

Adding a second verb requires the addition of an adverbializing suffix *-ya* to the initial verb *waththé-* 'to be beautiful', creating a complex verbal construction which might be more literally translated as 'standing beautifully' or simply 'existing beautifully'.[11] This type of construction can then be combined with the gendered enclitics to generate informal, somewhat formal, and very formal versions of even this basic greeting:

Table 4: Combining *hâ* with declarative enclitics to express formality

	'Good day, Hello'
informal between genders, used in a wide range of situations	*Ấba wathtéch*
very formal between genders	*Ấba wathtéya hâch*
informal between men or elders, or formally addressing an audience	*Ấba wathténo*
very formal between men, or very formally addressing an audience	*Ấba wathtéya hấno*

[9]Note that *aw* is used to represent the diphthong [aʊ]. Phonemic diphthongs are generally thought to be prohibited in Stoney Nakoda, but *haw* 'yes' is always pronounced with a diphthong even in slow careful speech.

[10]The intergender interrogative particle *hî* appears to be used primarily for forming tag questions and to be ungrammatical with open questions, but more investigation is required in order to confirm this.

[11]The suffix *-ya* is glossed as an adverbializer here, but it has other functions as well and generally facilitates building complex multi-verb constructions.

We see in this simple example that Stoney speakers can combine the gendered enclitics with auxiliary verbs like *hâ* to generate phrases with a wide range of formalities. These variations also provide some insight into the types of situations that are culturally salient, such as interactions within and between genders, between elders, and giving formal speeches. Note that the intramasculine enclitics are also used to express formality in certain situations, and so a more complete label might be *intramasculine/formal*.

4. Imperatives

4.1. Gendered imperative clitics

As seen in Table 3 above, there are five gendered imperative enclitics. Note that no enclitic is used to addressing a command to a single individual of a different gender, which is represented here by the null morpheme -∅. This means that bare verb roots are typically used as imperatives, and speakers frequently require a declarative enclitic (e.g. *-ch*, *-chwe/-che*, *-no*) when translating infinitive or generic forms of verbs into Stoney.

Table 5: Imperative enclitics

	Intergender	**Intrafeminine**	**Intramasculine**
singular imperative	*-∅*	*-we*	*-wo*
plural imperative	*-m*	*-mwe*	*-bo*

Note that while the intrafeminine plural is the combination of the intergender plural *-m* and the intrafeminine singular *-we*, the intramasculine replaces the *mw* sequence with a *b*. The following examples we see that the intramasculine singular imperative enclitic *-wo* can have different forms, depending on the formality of the situation:

(3) a. *ya*
ya=∅
go=IG.SG.IMP
'go!'[12]

b. *yáwe*
ya=we
go=IF.SG.IMP
'go!'

c. *yáwo*
ya=wo
go=IM.SG.IMP
'go!'

[12]Exclamation points are not used in the Stoney Nakoda forms because that punctuation mark is associated with a raised voice, which is considered to be both culturally inappropriate and unnecessary due to the existence of the imperative enclitics. Note that while plain imperatives can come across as somewhat terse or even rude in English, this is not generally the case in Stoney Nakoda.

d. *yo*
ya=wo
go=IM.SG.IMP
'go!'

In (3c) we see that the longer form *yáwo* is maintained in formal settings, but that this command can be reduced to a single syllable in informal settings by dropping the *a* and *w*, as seen in (3d). Note that this same pattern is seen in Table 3 when comparing the negative particles: intergender *hiyá*, intrafeminine *hiyáwe*, contrasted with intramasculine *hiyó*, where the expected *-aw-* sequence appears to have been elided. This pattern appears to occur with all verbs ending in a final *-a*, but may be more prevalent with high frequency stems. It does not appear to affect verbs ending in other vowels.

4.2. The imperative verb *hûjíya* 'come here'

As with some other Dakotan languages, some verbs are only used in imperative situations, such as when requesting or commanding others to come closer to the speaker. In Stoney Nakoda this is expressed using the imperative verb *hûjíya* 'come here', however the final *-ya* is only required in order to be polite (such as with elders), or in other formal situations.[13] This elision interacts with the gendered enclitics and the reduction seen in (3) to generate an extensive inventory of formal and informal realizations of this verb.

Table 6: Formal and informal variants of the imperative verb *hûjíya* 'come here'

		Formal	**Informal**
singular	intergender	*hûjíya*	*hûjí*
	intrafeminine	*hûjíyawe*	*hûjíwe*
	intramasculine	*hûjíyawo*	*hûjíwo*
plural	intergender	*hûjíyam*	*hûjím*
	intrafeminine	*hûjíyamwe*	*hûjímwe*
	intramasculine	*hûjíyabo*	*hûjíbo*

While these forms are generally predictable, there are a few interesting exceptions to take note of. First, the stress shifts to the first syllable in the informal variants, which is atypical for truncated stems in Stoney Nakoda. Secondly, the informal singular intramasculine form follows the same reduction pattern seen in (3d), where an *-aw-* sequence has been elided.

5. Interrogatives

In Table 3 above there are two sets of interrogative morphemes: the nonpast polar interrogative enclitics used for creating yes/no questions, and the interrogative particles *hî*, *hîwe*, and *hű́wo* that

[13]Note that *hûjíya* can only be used in situations where the listener is physically nearby and can respond promptly; it is judged as ungrammatical when used over the phone, for example. An elaborated translation such as 'come here (from your nearby position)' might more faithfully capture the meaning of this verb.

can optionally be used with open questions (similar to *wh*-questions in English). The latter are considered to be independent particles rather than enclitics because they bear their own intrinsic primary stresses, however they otherwise function much like the enclitics. The other important distinction is that the interrogative particles are optional, and are generally only used in formal situations.

(4) a. *Dokén yaű?*
doken ya-û
how 2SG-to.feel
'How are you?' (informal)

b. *Dokén yaű hîwe/hű́wo?*
doken ya-û hí́we/hú́wo
how 2SG-to.feel Q
'How are you?' (formal between women/between men)

Note that this usage is distinct from other Dakotan languages, where the interrogative particles are typically mandatory, even in open questions (also known as *wh*-questions, or *d*-questions in Stoney Nakoda since all of the interrogative terms start with *d*). The intergender particle *hî* seems to be used primarily for forming tag questions, and does not appear to serve as a conventional interrogative particle at all, however more investigation is required in order to confirm this.

Note the *î* to *û* vowel change in both of the intramasculine interrogatives: intergender *-nî* and *hî* become intramasculine *-nûwo* and *hű́wo*. While there are other cases of similar vowel changes particularly in informal speech, this is not a regular sound change throughout the language and the sequence *-îw-* does occur both within words and as a result of morphological processes (e.g. *îá* 'to speak, to speak a language', *îwá'ach* 'I'm speaking, I speak it').

6. Conjunctions and complement clauses

6.1. Conjunctions

By using different combinations of conjunctions, or by dropping them altogether, Stoney Nakoda speakers can indicate different degrees of formality. This is seen below, where the conjunction *gichí* 'with' can be dropped in very informal situations when only a pronoun and a proper name are being conjoined.

(5) a. *Mîyé Valerie gichí*
1SG.PRO Valerie with
'Valarie and I' (standard)

b. *Mîyé Valerie*
1SG.PRO Valerie
'Valarie and I' (very informal)

For longer lists, *gichí* is typically required, and additional conjunctions can optionally be included in order to indicate increased formality.

(6) a. *Valerie Carson Shanice gichí*
Valerie Carson Shanice with
'Valerie, Carson and Shanice' (standard)

b. *Valerie nâgú Carson ehűna, nâgú Shanice gichí*
Valerie CONJ Carson conj CONJ Shanice with
'Valerie, Carson and Shanice' (very formal)

Conjunctions can be difficult to translate, so they are not given precise glosses in (6b). Individual conjunction lexemes in Stoney often match best with short phrases in English, so we suggest the following translations for the conjunctions in (6b): *nâgú* 'in addition to, and'; *ehűna* 'also including'. Adding these conjunctions appears to be a way of embellishing an utterance, so a more stylistic gloss of this phrase might be: 'In addition to Valerie, Carson is also included, along with Shanice'. Note that while this translation might come across as somewhat long-winded in English, adding these conjunctions is simply a way of expressing formality in Stoney Nakoda, and does not imply an overly pretentious attitude.

6.2. Complement clauses

Similar to the conjunction *gichí*, the complementizer *-cha* can be dropped in informal situations:

(7) a. *dohâ hînîgecha wakách*
dohâ hînîga=cha[14] wa-ka=ch
INTS[15] be.bad=COMP 1SG-to.mean=IG.DECL
'I mean that it was terrible' (standard)

b. *dohâ hînîga wakách*
dohâ hînîga wa-ka=ch
INTS be.bad 1SG-to.mean=IG.DECL
'I mean (that) it was terrible' (informal)

Dropping the complementizer also results in the use of the unablauted form of *hînîga* 'to be bad'. More research is required in order to clarify the details of these patterns, as they may not hold across all situations.

7. Conclusion

In this paper we have seen a wide range of linguistic patterns that are used to express different degrees of formality in Stoney Nakoda. The choices speakers make with regard to kinship terms, gendered enclitics, conjunctions, complementizers and secondary verbs can all be used to indicate

[13]The complementizer *-cha* is treated as an enclitic in part because it does not receive stress. It also appears to trigger e-ablaut in Stoney Nakoda (contrary to some other Dakotan languages), causing *hînîga* 'to be bad' to become *hînîge* in this phrase.

[15]The intensifier *dohâ* can typically be translated as 'very, really', but it frequently interacts with the semantics of the words it modifies, in this case the verb *hînîga* 'to be bad'. It is probably more realistic to treat this combination as a verb phrase: *dohâ hînîga* 'to be terrible, to be awful, to be ugly, to be very bad'.

respect, the formality of the situation, or the familiarity between interlocutors. Many of the less formal patterns may appear to simply be reduced speech, however we want to emphasize that the Stoney Nakoda language is still used on a daily basis in these communities, including in formal situations such as feasts, eulogies, ceremonies, and powwows, as well as informal conversations. These different forms are frequently used by individual speakers, and elders often comment that different forms are acceptable depending on the situation. Since Stoney Nakoda is still so widely-spoken in the community, we felt it was important to document some of these usages in order to demonstrate the wide range of variation and to assist language learners in their effort to maintain the Stoney Nakoda language.

References

Bellam, Ernest Jay. 1975. *Studies in Stoney morphology and phonology*. Calgary, AB: University of Calgary M.A. thesis.

Cumberland, Linda A. 2005. *A grammar of Assiniboine: A Siouan language of the northern plains*. Bloomington, IN: Indiana University dissertation.

Harbeck, Warren. 1980. *Some agreement phenomena in Stoney*. Calgary, AB: University of Calgary M.A. thesis.

Shaw, Patricia A. 1980. *Theoretical issues in Dakota phonology and morphology*. London, UK: Routledge.

Shaw, Patricia A. 1985. Coexistent and competing stress rules in Stoney (Dakota). *International Journal of American Linguistics* 51(1). 1–18.

Statistics Canada. 2016. *2016 Census aboriginal community portrait – Stoney Tribe*. Census of Population. Catalogue Number: 412600001. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/abpopprof/infogrph/infgrph.cfm?LANG=E&DGUID=2016C1005454&PR=48 (20 May, 2021).

Ullrich, Jan. 2018. *Modification, secondary predication and multi-verb constructions in Lakota*. Düsseldorf, Germany: Heinrich Heine University Düsseldorf dissertation.

Ullrich, Jan & Jr. Black Bear Ben. 2016. *Lakota grammar handbook*. 2nd edn. Bloomington, IN: Lakota Language Consortium.

(Para-)Adpositional Morphosyntax in Siouan: A Case Study of Lakhota-Dakota-Nakota, Catawba, and Crow*

Noah Michael Coen
University of California, Los Angeles

Abstract: This paper examines the morphosyntactic status of adpositions and their kin in three Siouan languages—the Lakhota-Dakota-Nakota variety continuum, Crow, and Catawba—in hopes of providing a more nuanced account of the grammar of adpositions in Siouan. The data and analyses herein illustrate the need for linguistic examinations of Siouan adpositions to include applicative systems, as well as demonstrate that the descriptive work to date has been insufficient in regard to adpositional and applicative morphosyntax in these languages. This paper evidences far more diversity within the adpositional morphosyntax of Western Siouan languages than typically granted. Additionally, it demonstrates that the Eastern Siouan branch is not as dissimilar an outlier as it is often portrayed.

Keywords: adpositions, applicatives, Catawba, Crow, Lakhota, Dakota, Nakota

1. Introduction

1.1. Background

The data and analysis presented here explore the morphosyntactic status and behavior of adpositions and related constructions in three phylogenetically distant Siouan languages: the Lakhota-Dakota-Nakota language continuum (LDN henceforth), Catawba, and Crow. The result of this exploration is a more nuanced description of the grammar of adpositions in Siouan. Adpositional morphosyntax is understudied and often overlooked by linguists. Hagège (2010) draws attention to this throughout his monograph on the typology of adpositions. In addition to true adpositions, this study investigates applicatives. Throughout this paper, I refer to Siouan applicatives as an example of para-adpositional morphology. I do so because they are intimately related to adpositions and, as I discuss in section 1.3, are not true applicatives (*in Siouan*). Many Siouan languages do not have an extensive history of formal descriptions; this is especially true out-

*I would like to thank Dr. Ryan Kasak for supervising the independent study that led to this paper, meeting with me twice a week despite the difficulties of both the COVID-19 pandemic and it being his first semester at the institution.

side of the Dakotan subbranch.[1] Moreover, research on Siouan languages conducted before the 1960s is often difficult to parse, as authors adhered to unique, individualistic systems of phonetic description. This paper lies at the intersection of these understudied areas.

The Siouan languages—which constitute one of the world's primary language families—are traditionally split into two groups: Eastern (Catawban) and Western (Siouan "Proper"). The Eastern Siouan group split off from Proto-Siouan as long as 4,000 years ago and contains only two known languages: Catawba and Woccon, the latter of which is poorly attested (Kasak 2016:7, Rudin & Gordon 2016:3). The first linguistic group to separate from Proto-Western Siouan was the Missouri River Valley subbranch, whose modern descendants are Crow and Hidatsa. This split was followed by the Mandan language,[2]

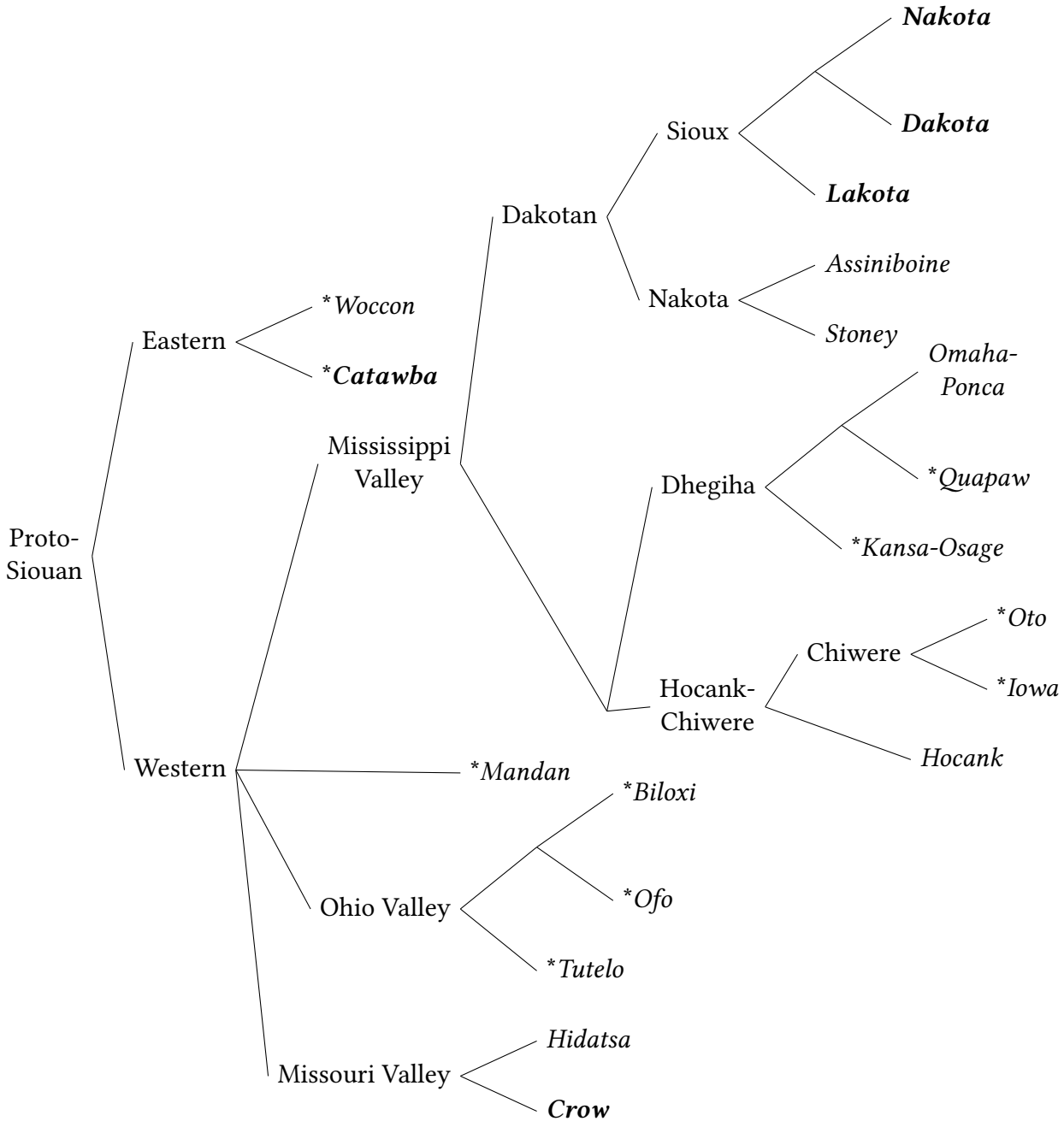
The relationships described above and illustrated in figure 1 are still being refined and reanalyzed. Yuchi—a language isolate spoken in present-day Oklahoma—has long been postulated to be a distant member of the Siouan family (notably by Sapir in 1929), but this theory is not widely accepted (Kasak 2016, Mithun 1999). In a recent manuscript, Kasak (2020a) augments the evidence in support of this postulation, providing two computational analyses of Siouan-Catawban-Yuchi phylogeny. Kasak's findings indicate not only that Mandan is more closely related to Crow and Hidatsa than to any other Siouan language, but also that Yuchi appears to be much more closely related to the Catawban languages (and thus the Siouan languages) than previously thought. Both new models suggest that Yuchi should be considered a subbranch of Eastern Siouan (Catawban), not a third subbranch of Proto-Siouan.

Adpositions, though nearly ubiquitous in human language, have not been subject to the same intensity and rigor of linguistic evaluation that many other word classes such as nouns and verbs have. Hagège (2010:2) goes as far as stating that his book was the first published monograph focused on adpositions and the typology thereof. However, much meaningful linguistic research on adpositions had been conducted before his work. Asbery's dissertation (2008) focused on the morphosyntax of case and adpositions. Hewson & Bubenik (2006) proposed a diachronic account of adpositions in the Indo-European language family. Hagège mentions the work of Kurzon & Adler (2008), but simply remarks that their work has a narrower scope than his own. However, it is their claim that extensive further research is necessary for linguists to arrive at an adequate theory of adpositional morphosyntax that Hagège appears to be echoing Kurzon & Adler (2008:1-3). The aforementioned research gap is certainly present in Siouan. There has yet to be a comparative study that examines adpositions in the Siouan language family. The *raison d'être* of this study is to contribute synchronic analyses of the (para-)adpositional morphosyntax of LDN, Catawba, and Crow that—when examined together—provide insights for the study of historical and comparative Siouan linguistics, as well as the typological study of adpositions more generally.

[1]That being said, Siouan linguistics has flourished over the past half-century, contributing not only to our understanding of Siouan languages but also to our understanding of linguistic theory. This period of scholarly vibrancy produced many of the works referenced herein.

[2]It should be noted that Mandan is frequently linked as a group with Crow and Hidatsa (Kasak 2016:8) which was in turn followed by the branching of the Ohio River Valley (Southeastern) and Mississippi River Valley subgroups. The Dakotan languages were the first to diverge from the Mississippi River Valley subbranch. Subsequently, the Winnebago-Chiwere and Dhegiha constituents of the Mississippi River Valley subbranch separated, forming their own subfamilies (Rudin & Gordon 2016:3, Park 2012:1-2). This is illustrated in figure 1, which uses data from Ullrich (2018:23), Rudin & Gordon (2016:3), and Kasak (2016:8).[3]

Figure 1: Siouan phylogeny



Siouan languages exhibit head-final syntactic structure; thus, free adpositions in these languages are postpositions rather than prepositions. Siouan adpositions appear to undergo enclisis in many of the family's languages; however, as discussed in sections 4 and 5, this is not universal. A variety of other adpositional phenomena are evidenced in Siouan herein, including proclisis, various combinational phenomena, and movement out of an adpositional phrase, inter alia.

In addition to true adpositions, this study investigates a set of Siouan preverbs referred

to as applicatives that are markedly similar to adpositions in both their morphosyntactic and semantic functions. I consider these an example of "para-adpositional" constructions, as their close relationship with adpositions is both diachronic and synchronic. In theoretical morphology, applicatives mandatorily increase the valency of a verb; they are often used to topicalize an oblique argument (Peterson 2007:1-3). However, expansion of a verb's argument structure does *not* always occur with Siouan "applicatives." Thus, as discussed further in section 1.3, these are not true applicatives, either. Some Siouanists, such as Kasak (2019), are beginning to consistently refer to these as preverbs, preferring the more correct and theory-neutral term. However, the vast majority of sources referred to herein refer to these as "applicatives."

There are four standard applicatives in Siouan languages (Helmbrecht 2006). The fourth of these, the benefactive, is not overt in Catawba or Crow; thus, this paper focuses on the three "locative" applicatives: the superessive, the inessive, and the instrumental. The superessive applicative most often denotes spatial location 'on top of' or 'above' something else. The inessive applicative typically corresponds to 'inside' or 'into' (Helmbrecht 2006). Finally, the instrumental represents a non-comitative instrumental relationship; however, it can also be used as a locative, meaning "against."

Helmbrecht & Lehmann (2008) propose a chronology of the development of internal affixation in Siouan based on their theory of isolated stem components (ISCs). For the purposes of this paper, it is not important to understand their theory of ISCs or the nuances of Siouan verbal morphology. It is only necessary to note that applicatives are one posited source of these components. Helmbrecht and Lehmann's conclusion delineates four stages that Siouan languages underwent in the development of ISCs. The relevant three are delineated below.

In stage one, Proto-Siouan, they claim that the now-grammaticalized applicatives were a preverbal constituent (such as a postposition). Helmbrecht and Lehmann do not assign a time-frame to stage two, stating only that it is still a reconstructed form; in this stage, they claim the aforementioned postpositions became "preverbs." Although "preverbs" often denote applicatives in Siouanist literature, here Helmbrecht and Lehmann are discussing proclitics. This is because in stage three—which also has no assigned time-frame but is said to have been 'historically observable in Hocąk and other Siouan language'"—they claim these "preverb" had become applicatives. This is presented in table 1 (Helmbrecht & Lehmann 2008:34-35).

Table 1: Helmbrecht and Lehmann's Four Stages

	Stage One	**Stage Two**	**Stage Three**	**Stage Four**
Time:	Proto-Siouan	*Not Stated*	*Not Stated*	Present
Status:	Postpositions	Proclitcs	Applicatives	ISCs

Helmbrecht and Lehmann's conclusion suffers from a lack of specificity. For instance, they state that their findings apply to "Hocąk and other Siouan languages," which indicates a broad interpretation. However, it is their claim regarding the stages in table 1 that is of particular importance to this paper. The picture Helmbrecht and Lehmann paint is one of clear-cut phases, with different word classes and morpheme types having diachronic relationships, but synchronic independence. This prompts a closer investigation of Siouan languages other than Hocąk in order

to determine how distinct these phases truly are.[4]

1.2. Towards a More Nuanced Typology

This survey of the (para-)adpositional systems in LDN, Catawba, and Crow reveals that adpositional morphosyntax has not been closely analyzed by Siouanists to date. The analysis herein of LDN—the most thoroughly documented Siouan language and that with the longest history of linguistic research—resulted in a number of novel findings. For instance, a group of discrete combinational processes involving adpositions are all described as 'incorporation' in the Siouanist literature. While LDN does exhibit true incorporation elsewhere in its morphosyntax, it does not in either of the combinational processes involving adpositions. The primary phenomena present in LDN's adpositional morphosyntax are compounding and enclisis. Most Siouanists have chosen to use 'incorporation' as an all-encompassing term that allows them to present the data without simultaneously presenting a morphosyntactic analysis. This paper thus presents a closer examination. The analysis in section 2 shows that, although historically related to applicatives, adpositions in LDN are synchronically distinct from them, supporting Helmbrecht & Lehmann's (2008) paradigm discussed above. Moreover, due to the aforementioned comparative robustness of research on LDN, it has been used as a template of sorts in descriptions of other Siouan languages. However, as demonstrated below, this can result in infelicitous analyses.

Catawba's adpositional morphosyntax has largely been ignored by the handful of scholars who have examined the language. Its use of proclisis is attested in the literature, but most work on Catawba—Rudes (2007) being an important exception—has involved lexical indexing rather than grammatical analysis. This is not surprising, as Catawba is primarily attested by Speck's (1934) transcriptions of folktales and lexicographical work is often a prerequisite for grammatical linguistic analysis. Moreover, apart from one article on onomastics by the late Blair Rudes, the Catawba language has not been the subject of published research in the twenty-first century. The analysis of Catawba herein serves to modestly remedy that, providing a novel analysis of its adpositional morphosyntax, as well as fodder for further research on the language. Catawba is regularly neglected in literature that claims to characterize the Siouan language family—such as Helmbrecht (2006)—as it differs significantly from many of its linguistic relatives.[5] Ignoring the Eastern branch of Siouan languages allows for cleaner conclusions to be drawn, but it invalidates wide-reaching claims about the entire language family. Notably, Catawba does not have overt applicatives, contra the claim in Helmbrecht (2006) that all Siouan languages do.

The analysis of Crow herein further demonstrates that the problem of underdescribed adpositional systems is endemic to the "core" (Western) Siouan languages, as well—not just the Catawban (Eastern) branch. Crow has a remarkably flexible system of adpositional morphosyntax, allowing left-anchored, right-anchored, and bidirectionally-anchored compounding in addition to free-standing postpositions. This constitutes a rejection of the attempted general characterizations of the Siouan family from within the "core" (Western) Siouan branch itself. The analysis of several unexplained (or insufficiently explained) adpositional structures in previously-

[4]Note that this is not the focus of Helmbrecht & Lehmann (2008) conclusion; rather, it is an implication they make en route to their conclusion about ISCs and verbal morphology.

[5]Kasak (2020a) and this paper argue that this difference may not be as extreme as previously thought. Furthermore, I believe the *perception* that Catawba is only peripherally related to the "core" Siouan languages contributes as much to Catawba's de facto exclusion as the actual linguistic variation.

elicited data results in positing the presence of topicalization movement, which helps account for irregularities noted by scholars in the past. Additionally, the boundary between adpositions and applicatives in Crow is blurry at best, suggesting that a more nuanced analysis of Siouan applicatives—one that analyzes them as para-adpositional—is necessary.

1.3. Theoretical Orientation

There are several aspects of this study that require preemptive clarification: the conceptualization of a "word," the conventions of syntactic notation used, the parameters of the combinational phenomena discussed, and the usage of the term 'adposition.' The first two of these elements lie at the center of intense, ongoing theoretical investigation and debate. This paper does not make cross-linguistic claims about the nature of wordhood, nor about the innate human faculty for language and its best syntactic representation. Nonetheless, it must adopt frameworks for both aforementioned components.

In this paper, I identify two discrete categories of 'word': prosodic words and morphological words. Prosodic words are defined herein as sentential constituents that have a single primary lexical stress. It is critical to note that this definition refers to lexical stress alone, not phrasal pitch accent or prosodic emphasis. This distinction is especially important for the discussion of Catawba in section 4.

Morphological words are defined herein as a group of one or more morphemes that always co-occur in the same pattern and that are synchronically unanalyzable. This conceptualization is strongly influenced by Dixon & Aikhenvald's (2003:18-25) notion of "grammatical words."

The syntactic notation used throughout this paper is best described as a kind of "pseudo-minimalism." There are several space-consuming syntactic representations within this paper, which caused formatting issues when using "pure" X-bar theory. Switching to a paradigm more closely aligned with the minimalist program allowed these formatting issues to be resolved without sacrificing any substance or altering any theoretical claim.

There are a variety of morphosyntactic phenomena in Siouan that involve combining more than one morpheme to create a single 'word.' Such processes are almost exclusively referred to as "incorporation" in Siouanist literature to date. Olthof's dissertation on incorporation defines the phenomenon as "the inclusion of one lexical element in another lexical element such that they together constitute a single word" (Olthof 2020:71, 131-132). The key word in Olthof's description is '*in*'; English words like 'firetruck' and 'bookstore' do not fall into this category. Olthof gives the following example from Chukchi (Olthof 2020:53). In example (1), incorporation is not present ('to catch' [the hare]). Example (2) expresses an almost identical meaning using an incorporation construction ('to hare-catch').

(1) *ʔətt-e* *piri-nin-∅* *melota-lɣən*
dog-ERG catch-3SG>3SG-PST hare-ABS.SG
'The dog caught the hare.'

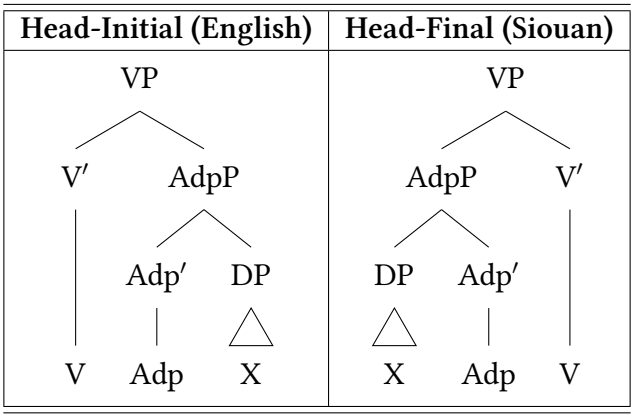
(2) *ʔətt-ən* *milute-piri-ɣʔi-∅*
dog-ABS hare-catch-3SG.SBJ-PST
'The dog caught a hare.'

Siouan languages do exhibit incorporation—e.g. the word /cʰąlí-wakpà/ 'to tobacco-cut' (to cut tobacco) in LDN (Boas & Deloria 1939:70)—but not in their adpositional morphosyntax.[6] Other phenomena, namely compounding and cliticization, do. Compounding is a phenomenon in which two constituents, each with their own primary lexical stress, merge to form one prosodic word. The new primary lexical stress may fall on either original constituent of the newly formed compound.

Cliticization is a similar, but discrete morphological operation. In the examples of enclisis discussed herein, an adpositional enclitic is attached leftward, onto its governed term. Proclitics differ from these by attaching rightward, onto the verb dominating the adposition. Unlike constituents of compounds, clitics can never be stressed.[7] Neither compounding nor cliticization is true incorporation.

An adposition is defined within this paper as follows: a maximal projection that forms an adpositional phrase with the determiner phrase (DP) it governs and that denotes a relationship between that governed DP and the phrase that most immediately dominates it. This is generally a verb phrase (VP), but it can also be a noun phrase (NP) or an adjective phrase (AP) (Hagège 2010:8). This conceptualization of adpositions is rather uncontroversial. The theoretical foundations of this definition are influenced by Hagège (2010); however, it does not conform to the terminology used therein.[8] The structure described above is perhaps clearer when illustrated rather than articulated. This is depicted in figure 2.

Figure 2: Prototypical Adpositional Phrases



Hagège (2010) claims that adpositions are, minimally, unique morphological words. The data presented here support this claim if one subscribes to the definition of a morphological word delineated above which, in accordance with both Dixon & Aikhenvald (2003:24-27) and Booij (2005:202), includes clitics. However, the Siouan data contradict this if one does not designate clitics as morphological words. Hagège himself does not assign the label of 'adposition' to preverbs and clitics (Hagège 2010:62-63). This illustrates the importance of considering this paper's theoretical assumptions when examining the claims herein.

[6]The situation in Crow is more complex, but I argue that it does not appear there, either.

[7]This rule does have exceptions. For example, an enclitic can be stressed in Modern Greek if a second enclitic is attached to it. See Anderson (1992) and Anderson (2005:24) for more information.

[8]For example, Hagège refers to the VP (or NP or AP) that dominates the adverbial phrase as a "head," which is a non-standard description. Throughout this paper, I will use "head" to refer to a maximal projection.

1.4. Notes on Research Methodology

In addition to the academic literature on Siouan linguistics to date, the principal sources referenced herein are transcriptions of folktales and other narratives told in LDN, Catawba, and Crow. It is often difficult to discern whether transcribed texts accurately portray phonetic reality. For example, in the Speck (1934) texts—the main source of Catawba data—primary, secondary, and tertiary stress are not orthographically differentiated. As highlighted by the discussion of stress above, prosodic data are often vital when conducting morphosyntactic analysis; the absence of this information is discussed where relevant throughout. Regardless, large quantities of these data can provide researchers with phonological and prosodic insights, allowing us to produce salient analyses. This is expanded upon significantly in section 4.

Despite the difficulties it presents, written material is a captivating medium for linguistic research—particularly attempts at recording oral traditions in the realm of folklore, mythology, and fables. These genres are the central sources of extant texts in the Siouan languages discussed herein, as well as in many other understudied languages. The registers used in these texts differ from the register of casual speech. Thus, it must be noted that morphosyntactic phenomena identified from analysis of these genres could result from the language play typical of storytelling and narration.[9]

2. Evidence from Lakhota-Dakota-Nakota

2.1. Overview

The LDN variety continuum is perhaps the best-documented of the Siouan languages, and this is true of its adpositional system, too. However, most work on adpositions in LDN is concerned with their free postpositional forms and the integration of pronominal elements therein. Far less studied are the processes of compounding and enclisis, which are seldom discussed in the Siouanist literature to date. In contrast with Catawba and Crow, LDN does not exhibit proclitic or preverbal constructions in its adpositional morphosyntax.[10] Moreover, LDN utilizes a robust system of locative applicatives that are grammatical both with and without a preceding postpositional phrase, though it is not clear whether there is productive semantic variation between using solely a postposition, solely an applicative, or using both. From the data analyzed within this study, the choice appears to be lexically determined. These applicatives, although semantically and historically related to postpositions, do appear to support Helmbrecht & Lehmann's (2008) theory that these grammatical constituents should be synchronically treated as discrete phenomena.

[9]Language death is occurring rapidly world-wide. As such, analysis of (often less-than-ideal) archive material is becoming increasingly important for the field of language documentation (Bowern 2018). Thus, written material is not only vital for philologists; it is also increasingly relevant in the fields of language documentation and revitalization.

[10]There is one potential example of proclisis (other than the applicatives) that is now fossilized in a verb stem. The word /akąyaka/, 'to ride,' is more precisely glossed as [on=sit], but it does not appear to be synchronically analyzed as such by speakers (Deloria 1932:237).

2.1.1. The Derivation of Adpositions

There are at least two common sources of adpositional derivation in LDN: adverbs and verbs. Adpositions can be derived from adverbs via the addition of the prefix /i-/ (Ingham 2003:41, Ullrich 2018:62). Consider the following examples.

(3) a. *hakáb* (ADV)
afterwards
b. *ihakáb* (ADP)
behind, after

(4) a. *mahél* (ADV)
inside
b. *imáhel* (ADP)
inside

In example (3), the adposition 'behind, after' is derived from the adverb 'afterwards' via the addition of the prefix /i-/ (LLC 2021:41, Ingham 2003). This derivational prefix has become so productive that native speakers sometimes add /i-/ to lexemes that are already free, non-derived postpositional forms. This is evidenced by example (4), in which the word-class of /mahél/ does not change, but the prefix /i-/ is still added and the locus of lexical stress is subsequently shifted (Ingham 2003:41).

(5) a. *iyúweğa* (V.INF)
to.cross.over
b. *iyúweȟ* (ADP-like)
across

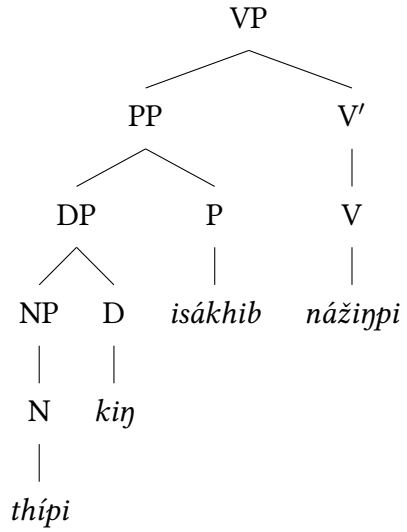
Another source of adpositional derivation is the verb, as shown in example set (5). This process is described by Ingham (2003:41) as "plain stems of verbs... [being] used in a participle-like construction. [These] can be regarded as in a transitional status [between verbs and adpositions]." Very little has been written about this derivational process and further research is needed to provide an adequate description. Unlike /i-/ prefixing, this does not appear to be a productive process.

2.2. Free Postpositions

Free postpositions are the most common type of adposition in the LDN data examined herein. The tree below depicts a simple postpositional phrase and its clausal environment.

(6) Simple postpositional phrase within a clause
a. Gloss
thípi kiŋ isákhib nážiŋpi
house DEF next.to they.stand.PST
'They stood next to the house.' (Ullrich & Black Bear 2016:380)

b. Syntactic diagram



(7) *maza oŋ*
iron of
'[made] of iron' (Riggs 1895:52, 77)

(8) *thiyópa ikhíyela yaŋké.*
door near sit.PST
"He sat near [the] door." (Ullrich 2018:380, Ingham 2001:91, 220)

This type of postposition is thoroughly attested. The following are a brief selection of the postpositions that can be found in the Deloria texts: /étkiya/ 'towards,' /ogná/ 'through,' /ekta/ 'to,' /etą́hą/ 'from,' /op^{h}áya/ 'along (1),' /oȟlathe/ 'below,' and /aglágla/ 'along (2),' inter alia (Deloria 1932:19, 28, 30, 65, 213, 234, 267). Note that the use of a postposition decreases the likelihood that the governed DP will contain an overt determiner head. Example (6) depicts an exceptional case in which the definite marker is used, as this can occur (Ingham 2003:40). Example (8) depicts a case in which the use of an adposition results in a null D-head.

(9) *ilázata*
i-lázata
against-behind
'behind' (Deloria 1932:109)

(10) *ilázatalaȟcį*
i-lázata=laȟcį
against-behind=INTENSE
'directly behind' (Deloria 1932:246)

Section 2.1.1 illustrated that the prefix /i-/ could derive adpositions from adverbs. However, as will be discussed in section 2.4, the prefix /i-/ is also an applicative in LDN. One of the senses of this applicative is the locative 'against.' This marker can be attached to postpositions to alter or emphasize the spatial aspect of their semantics. This is the case in example (9). The postposition /lázata/ by itself means 'behind' (Deloria 1932:67); the addition of the locative /i-/

adds the sense of being up against something, directly behind it. In example (10), the intensifier enclitic /=laȟcį/ is attached to /ilázata/ 'behind,' illustrating the grammaticality of attaching enclitics to free postpositions in LDN.

Furthermore, in addition to following determiner phrases, LDN's postpositions can follow stative verbs. However, the syntactic processes that result in this surface structure have not yet been analyzed. Two plausible analyses are explored below.

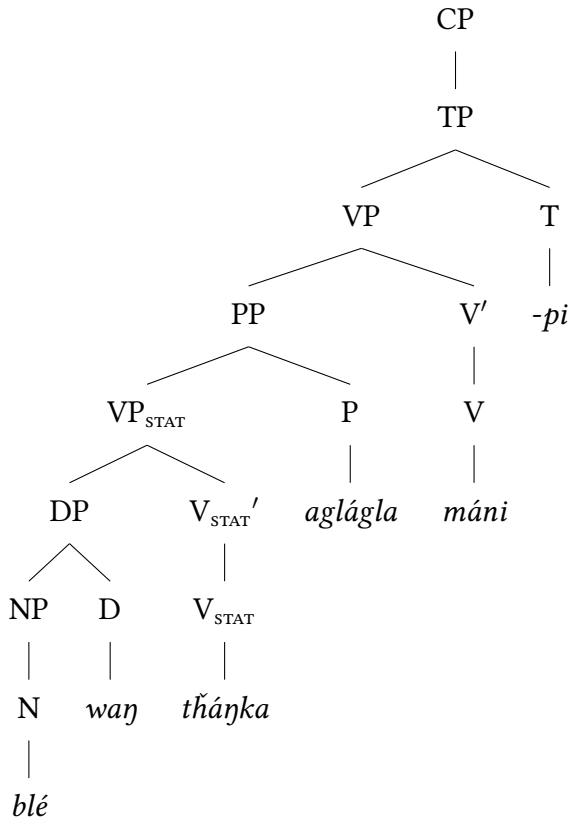
(11) Stative Verbs Preceding Adpositions

a. Gloss

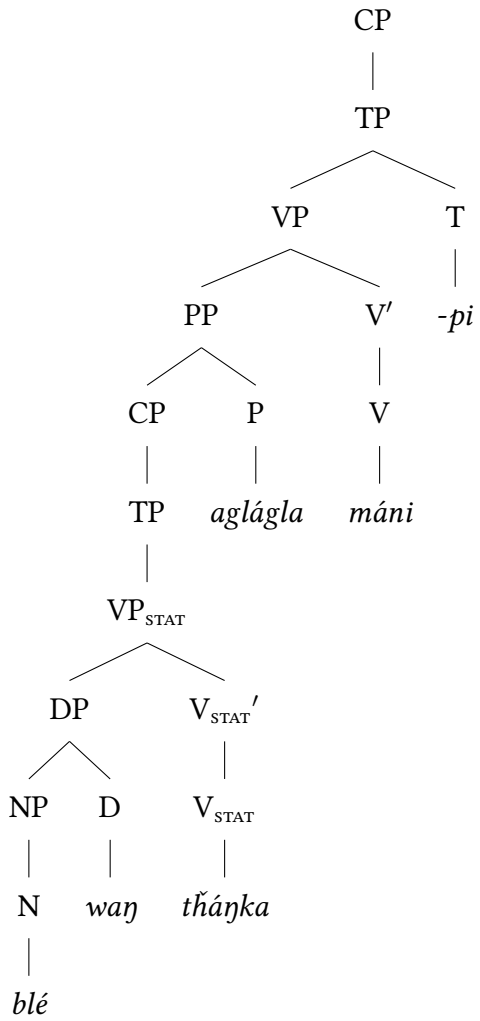
blé waŋ tȟáŋka aglágla mánipi
lake INDEF big.V.STAT along walk.PST.PL

'They walked along a big lake.' (Ullrich 2018:143)

b. Adpositionally governed VP_{STAT}



c. CP complement of PP



In example (11b), the VP_{STAT} is the constituent governed by the postposition. As the stative verb functions adjectivally, one might expect this constituent to fall within a determiner phrase, which would in turn be governed by the postposition. This type of DP-structure *is* grammatical in LDN; an example is shown in (12), below.

(12) *Hokšíla čikčík'ala kiŋ*
Hokšíla čik∼čík'ala kiŋ
boy R∼to.be.small DEF
'the small boy' (Ullrich 2020a:412)

Note that the stative verb in example (12) lies between the noun and the determiner. However, this is not the syntactic structure exhibited in example (11). Consequently, other analyses must be explored. The strongest alternative analysis is that postpositional phrases in LDN can take CP complements. A tenseless, non-finite clause governed by a postpositional head provides a salient explanation for this phenomenon. This structure is illustrated in example (11c).[11] How-

[11]The existence of CP complements in adpositional phrases is not unique to Siouan; for example, we find them in Dutch (Broekhuis 2015).

ever, this is not the only possible analysis. For example, it is plausible that the stative verb is an extra-syntactic parenthetical. More data needs to be elicited for further research.

It is worth noting that free adpositions in LDN occasionally lack an overt governed phrase. This occurs only when there is information from earlier in the discourse that allows the participants of the conversation to infer the entity being discussed. This happens in English, as well. If someone says, "We went across." in a discussion about a creek or a bridge, it is clear that the speaker went *across the creek* or *across the bridge*, respectively.

2.2.1. Pronominal Affixation

Another feature of many free postpositions in LDN is pronominal affixation.[12] This results in several noteworthy morphosyntactic and morphophonological phenomena. Pustet (2000:180) claims the postpositions that *can* adjoin with pronouns have narrower semantic scopes than those that cannot, which often have non-specific locative meanings.

(13) Plural marking of affixed postpositional patients

a. Singular

nihakab *iyaye*
ni-hakab iyaye
2SG.PAT-behind go

'They.SG are walking behind you.' (Pustet 2000:162)

b. Plural

nihakab *iyayapi*
ni-hakab iyaya-pi
2SG.PAT-behind go-PL

'They.SG are walking behind you guys.' (Pustet 2000:162)

Note that if the pronoun is plural, the person marker prefixes to the postposition, but the plural marker suffixes onto the verb (Pustet 2000:162). If the pronominal prefix ends with the same vowel that the postposition begins with, elision will occur (Pustet 2000:161-162).

(14) *ihakab ųkiyaye*
ihakab ųk-iyaye
behind DU.PAT-go
'They.SG are walking behind the two of us.' (Pustet 2000:165)

(15) **etkiya pak^{h}ab iyemaniye*
etkiya pak^{h}ab iye-ma-ni-ye
toward push send-1SG.PAT-2SG.PAT-CAUS
'They.SG pushed me toward you OR ...you toward me.' (Pustet 2000:165)

[12]The affixation of pronouns onto word classes other than verbs is common among Siouan languages (Kasak 2020b).

Another way to denote the person(s) governed by a postposition is to add a concordant patient marker onto the verb (Pustet 2000:164-165). However, this becomes ungrammatical if the verb itself has a patient. This is depicted in examples (14) and (15).

(16) *miye etkiya pak^{h}ab iyeniye*
miye etkiya pak^{h}ab iye-ni-ye
1SG.PAT toward push send-2SG.PAT-CAUS
'They.SG pushed you toward me.' (Pustet 2000:166)

(17) *miye c^{h}a ihakab iyaye*
1SG.PAT EMPH behind go
'It is I that they.SG were walking behind.' (Pustet 2000:168)

One method of solving this "problem" is to use an independent pronominal patient marker for the adposition's governed entity. This is shown in example (16), above (Pustet 2000:166). However, the use of independent pronouns is not confined to situations in which both the adposition and verb have a patient. This is also the construction used in tandem with the emphatic particle *c^{h}a* to denote focus, as shown in example (17) (Pustet 2000:168).

(18) *wic^{h}ihakab iwic^{h}ayaye*
wic^{h}i-ihakab i-wic^{h}a-yaye
3PL.PAT-behind INS-3PL.PAT-go
'They.SG are walking behind them.' (Pustet 2000:168).

When only the postposition has a patient, patienthood can be marked doubly—on both the postposition and the verb. Pustet makes no claims about the semantic effect this elicits, noting that previous researchers seem to have ignored this construction (Pustet 2000:168). This presents a fascinating area for future research at the morphosyntax-semantics interface.

2.3. Combinational Phenomena

2.3.1. Preliminaries

Both N+ADP compounding and adpositional enclisis are grammatical processes in LDN. Siouanist literature does not discuss these compositional phenomena in any detail, instead choosing to group these related processes under the title 'incorporation.'[13] Compounding and enclisis are prosodically distinct, meaning they differ at the suprasegmental level. In compounding, a noun and an adposition with individual primary stresses merge to form one prosodic word with one primary lexical stress, which can fall on a nucleus from either original constituent. In enclisis, however, the adposition is prosodically deficient and must attach to the nearest word in the phrase it governs. Enclitics may not receive primary lexical stress.[14] Clearly, the processes are closely related, but their differences at the morphology-phonology interface are important.

[13]Section 1.3 provides further commentary on this terminology.

[14]The concept of cliticization as discussed herein is influenced strongly by Anderson's (1992, 2005) discussions of "phonological clitics."

2.3.2. Compounds

Compounding is a common structure in LDN, found throughout the texts investigated in this study. The four examples below depict the syntactic phenomenon of adpositional compounding in LDN when the attached noun is monosyllabic.

(19) *mniáglagla*
mni-aglágla
water-across
'across [the] water' (Deloria 1932:68)

(20) *tʰimáhel*
tʰi-mahél
house-in
'in [a] house' (Rood & Taylor 1996:452)

(21) *cʰąʔákʰotąhą*
cʰą-akʰótąhą
woods-across
'across [the] woods' (Rood & Taylor 1996:452)

(22) *cʰąáglagla*
cʰą-aglágla
woods-along
'along [the] woods' (Deloria 1932:40).

Note that the primary stress of the resulting compound always falls on the adposition in these cases. In compounding, LDN's strong tendency to place primary stress on the second nucleus of a prosodic word appears to hold. Rood & Taylor (1996:452) explicitly describe the process of conjoining adpositions and the determiner phrases that they govern as "compounding"; however, instead of calling the resulting word a compound, they call them adverbs. This makes sense, as the constituent created typically describes a verb. This view implies that these constructions are not only single prosodic words, but also single morphological words, which I do not believe to be accurate. Thus, I refer to these as compounds throughout this paper.

(23) *pahá-ektà*
pahá-ektá
hill-at
'at [a] hill' (Ullrich 2018:136-137)

(24) *pahá-akàŋl* (provided as *pahá-akáŋl*)
pahá-akáŋl
hill-on
'on [a] hill' (Ullrich 2020b)[15]

(25) *wakpála-opʰàya*
wakpála-opʰáya
creek/stream-along
'along [a] stream' (Deloria 1932:19)

(26) *wakpála-aglàgla*
wakpála-aglágla
creek/stream-along
'along [a] stream' (Deloria 1932:146)

Adpositional compounding can also occur with polysyllabic nouns. In these cases, the primary stress falls on the governed term—not the adposition—unlike the examples with monosyllabic nouns (Boas & Deloria 1939:21). This is due to LDN's pervasive left-aligned iambic stress, as mentioned above. When the nominal constituent of the compound has more than one syllable, it will contain the stressed nucleus of the first iamb; this demotes the stressed syllable in the adpositional constituent of the compound to secondary stress. As evidenced by examples (23) and (24), individual scholars vary the notation in which they record the prosodic features of compounds from paper to paper. Ullrich marks both accents as primary and refers to the combining

[15]Example (24) was generously provided by Dr. Jan Ullrich in personal correspondence.

process as incorporation in example (24). However, the structure of this example is identical to the structure of the numerous examples in Ullrich (2018) and Deloria (1932), such as example (23), which leads me to posit that the postpositional accent is likely secondary.[16]

(27) *mní wą aglágla*
water INDEF across
'across a [body of] water' (Deloria 1932:74)

As evidenced by example (27), compounding does not occur when a determiner is used. More research needs to be conducted on the precise semantic variation in usage, but the current evidence points to speakers choosing which construction to use based on the importance of the [±definiteness] feature of the noun in a given utterance.

(28) *sič^{h}óla*
si-č^{h}óla
shoes-without
'barefoot' (Ullrich 2018:136-137)

(29) *hač^{h}óla*
ha-č^{h}óla
clothes-without
'naked' (Ullrich 2018:136-137)

(30) *míla č^{h}óla*
knife without
'without [a] knife' (Deloria 1932:124)

(31) *huŋská č^{h}óla*
leggings without
'without leggings' (LLC 2021)

The morphosyntactic usage of /č^{h}óla/ ('without') is almost identical to that of the adpositions in the compounds discussed above. However, some scholars suggest that /č^{h}óla/ is always bound, implying that examples (30) and (31) are ungrammatical (Ingham 2003:40). Because of the fact that /-čhóla/ is always primarily accented, it cannot be an enclitic.[17] Thus, this would force us to describe /č^{h}óla/ as a derivational suffix that derives adjectives from nouns while adding the semantic notion of 'without.' However, this is not the situation that the Deloria texts present.

As illustrated in the four examples above, /č^{h}óla/ appears to attach to a noun, forming a compound with it only when the adjoining noun is monosyllabic. In these cases, since the first nucleus of /č^{h}óla/ is the second syllable, it maintains its primary stress (Boas & Deloria 1939:21). Polysyllabic nouns, however, contain (minimally) a complete iamb; this would inhibit /č^{h}óla/ from simultaneously compounding with one *and* maintaining its primary stress. To avoid this, /č^{h}óla/ remains a free-standing prosodic word in these scenarios, with both the noun and /č^{h}óla/ maintaining their own full primary lexical stress. Thus, the difference between /č^{h}óla/ and the other compounds discussed herein is that there is a lexeme-specific rule that prevents /č^{h}óla/ from compounding with polysyllabic nouns.

(32) *holázatakiya*
ho-lazáta=kiya
tipi.circle-behind=towards
'towards the back of the tipi circle' (Deloria 1932:233)

[16]Note that the two examples from Deloria (1932) have identical glosses but use different postpositions; there are multiple prepositions meaning 'along' with only slight semantic differences.

[17]As evidenced by examples (30) and (31), there seems to be a lexical constraint on /č^{h}óla/ that forces the /ó/ to always carry primary stress. Further research is needed to determine whether there are other words with similar prosodic requirements.

Just as LDN's enclitics can attach to free postpositions, they can attach to the postpositional morpheme of a compound. Example (32) depicts the addition of an adpositional enclitic onto a N+ADP compound. Adpositional enclitics will be discussed further in section 2.3.3; for this example, only its status as a clitic is important. The meaning of this utterance, and many others like it, is compositional. This is much like the English preposition "into," but with even less semantic drift and fossilization.

2.3.3. Enclisis

Enclisis is not a common morphosyntactic realization of adpositions in LDN and in fact has yet to be described as such in the Siouanist literature.[18] However, there are at least two constructions in LDN in which enclisis does occur: /=kiya/ 'towards,' as already seen in example (32) , and /=ta/ LOC. This is a fertile area for further research, particularly if one has access to native consultants or archival recordings and can thus perform suprasegmental analysis.

(33) *iyúweȟtakiya*
iyúweȟta=kiya
opposite.shore=towards
'towards [the] opposite shore' (Deloria 1932:29)

(34) *enánakiya*
enána=kiya
here.and.there=towards
'towards various locations' (Deloria 1932:104)

(35) *holázatakiya*
ho-lazáta=kiya
tipi.circle-behind=towards
'towards the back of the tipi circle' (Deloria 1932:233)[19]

(36) *wic^{h}áša wą étkiya*
buffalo INDEF =towards
'towards some buffalo' (Deloria 1932:99)

The enclitic /=kiya/ functions as a canonical prosodic clitic (Anderson 1992).[20] It never appears as an independent prosodic word, nor does it ever carry stress after undergoing encliticization. This clitic appears to be a form of /etkiya/, an analogous free postposition also meaning 'towards' (Deloria 1932:30, 99). This is exemplified in example (36).

The morpheme /=ta/ is a versatile locative meaning 'to, on, or at.' There are at least two plausible explanations for the morphosyntactic behavior of /=ta/: enclisis and case-marking. Enclisis triggers a null determiner, just like many of the examples above, and is the simplest explanation.[21] An alternative explanation is that /-ta/ itself is in the D-head as a locative case marker. This analysis has not been posited by contemporary scholars of LDN, and Siouan languages are typically caseless. The late Regina Pustet (2000) briefly mentioned that adpositions could be developing into case markers in LDN, but she never expanded upon this theory. /-ta/ can be analyzed as a result of this process.

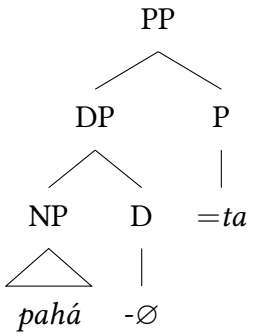
[18] Note that this is specifically *adpositional* enclisis. Other forms of enclisis have been discussed.
[19] Note that this example is repeated from section 2.3.2.
[20] "Canonical" in the context of the theoretical orientation of this paper, as discussed in section 1.3
[21] Ullrich (2020b) states that this null determiner results in semantically opaque definiteness.

(37) *tiŋtata*
tiŋta-ta
prairie-LOC
'on/at/to [a] prairie' (Riggs 1895:52)

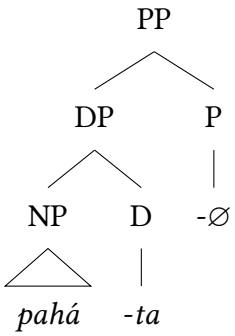
(38) *paháta*
pahá-ta
hill-LOC
'on/at/to [a] hill' (LLC 2021)

Examples (37) and (38) depict common instances of this morpheme. Both aforementioned morphosyntactic theories—enclisis and locative case-marking—are illustrated by examples (39) and (40), respectively.

(39) Enclisis



(40) Locative Case



Another possibility to consider is that /-ta/ could be in a transitory state between enclisis and case-marking. The lack of determiner usage with enclisis makes it difficult to differentiate the two structures syntactically.

2.4. Applicatives

As discussed briefly in section 1, LDN has three locative applicatives: the superessive ('above'), the inessive ('inside'), and the instrumental ('against' or 'by means of'). These are represented by /a-/, /o-/, and /i-/, respectively, and are shared by many Siouan languages (Ingham 2003:26-27).

(41) *ieekiya*
i-eekiya
INS-pray
'to pray for [something]' (Riggs 1895:53)[22]

(42) *ohnaka*
o-hnaka
INESS-place.something
'to place something into [something else]' (Riggs 1895:53)

(43) *amani*
a-mani
SUPERESS-walk
'to walk on [something]' (Riggs 1895:53)

While the instrumental and inessive applicatives tend to take inanimate patients, the superessive takes both animate and inanimate patients freely. Additionally, there are numerous

[22]The gloss 'for' appears to be an extension of "against."

examples of applicatives becoming fully fossilized within a verb. In many of these cases, semantic drift has obscured the semantic connection between the verb's contemporary meaning and the fossilized applicative's semantic content (Boas & Deloria 1939:42). This is the case in example (44) below.

(44) *ic^{h}ága*
*i- c^{h}ága
*INS- grow.INF
'to grow' (Boas & Deloria 1939:42)

The middle lines of example (44)'s gloss are misleading, though. This is because—due to the aforementioned fossilization—there is no longer a morpheme boundary where example (44) suggests. A more accurate version is presented in example (45).

(45) *ic^{h}ága*
grow.INF
'to grow' (Boas & Deloria 1939:42)

While LDN's applicatives almost certainly developed from postpositions (Helmbrecht 2006), they appear to have a broader semantic scope and more morphosyntactic versatility than their adpositional relatives. In some cases, the noun that the superessive puts in relationship with its attached verb is deep in the previous clause without apparent movement (Deloria 1932:48). LDN's various forms of adposition discussed above must be immediately adjacent to the phrase they govern.

(46) *c^{h}ąíyali*
c^{h}a-i-a-li
tree-INS-SUPERESS-climb/step
'to climb up against the tree' (Deloria 1932:117)

Example (46) illustrates that—when not separated by a determiner, free postposition, stative verb, or other sentential unit—verbs with attached applicatives can compound with the preceding noun. If /c^{h}a/ was not compounded with the applicativized verb, the stress would fall on the /a/ of (/iyáli/), not the /i/ (/íyali/), suggesting that this is true compounding and not a clerical error (Deloria 1932:117).

3. Discussion of Lakhota-Dakota-Nakota

LDN is implicitly (perhaps even subconsciously) the "de facto" language of reference among Siouanists (Rankin et al. 2003). The name of the entire language family—Siouan—comes from the exonym for LDN's speakers: the Sioux. LDN is one of the most thoroughly documented Siouan languages and has published grammars going back more than a century (Riggs 1895). As a result of this, preeminent scholars of LDN—such as Jan Ullrich, Bruce Ingham, Franz Boas, and David Rood, *inter alios*—often agree on the functions and descriptions of its basic morphosyntactic phenomena. For example, as discussed at length in section 2.3, almost all of the aforementioned scholars refer to any phenomenon relating to word-combining as 'incorporation.' Despite this

widespread scholarly agreement, the analysis in section 2 argues for the existence of compounding and enclisis as distinct morphosyntactic phenomena. Section 2 also postulates the existence of a locative case marker in LDN, another phenomenon yet to be seriously considered by contemporary scholars.

Morphologically independent postpositions are the most common form of adposition in LDN. These constituents govern a determiner phrase and are typically dominated by a verb phrase. The analysis above suggests that these assign a [+ambiguous definiteness] feature to the determiner phrase they govern, resulting in only rare uses of determiners. Bruce Ingham hints at this, as discussed in section 2.2; additionally, personal observation from the Deloria texts suggests adpositional co-occurrence with determiners is very uncommon (Deloria 1932). Example (7)—*maza oŋ* ('[made] of iron')—illustrates a simple postpositional phrase with a null determiner. When a postposition is morphologically free, this rule is violable, but usually still holds. A violation of this rule is outlined with 'next to the house' in (6a), where /kiŋ/ (DEF) appears in the surface structure. Additionally, morphologically independent postpositions in LDN can attach the same pronominal affixes that verbs take. Example (13) depicts a simple case of this phenomenon in which /hakab/ ('behind') is prefixed with /ni-/ (1SG.SBJ). Section 2.2.1 discusses more complex examples.

Adpositions can directly follow stative verbs, which function adjectivally in LDN. This is examined at length in example (11), where the postposition 'along' follows the stative verb 'to be large.' Research to date has only mentioned this construction and listed examples; the underlying syntactic structure has not been formally analyzed.[23] Example (11) evaluates multiple analyses, but the most likely structural motivation is that the postpositional phrase headed by 'along' takes a clausal complement, as delineated in example (11c).

Adpositions can be derived from adverbs in LDN simply by adding the prefix /i-/. This markedly productive construction is depicted in example (3), in which the adposition 'behind, after' is derived from the adverb 'afterwards' by attaching the prefix /i-/. Some adpositions have even developed a second form prefixed with /i-/ due to speakers reanalyzing the adpositional base as an adverb and subsequently adding the /i-/ prefix to ensure the word's adpositional morphosyntactic functions. This is illustrated in example (4), in which both /mahél/ and /imáhel/ mean 'inside.'

Compounding—under the term "incorporation," as discussed in sections 1.3 and 2.3.2—is a well-documented phenomenon in LDN. Despite this, the scholarly work I encountered all referred to the products of compounding as adverbs, not compounds. Only Ingham (2003) even refers to the process as compounding. Scholars' choice to not use more specific language was likely intentional, as it allowed them to present data without making an intentional claim about the morphosyntactic phenomena therein. In compounding, a noun and a postposition—each with their own underlying primary stress—are conjoined, creating a single prosodic word.[24] When an adposition is compounded with a monosyllabic noun, the primary stress is placed on the first syllable of the adposition, as illustrated in example (22). When compounding occurs with poly-

[23] Jan Ullrich has analyzed the underlying structure of stative verb phrases, but adpositions were not part of this analysis (Ullrich 2020a). Additionally, stative verbs directly following a determiner have—to my knowledge—not been analyzed.

[24] What I call a "prosodic word" and a "morphological word" here would be a type of phonological word and a grammatical word, respectively, in the typology of Dixon and Aikhenvald 2003. This is discussed in greater detail in section 1.3.

syllabic nouns, the primary stress falls on the second syllable of the noun, as shown in example (23). This patterning is due to LDN's pervasive left-aligned iambic stress. In compounds with monosyllabic nouns, the first iamb is split by a morpheme boundary; with polysyllabic nouns the entire iamb falls within the nominal constituent. Section 2.3.2 discusses this in greater detail.

The use of the term "incorporation" by scholars of LDN extends to their descriptions of enclisis, as well. Compounding and enclisis are distinct phenomena, a fact the term "incorporation" belies. Section 2.3.3 illustrates and delineates the discrete prosodic features that engender this distinction. It should be noted that the determiner phrases governed by adpositions in both compounding and enclisis cannot contain overt determiners; as mentioned above, this rule is only violable in the case of independent postpositions. Adpositional enclitics in LDN are phonological clitics[25]—not morphosyntactic clitics—under the theory of A-Morphous Morphology proposed by Anderson (1992, 2005). Thus, no intra-clitic syllable can receive primary stress, as the morpheme is prosodically deficient and attaches to the already-stressed noun that precedes it. Example (33) in section 2.3.3 exemplifies these properties with the enclitic /=kiya/ ('towards').

The enclitic /=ta/ (LOC) has plausibly become a locative case marker. Because enclisis disallows the presence of an overt constituent in the D-head, the syntax is ambiguous. The possibility of a locative case-marker is notable because it is not discussed in the major grammars of LDN. Moreover, Siouan languages generally do not have phonetically-realized morphological case markers. If /=ta/ is not yet a full case marker, it may be in a transitory state between this and an enclitic. More data is needed for further analysis.

LDN exhibits three locative applicatives—the superessive, the inessive, and the instrumental—which I consider "para-adpositional" phenomena. This is because they often provide information semantically similar to that provided by adpositions; moreover, this set of preverbs almost certainly developed from free postpositions. The usage of each aforementioned applicative is examined in section 2.4. The presence of these aligns with Helmbrecht's (2006:4) claim that these three types of locative applicative are found in all Siouan languages. The declining productivity and increasing semantic ambiguity of these applicatives support the theory that they are progressing towards fossilization, as suggested in Helmbrecht & Lehmann's (2008:34-35) diachronic hypothesis (discussed in section 8). The clear distinction between applicatives and adpositions morphosyntactically also supports Helmbrecht & Lehmann's (2008:34-35) implication that these phenomena ought to be treated separately in synchronic analyses. However, LDN's support of these claims is not entirely surprising, given that Helmbrecht would likely be more familiar with LDN and have more access to data from LDN than any other non-Hocąk Siouan language.[26]

If the Siouan language with the most significant history of linguistic work and documentation has significant gaps in the analysis of its adpositional morphosyntax, then it is plausible that a Siouan language studied by only a small handful of scholars over the past century would, as well. This prompts the examination of Catawba's (para-)adpositional morphosyntax, which is explored in sections 4 and 5, below.

[25]The nomenclature for these used throughout this paper is "prosodic clitics."

[26]Hocąk is Helmbrecht's primary language of interest in the Siouan family. This statement does not reflect a general abundance of Hocąk data.

4. Evidence from Catawba

4.1. Overview

The morphosyntactic status of adpositions in Catawba is markedly distinct from that of the so-called "core" Siouan languages (the Western branch). This is unsurprising, given its early split from the group (see figure 1). In Catawba, the attachment of adpositional proclitics onto verbs is by far the most robust form of adposition-marking. Free postpositions, while they do occur in the extant corpus, are relatively rare. The adnominal enclisis of adpositions does occur, but this is far less common than the existing transcriptions suggest and is not appreciably productive. Furthermore, this investigation revealed several data in which Catawba makes use of post-verbal adpositions, both as verbal enclitics and as free prosodic words.[27]

4.2. Proclisis

As stated above, the primary method of adpositional marking in Catawba is the attachment of postpositional proclitics onto the verb that dominates them. The following examples illustrate the typical usage of these proclitics.

(47) *huktúkəre*
huk=tuk-re
down=fall.down-IND
'[it] falls down' (Speck 1934:2)

(48) *duhotiiriie*
duk=ho-tiiriie
back=come-NARR
'[it] came back' (Speck 1913:323)

(49) *hukáii*
huk=káii
down=throw
'throw [it] down' (Speck 1913:324, Rudes 2007:34-35)

(50) *dugdánire*
duk=ra-ni-re
back=go-1SG.OBJ-IND
'back to me' (Speck 1934:3, Rudes 2007:44)

As depicted in examples (47) through (50), Catawba's adpositional proclitics attach rightward, onto the left end of a verb. This is often accompanied by phonological changes, which the academic literature on Catawba has thus far neglected.[28] Example (47), for instance, shows that the indicative suffix /-re/ requires a preceding vowel. When a vowel does not precede it

[27] While outside the scope of this paper, this research potentially revealed a morphosyntactic phenomenon yet to be documented in Catawba: switch-reference marking via the suffix /-uk/ (sometimes realized as /-ik/ or /-ək/ due to u-i variation and reduction, respectively). Apart from Rudes (2007), no analysis of Catawba has investigated this morpheme. Rudes claims that it is a resultative marker and also states that this morpheme accounts for the word-final /k/ in /únik^{h}/ (Rudes 2007:77-78). While the extant data do not refute his analysis, I maintain that this could be switch-reference and look forward to researching it further. It should also be noted that this is a plausible cognate for Mandan's different-subject switch-reference marker, /-ak/, which would provide additional morphological evidence for the new computationally-modeled phylogeny of the Siouan languages developed by Kasak (2019:313-314, 2020a).

[28] This is common in Catawba, as it has only been studied by a handful of scholars, most of whom worked on the Catawba lexicon. The most recent of these are Kathleen Shea and the late Blair Rudes, both of whom worked on Catawba in the late twentieth century.

in the underlying structure, a schwa is epenthisized, resulting in /-əre/.[29] In example (48), the morpheme-final /k/ is syncopated. Example (49) illustrates a pervasive phonological process in Catawba: geminate deletion. We see voice assimilation across a morpheme boundary in example (50), in which morpheme-final /k/ becomes /g/, acquiring the [+voice] feature of the following morpheme-initial /d/.[30]

(51) *yapawą́mǫhere*
yəpə=wą́-mǫ-h-re
up.and.down=jump-sing-3.SBJ-IND
'Jump up and down [while] singing'
(Speck 1934:9, Voorhis n.d.122,124, Shea 1984:336)

The example above adheres to the same rules as the previous four examples. The only difference is that in this case, the postposition procliticizes onto a serial verb.[31]

(52) *ntúgbakóre*[32]
n=tuk=buk-re
then=inside=put-IND
'then put [it] in' (Speck 1934:8, Shea 1984:303)

The attachment of the proclitic /n=/ ('then') onto the adpositional proclitic /tug=/ illustrates the grammaticality of clitic-stacking in Catawba. Speck marks not only primary, but also secondary and tertiary stress with an acute accent, so the presence of /ú/ instead of /ù/ is not particularly concerning evidence against proclisis, as we do not know the syllable on which Speck heard the primary accent. Moreover, I believe Speck has a tendency to confuse lexical stress with prosodic emphasis, which is common for native English-speakers. This is expanded upon throughout section 4.

(53) *búrukyáamuhiiwą́hahe*[33]
buruk=yaamu=hii-wą-ha-he
back.again=into.water=-3SG.SBJ-jump-INCEP-CONT
'Back into [the] water he jumped' (Speck 1913:323, Rudes 2007:18-19, 71-72)

(54) *mǫhuktuikəre*
mǫ=huk=tuk-re
in=on=fall.down-IND
'fall onto...' (Speck 1934:1)

[29]Kasak (2020b) suggests that this may be an instance of Dorsey's Law, as many Siouan languages are subject to this (Dorsey 1885).

[30]These are merely cursory phonological observations that are evidenced by the data directly pertaining to adpositions. I hope to examine this further in the future.

[31]Serial verbs are a common structure in Catawba and appear to be semantically transparent, corresponding roughly to a coordinated verb pair in English ('He *eats and drives* at the same time.').

[32]This is how Speck transcribed this word. I believe the transcription is more likely /ntùgbakóre/.

[33]As with example (52), we do not know which marked stress is primary. My argument suggests it is on /wą́/. Furthermore, one reviewer suggests that The *mu* in *yaamu* is really a locative adposition, where *yaa* is 'water.' Furthermore, this reviewer also suggests that the *hii* is not a subject marker, but the determiner *hii* 'yonder,' given that *wą* generally takes subject suffixes instead.

One notable morphosyntactic feature in example (53) is prefixal subject-marking. Catawba has full, productive systems of both prefixal and suffixal person-marking, as outlined in Rudes (2007).[34] Example (53) does not differ significantly in structure from example (52); however, in this case, it is a second postposition being procliticized onto the postposition closest to the verb. The semantics of this construction are straightforwardly compositional. As in the previous example, the placement of an acute accent mark within both proclitics is not problematic, as Speck did not distinguish stress tiers and these are most likely secondary and tertiary stress.

It could be argued that /búruk/ in example (53) is an independent prosodic word; /búruk/ is irregular in that its free form and proclitic form only differ prosodically (/búruk/ has primary stress; /buruk=/ does not). Moreover, the first /u/ is where we would expect the stress to fall in its free form (Rudes 2007:18-19). However, as evidenced by example (54), even if this /búruk/ is a free-standing postposition, it does not change the fact that stacked proclisis is grammatical in Catawba.

4.3. Free Postpositions

In addition to postpositions being able to procliticize onto verbs in Catawba, they can also appear as free prosodic words. Free postpositions in Catawba appear to assign a [+ambiguous definiteness] feature to the preceding noun. The result of this in the surface structure is a null determiner head; however, as evidenced by example (57), this rule is violable. Catawba's free postpositions typically contain their corresponding proclitic form along with an additional syllable. Rudes argues that this extra syllable is underlyingly /-ya/, /-yi/, or /-ku/ and calls these morphemes "adverbializer" suffixes, despite identifying the words they create as free postpositions (Rudes 2007:18-19). My analysis does not support this theory. Only a small number of Catawba's free postpositions end in morphemes that are probably derived from /-ya/, /-yi/, or /-ku/. Some, such as /hitak/ in example (55)—whose proclitic form is /tak=/—even have the extra syllable on the left. Others, like /buruk/ (as discussed in example (53)), do not add a syllable at all. Of the four examples below, none appear to have morphemes derived from /-ya/, /-yi/, or /-ku/. However, further diachronically-focused research is necessary to determine the morphemic status of the additional syllables in these free adpositional forms.

(55) *iswą hitak*[35]
river down
'down [a] river' (Speck 1934:36)

(56) *súk hapáng*[36]
house above
'above [a] house' (Gatschet 1900:533)

(57) *yątci kį sukhǫ́*[37] *wąre*
yątci kį sukhǫ́ wą-re
stream the over sit-IND
'[It] sits over the stream.' (Speck 1934:10)

(58) *yancámǫntu*
yancá#mǫ́ntu
creek#in
'in [a] creek' (Speck 1934:3, Shea 1984:301)

[34]This is not an uncommon feature among Siouan languages. Crow has two pronominal paradigms (Graczyk 2007:60). While working with Dr. Marcia Haag and Dr. Dylan Herrick on their Osage (Siouan, Dhegihan) fieldwork, we encountered double subject-marking, with some speakers using both paradigms simultaneously.

[35]/hitak/ corresponds to the proclitic /tak=/

[36]Both /hapáng/ and /hápki/ correspond to the proclitic /hap=/

[37]/sukhǫ́/ corresponds to the proclitic /suk=/

Though Speck writes the above as if /mǫntu/ is an enclitic attached to /yancá/, I believe these are separated by a word boundary. This is because /mǫtu/ is the free form of /mǫ=/, the proclitic for 'in.'[38] However, as with many phenomena in Catawba, the lack of audio data inhibits unequivocal descriptions.

(59) Enclisis, Morphological Independence, or Proclisis?

a. *sakhapkii*
sák hápki
hill up
'up [a] hill' (Speck 1913:322)

b. *sák hápki*
hill up
'up [a] hill' (Speck 1934:84)

c. *hápkiiwá*
hápki wą́
above sit
'[to] sit above' (Speck 1913:323, Voorhis n.d.112, Rudes 2007:18-19)

The three examples above were all recorded by Speck. However, in example (59), /=hapkii/ is written as an enclitic; in example (59b), it is written as its own morphological word; and in example (59c), it is recorded as a proclitic. By my analysis, the postposition is prosodically independent in all three instances. As mentioned above, /hápki/ is the free form of the proclitic /hap=/. I believe this variation in transcription is due to monosyllabic words not receiving strong primary stress in casual speech. Moreover, it is easy for English speakers to confuse prosodic emphasis with stress, as both involve similar suprasegmental features. This is what most likely caused the lack of consistency in Speck's transcription. However, this is simply a general characterization based on my research and more data is needed to draw definitive conclusions.

4.4. Enclisis and Complex Incorporation

4.4.1. Enclisis

Despite the adnominal attachment of adpositions being recorded frequently in Speck's (1934)) transcriptions, my analysis suggests that enclisis was not a productive morphosyntactic process at the time of his work on Catawba. Many apparent examples of enclisis recorded by Speck have nearly identical corresponding examples in which the adpositional form is free. This was illustrated in examples (59) in the previous section.

(60) Enclisis or Free Postposition?

a. *íiswąhiiák*
íiswą hiiák
river over
'over [a] river'
(Speck 1913:329, Shea 1984:173)

b. *íiswą hiiák*
river over
'over [a] river'
(Speck 1934:91, Shea 1984:173)

[38] /mǫntu/ is more commonly written as /mǫtu/.

Example (60) also depicts this transcriptional inconsistency. Despite this, there does appear to be one clear example of enclisis; however, I believe this is a fossilized form, not a productive enclitic.

(61) *íiswątak*
íiswą=tak
river=down
'down [a] river' (Speck 1934:1, 14, 15, 39, 72)

In section 4.3, we saw that /hitak/ was the long form of the proclitic /tak=/. This is the same clitic morpheme, but used enclitically as /=tak/. The word /íiswątak/ occurs often in the stories documented by Speck. While the indices by Voorhis (1992, n.d.) and Shea (1984) include the word, its usage has not been analyzed contextually in Siouanist literature.[39] My impression is that /íiswątak/ potentially underwent a mild semantic bleaching process, resulting in it being used to convey "elsewhere, not here, over there." Another analysis could be polysemy; perhaps "across/down [the/a] river" has been metaphorically extended to mean "somewhere other than here" or "not in this immediate vicinity," creating a polyseme. Further research is necessary in order to make stronger claims.

For both of these theories, /íiswątak/ appears to have undergone fossilization—which in this case is pseudo-adverbialization—before enclisis became ungrammatical in Catawba. Fossilization would have deleted the morpheme boundary between /íiswą/ and /=tak/, so native speakers would not have found this construction ungrammatical despite the ungrammaticality of enclisis in Catawba.[40] Even if none of the aforementioned hypotheses reflect reality, /íiswątak/ still appears to be the only consistent example of enclisis in the extant Catawba data. This suggests that enclisis was once a grammatical morphosyntactic feature, but that it is no longer productive.

4.4.2. Complex Incorporation

Multiple times throughout the texts Speck transcribed, he writes N+ADP+V combinations as a single word.

(62) One, Two, or Three Prosodic Words? (N+ADP+V, N+ADD=V, or N ADP V?)

a. *yapháp dáre*
yap hàp=dá-re
tree up=go-IND
'go up [a] tree' (Speck 1934:16)

b. *yap háp cáre*
yap hàp= cá-re
tree up= climb-IND
'climb up [a] tree' (Speck 1934:16)

The examples in set (62) differ by only one morpheme, resulting in their glosses differing by only one lexeme: 'go' vs. 'climb.' The semantic senses of the utterances are quite similar. Moreover, these phrases were recorded within the *same story*. However, example (62a) is written

[39] To my knowledge, there has not been any semantic study done on any word in Catawba. I say this to reiterate that this is not a failing of any previous scholar; it is simply representative of the dearth of material.

[40] This could be compared to the transformation in English from "down (the/a) stream" to "downstream," though the polysemization theory goes one step further.

as a single word, while example (62b) is written as three separate words. I am not convinced either parsing is correct.

(63) *katukéhəre*
ka tuk=ké-h-əre
hole inside=put-3SG.SBJ-IND
'Hole in put' (Speck 1934:15, Voorhis n.d.118)

(64) *yaphapkǫere*
yap hap=kǫ-ere
tree up=go-IND
'go up [to a] tree' (Speck 1934:7)

Examples (62a) and (62b) are not the only examples of this. The two examples above are morphosyntactically identical. Thus, I have parsed all four in the same manner. I analyzed all four examples as consisting of a morphologically and prosodically free noun followed by a procliticized postposition + verb unit, resulting in two prosodic words. As mentioned in section 4.3 and elsewhere, I believe Speck tends to transcribe compounds when the primary stress on monosyllabic words is not particularly strong, resulting in the confusion of lexical stress with the suprasegmental effects of emphasis. However, this is a general characterization based on my review of printed material; to conclude with more confidence, one would need access to prosodic data.

(65) *hícəpąhúkcę́həkʼ*
hícəpą huk=cę́-h-uk
slobber down=pull-3SG.SBJ-SW.REF.DIFF.SBJ
'[His] slobber fell down...' (Speck 1913:323, Shea 1984:266)

It is unclear why Speck transcribed a single prosodic word for the utterance glossed in example (65). In his footnote, he transcribes 'slobber' as /hícəpą́/, with a stress on both the /i/ and /ą/. Note that the /ą/ is unmarked. It is plausible that Speck expected to hear a stress on this /ą/ and did not, and thus believed it to be compounded onto the verb phrase. As already mentioned, Speck may tend to confuse phrasal emphasis and lexical stress. Because the suprasegmental effects of phrasal emphasis likely would have affected /cę́/, it would not be surprising if this were an example of that confusion. Again, however, this is solely conjecture based on intuitions from researching the corpus of extant Catawba data; unfortunately, no definitive conclusions can be drawn without access to recordings.

(66) *hapáawǫ́həʼdúgrehatíiriie*
Hapa=wą-hə duk=re-ha-tiiriie
out.on.the.bank=jump-INCEP back=look-3SG.SBJ-NARR
'He jumped out onto the bank, looked behind…' (Speck 1913:323, 326)

(67) *hukáiiʔhagwarúphə*
huk=káiiʔ hagda+warúp-hə
down=throw pick.up+grab-3SG.SBJ
'…throws [it] down, grabs [it]' (Speck 1913:324)

Examples (66) and (67) are particularly noteworthy, as Speck's transcriptions suggest that two full verb phrases are compounded together. Serial verbs appear to be quite common in

Catawba, as discussed in section 4.2, but these would be the only examples of two *verb phrases* combining. Consequently, my analysis does not align with Speck's. As delineated in example (66), I consider there to be two independent prosodic words, each consisting of a postposition procliticized onto a verb. The postposition /duk=/ in the prosodic word /dugrehatiiriie/ in (66) lacks a governed noun, suggesting one of two phenomena. This is most likely an example of NP-dropping. Siouan languages have a strong tendency to drop lexical information that has already been introduced into the discourse (Kasak 2020b). This seems to be evidence that Catawba does the same, as it is clear from context that the subject is looking behind *himself*. Another possibility is that /dugre/ has undergone a degree of fossilization, similar to particle verbs in English. In this case, it would not necessarily require a governed term. As with numerous examples already discussed in section 4, I presume the suprasegmental effects of phrasal emphasis to be the source of Speck's unexpected transcriptions here. Note that the actions of these verbs are occurring simultaneously (or, if not, practically so). In the first example, both are marked with the inceptive (INCEP) aspect. It makes sense that these verb phrases would share a single phrasal point of emphasis.

4.5. Post-Verbal Adpositions

There are two instances in the extant data which contain an adposition following a verb rather than preceding it. The first appears to serve a clear semantic purpose. The second, however, is far less transparent.

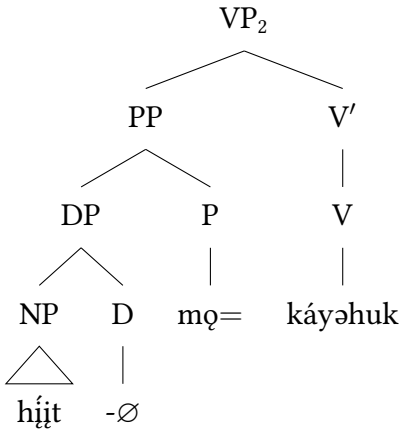
(68) *hapkáii?iitíiriie* *hápkii*
hap=kai-?ii-tíiriie hapkii
up=throw-3PL.SBJ-NARR up
'Up they put him, way on top' (Speck 1913:323)

In example (68), the addition of the independent postposition /hápkii/ after the verb serves to reiterate and emphasize the spatial relations between the patient and their environment (in this case, between an opossum and a scaffold). Note that the verb phrase already contains /hap=/, the proclitic form of /hápki/. This could also be an effect of register, as emphatic devices like repetition are common in storytelling.

(69) *káyəhuk* *hi̱tmǫtúkhətíiriire*
káyi?-h-uk hí̱i̱t mǫ=tuk-h-atiiriire
throw-3SG.SBJ-SW.REF.DIFF.SBJ face in=fall.down-3SG.SBJ-NARR
'He threw [it] in [his] face [and] he[dif] fell down.'
(Speck 1934:322, 326; Shea 1984:229, 292)

Example (69) is much more grammatically complex. One would expect the noun /hí̱i̱t/ and the postposition /mǫ=/ to precede the verb /káyəhuk/, as the semantic output implies they are dominated by it syntactically. It is clear that the surface structure of example (69) does not match the underlying structure because of the head-final nature of Catawba's syntax. The patient did not "fall down in [his] face," the agent "threw [it] in [his] face." Thus, we would expect the structure illustrated in example (70).

(70) Syntactic diagram (VP_2)



However, this is not the surface structure. The sentence undergoes some process in which the postpositional phrase is dislocated to the right side of the verb phrase that dominates it, deviating from Catawba's regular syntactic structure. This is yet another phenomenon that requires further research; unfortunately, there may not be enough extant data to conclusively answer this question.[41]

4.6. Relationship to Applicatives

Locative applicatives are a hallmark of Siouan morphosyntax, but Catawba appears to have either never developed them or to have developed them and subsequently lost them. Two of the applicatives found in other Siouan languages (for example, those described in section 2.4) have possible cognates in Catawba. The proclitic /sak=/ ('above') is plausibly a cognate of the superessive applicative (typically /a(a)-/), and /mǫ=/ ('in') is potentially cognate with the inessive applicative (typically /o(o)-/). Despite these potential etymological relationships, the extant Catawba data strongly suggests that Catawba did not have applicatives at the time of Speck's transcriptions.

5. Discussion of Catawba

The evidence presented above differs significantly from the LDN data delineated in section 2. Proclisis of a postposition onto the verb that dominates it is Catawba's predominant adpositional construction. However, this is not the only morphosyntactic locus in which adpositions appear. Independent, free-standing postpositions are grammatical in Catawba, as well. While Speck's (1934) transcriptions seem to suggest the presence of enclisis, bidirectional compounding, and phrasal compounding, I do not believe any of these phenomena are truly manifest in these texts. Additionally, the applicatives (preverbs) discussed in both section 1.3 and section 2 are absent from Catawba, categorically rejecting Helmbrecht's (2006:4) claim that the three locative applicatives are present in "all Siouan Languages."

[41] One possibility is that /hį́įt mǫ/ could be an extra-syntactic parenthetical. However, given the presently accessible data, this is impossible to prove. Moreover, parentheticals are a markedly controversial phenomena which—to my knowledge—have not been discussed in the Siouanist literature.

Proclisis is overwhelmingly the preferred adpositional construction in Catawba. This is outlined and exemplified in section 4.2. In example (48), for instance, the proclitic /duk=/ (‘back’) is attached to the verb /ho/ (‘to come’). In addition to standard ADV=V proclisis, adpositions in Catawba can procliticize onto serial verbs, as illustrated in example (51). Non-adpositional proclitics can undergo proclisis onto adpositional proclitics, as shown in example (52) where /n=/ (‘then’) is procliticized onto /tuk=/ (‘inside’). Moreover, stacked adpositional proclisis—the phenomenon in which one adposition undergoes proclisis onto another adposition that is already procliticized onto a verb—is grammatical in Catawba, as well. This is illustrated in example (53), in which /buruk=/ (‘back.again’) is procliticized onto /yaamu=/ (‘into.water’) and example (54), in which /mǫ=/ (‘in’) is procliticized onto /huk=/ (‘down’). Example (53) also demonstrates that adpositional proclisis can occur onto verbs with prefixal person marking, as we see /yaamu=/ attach to /hii-/ (3SG.SBJ). This illustrates the marked productivity of adpositional proclisis in Catawba.

Free postpositions, though not as common as postpositional proclitics, are also grammatical in Catawba. The use of a free postposition assigns a [+ambiguous definiteness] feature to the preceding noun, resulting in a null determiner head. This is evidenced by the habitual absence of Catawba’s determiners /kį/ (DEF) and /hį/ (INDEF) when a DP is governed by a postposition. However, this rule is violable, as shown in example (57), in which /kį/ (DEF) appears in the surface structure. When free postpositions occur, they are almost invariably one syllable longer than their corresponding proclitic form. Contra Rudes (2007), these “added” morphemes vary significantly from word to word, and it is highly unlikely that these morphemes are all adverbializers. Moreover, processes such as grammaticalization rely on the fact that as time goes on, free constituents often become bound. It would be rather surprising—though certainly not impossible—for a proclitic to take a suffix and become a free form.

On the other hand, the correspondence of both /hapang/ and /hapki/ to /hap=/ complicates this (see section 4.2). However, neither Voorhis (n.d.) nor Voorhis (1992)—two of the three best indices of the Catawba lexicon—include the word /hapang/. Moreover, Shea (1984:132)—the third member of that set—only includes it in regard to the exact sentence from Gatschet’s (1900) grammatical sketch that I cite in section 4.3, opting not to include it as its own word in her lexicon section. I have not encountered this word anywhere in the Speck (1934) texts, and it seems that neither Voorhis nor Shea did, either. This singular mention of /hapang/ appears to be the only extant evidence of its existence. Thus, the fact that both /hapang/ and /hapki/ correspond to /hap=/—though notable—is not well-attested.

The enclisis of adpositions onto the noun they govern is recorded frequently by Speck, and no subsequent scholar of Catawba appears to have questioned this. However, I do not believe enclisis to be a productive nor a common process in Catawba. Speck’s examples of enclisis consistently use the free form of a postposition (as in example (60)) or consist of a noun *and* a verb with a procliticized adposition (as in example (62)). There is only one example that appears to be true enclisis—example (61), *íiswą=tak* (‘river =down’)—but this appears to be a fossilized form and thus is not indicative of productive enclisis (see section 4.4.1 for further details). Regarding Speck’s numerous examples of N+ADP+V compounding and his occasional example of ADP+V+ADP+V compounding, I do not believe any to be parsed accurately. In these cases (as discussed at length in section 4.4.2), there are likely multiple independent prosodic words, as there should be a word boundary before the adposition(s). The “single” primary stress that Speck recorded was likely the locus of prosodic emphasis, not lexical stress. This process is illustrated

in example (66) and the ensuing discussion.

Although extremely rare, there are two notable cases of adpositions occurring post-verbally in the Speck (1934) texts. The motivation for this in the first example is emphasis, which I believe to be an extra-syntactic storytelling device in this instance. In example (68), the adposition /hap=/ (‘up’) is procliticized onto the verb, then the free form /hapki/ (‘up’) directly follows the verb. Note that the verb is marked with the NARR declension, supporting the storytelling theory. In the second post-verbal adposition, example (69), the constituents of a verb phrase (V′ and a postpositional phrase) switch places. There is only one example of this in the extant Catawba texts, so no definitive conclusions can be drawn. This is likely an example of right-dislocation.

This analysis demonstrates that Catawba's syntax differs markedly from other Siouan languages, such as LDN. However, this is not particularly surprising, as it explains why Siouanists tend to treat the Eastern (Catawban) branch as an inconsequential outlier. Thus, to have a more nuanced conversation about Siouan adpositions, one must explore another “core” (Western) Siouan language beyond LDN. Section 6 provides this with Crow.

6. Evidence from Crow

6.1. Overview

The adpositional morphosyntax of Crow exhibits marked variation and flexibility. The GOAL postposition in Crow is illustrative of this fact, as it can take all of the following forms: /-ss-/, /-ssee/, /kuss-/, and /kusseé/ (Graczyk 1989:8). Note that the first example, /-ss-/, is morphologically anchored[42] on both sides, the second and third examples are each anchored in a single direction, and the fourth example is a free postposition. Moreover, Crow has cognates to the LDN superessive, inessive, and instrumental applicatives discussed in sections 1.1 and 2.4, but the boundary between these applicatives and Crow's postpositions is rather inchoate. Because of this lack of clarity, it is most elucidative to begin the discussion of Crow's (para-)adpositional morphosyntax with its applicatives.

6.2. Applicatives

The most thorough grammar of Crow to date spends merely four short paragraphs on applicatives, which it labels “locative prefixes” (Graczyk 2007:88-89). These are /a(a)-/ (SUPERESS), /o(o)-/ (INESS), and /i(i)-/ (‘against’). Note that the applicative /i(i)-/ does not have an INS meaning. Graczyk claims that these constructions are the result of postpositions incorporating with the verb that dominates them and subsequently fossilizing, which is generally consistent with Helmbrecht and Lehmann's theory (Graczyk 2007:88-89, Helmbrecht & Lehmann 2008). The data evidencing Crow's applicatives are overwhelmingly examples of fossilization. Graczyk (2007:89) notes that many examples have no modern trace of a locative meaning. Consider the following examples.

[42] I use “anchored” here to mean morphologically bound. This avoids confusion with syntactic binding when both concepts are discussed simultaneously.

(71) *óolichi*
'to envy' (Graczyk 2007:91)

(72) *íkuchki*
'to plan' (Graczyk 2007:90)

(73) *áachiwi*
'to climb' (Graczyk 2007:89)

Note that examples (71), (72), and (73) are not parsed morphemically, as there is no longer a morpheme boundary between the applicative and the following verb. However, this does not illustrate the entire picture.

(74) a. *shuá*
shuá
spit (V)
'to spit' (Graczyk 2007:89)

b. *áasshua*
áa-shua
SUPERESS-spit
'to spit on [smth]' (Graczyk 2007:89)

As depicted in example (74), applicatives are not always fossilized. This is illustrative of variability within Crow's (para-)adpositional morphosyntax.

6.3. Compounding

Adpositions in Crow are frequently compounded with both nominal and verbal elements. The extant data exhibit numerous examples of all three variations of compounding that I propose: left-anchored adpositions (discussed in section 6.3.1), right-anchored adpositions (discussed in section 6.3.2), and bidirectionally anchored adpositions (discussed in section 6.3.3). This is yet another example of the fluidity in Crow's (para-)adpositional morphosyntax.

6.3.1. Left-Anchored

Many adpositions in Crow can be compounded leftward, onto the noun they govern rather than the verb that dominates them. Consider the following example.

(75) *hilaakée*
hili-aakee
this-SUPERESS
'now' (Graczyk 2007:71,110,368)

Example (75), above, depicts the leftward compounding of a form related to the superessive applicative (discussed in examples (74) and (73) in section 6.2).

(76) *hilíssee*
hilí-ssee
this-GOAL
'towards this' (Graczyk 2007:80)

(77) *éekhkoon*
éekhkoo-n
that-LOC
'in/on there' (Graczyk 2007:81)

(78) *hilihtée*
hili-htée
this-SPEC.LOC
'in/on right here' (Graczyk 2007:82)

(79) *baleeaak*
balee-aak
1PL-COM
'with us' (Graczyk 2007:388)

(80) *Bill binnáasketaa díilik*
Bill bin+náask-etaa díili-k
Bill water+bank-along walk-DECL
'Bill was walking along the shore.' (Graczyk 1989:2)

As illustrated by examples (76) through (80), leftward compounding occurs in a wide variety of Crow's postpositions. These examples further evince the grammaticality of left-anchored adpositional compounds in Crow. Example (80) is slightly more intricate than the other cases in that the postpositional compounding occurs onto a compound instead of a monomorphemic word. However, the morphosyntactic mechanism does not differ from the others.

6.3.2. Right-Anchored

In rightward compounding, a postposition attaches onto the verb that dominates it (while still forming a postpositional phrase with the DP it governs).

(81) *áakeela*
aakee-la
SUPERESS-be.at
'be on top' (Graczyk 2007:186)

Example (81) depicts the superessive /aakee/ undergoing right-anchored compounding. Recall that in example (75), the superessive underwent left-anchored compounding; in example (74), the it was a productive and semantically overt applicative; and in example (73), it was a semantically null, fossilized former-applicative. This is yet another piece of evidence that Crow's (para-)adpositional morphosyntax is incredibly flexible and it is difficult to demarcate boundaries therein.

(82) *aashúua iihúppiiliawaak*
aashúua ii-húppii-lia-waa-k
its.head INS-soup-make-1A-DECL
'I will make soup with its head.' (Graczyk 2007:386)

The /ii-/ in example (82) *is* the instrumental (INS) form, unlike in the applicative section, above, in which is exclusively had the locative meaning 'against.' The word /húppii/ ('soup') is a regular noun that has been incorporated into the verb /lia/ to form a verb meaning roughly 'to soup-make' (Graczyk n.d.287). Graczyk (2007:386) states that this is the free postposition /ii/ incorporating into this already-incorporated verb. However, if this were incorporation, /ii-/ would simply be the instrumental applicative, which is not attested in Crow. This leaves two plausible possibilities: this is an example of right-anchored compounding *or* the /i(i)-/ applicative represents both 'against' and INS. I believe this phenomenon to be the former, and that appears to be Graczyk's intended description; however, the latter possibility cannot be ignored.

(83) *baakáateesh aakhawassdáawaatak*
baakáatee-sh aak-hawass-dáaw-aat-ak
child-DET COM-around-travel-APPROX-SW.REF.SAME.SBJ
'...travel around with this child.' (Graczyk 2007:388)

In example (83), the comitative postposition is compounded to the right. Notably, it is compounded onto another postposition: /hawass/ ('around'). However, the latter appears to be strongly collocated with the verb, forming an approximate semantic equivalent of the English particle verb 'to travel around.'[43]

6.3.3. Bidirectional

Another construction exhibited by Crow is adpositions that undergo both the process discussed in section 6.3.1 and that of section 6.3.2, resulting in bidirectionally anchored postpositions.[44]

(84) *baáhpe héelahkeetaawasaailuk*
baáhpe héelahkee-taa-wasaa-i-lu-k
rock side-PATH-run-HAB-PL-DECL
'They run alongside the rocks.' (Graczyk 2007:384)

Example (84) illustrates the adposition /taa/ (PATH) not only attached rightward onto the verb dominating it, but also leftward onto the noun it governs. Additional examples of bidirectionally anchored adpositional compounding appear in sections 6.5 and 6.6, below.

6.4. Free Postpositions

In addition to the aforementioned varieties of compounding, Crow exhibits postpositions that constitute their own prosodic words. In prototypical instances, free postpositions follow the determiner phrase they govern and immediately precede the verb that dominates the adpositional phrase.

(85) *iseé ii*
his.arrow INS
'with his arrow' (Kasak 2019:195)

(86) *binnaxché kusseé*
fence GOAL
'towards [a] fence' (Graczyk 1989:81)

(87) *bishée áappaa déek*
bishée áappaa dée-k
buffalo COM go-DECL
'...goes with buffalo.' (Graczyk 2007:362)

(88) *amníam biaxsée bilé*
amnía-m biaxsée bilé
bank-DET under water
'...water under that bank...'
(Graczyk 2007:362)

[43]This is simply an observation from the data I encountered during this research. Further specific study is needed to thoroughly examine the lexical status.

[44]This would be considered by many to be incorporation. Among this group are many Siouanists and Crow specialists, including Graczyk (2007). However, some researchers dispute this claim (Gebhardt 2019). The long words in which incorporation is assumed may in fact represent a single pitch accent, not a single lexical accent.

(89) *baattáche aák*
rawhide com
'with rawhide' (Graczyk 2007:282)

Examples (85) through (89) provide a representative overview of Crow's free postpositions. Examples (85) and (86) are minimal examples in which a postposition is preceded by the noun it governs. In the example (87), this is expanded by adding the verb that dominates the postpositional phrase. Example (88) is a slight modification of this in which the dominating phrase is a DP instead of a VP. The final example, (89), depicts the comitative /aak/ as a free postposition; note that this is the same sense as the postposition in example (87), but that takes in a different form.

(90) *áakeen*
áakee=n
SUPPERESS=LOC
'on top of [smth]' (Graczyk 2007:46)

There is also an adpositional enclitic, /=n/, that can attach to adpositions that otherwise must be bound, allowing them to appear as free postpositions. This is illustrated in example (90) with the superessive.

6.5. "Missing" DPs

In several examples of utterances containing adpositions, the literature on Crow refers to the governed term of the postposition as "missing," wherein the listener must imply it from context (Graczyk 1989, 2007). While inference is certainly plausible, I argue that the governed term is not missing at all. Rather, the morpheme in question is a derived noun that has undergone conversion from an adposition without segmental changes.[45] Consider the examples below.

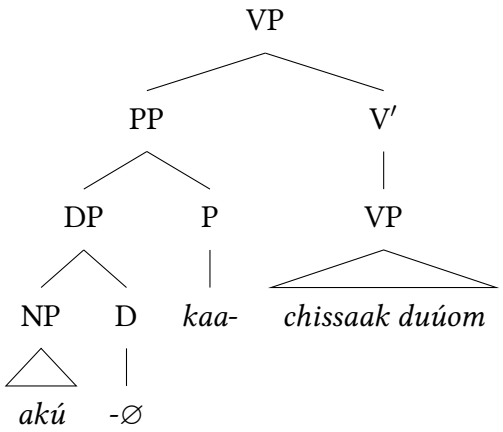
(91) *awúuassshiichih*
awuua-ss-shiichi-h
inside.N-GOAL-throw-IMP
'Throw it in [the] inside! [of the hoop]' (Graczyk 1989:3)

(92) *akúkaachissaak* *duúom*
aku-kaa-chissaa-ak duu-o-m
beyond.N-SOURCE-return-SW.REF.SAME.SBJ come-PL-SW.REF.DIFF.SBJ
'They came back from [the] beyond. [beyond the hill]' (Graczyk 2007:384)

In examples (91) and (92), there are two consecutive postposition-appearing morphemes in which the leftmost appears to be lacking a governed phrase. However, I believe that the leftmost "postposition" itself is plausibly the term governed by the adjacent postposition. Per this analysis, the governed term undergoes conversion and functions as a noun. Example (92) is depicted syntactically in example (93) below.

[45]Conversion without any segmental change is common in English. When this shift in word-class includes prosodic changes, it is said to have undergone suprafixing, referring to the altered suprasegmental features.

(93) Syntactic diagram for example (92)



As illustrated in the syntax tree in example (93), this theory allows for all structural necessities to be filled without compromising the semantics of the utterance.

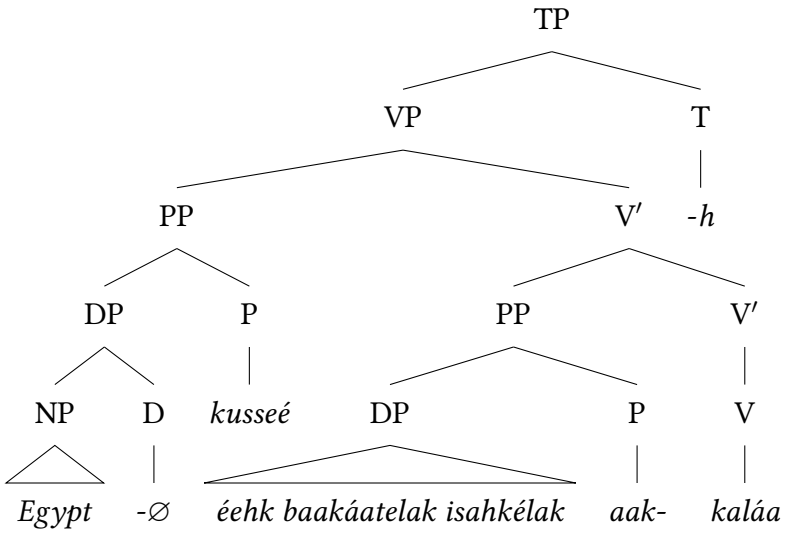
6.6. Complex Cases

Another noteworthy phenomenon in Crow's adpositional morphosyntax is the grammaticality of stacked procliticized adpositions with *distinct* governed terms. Consider the following utterance and subsequent illustrations.

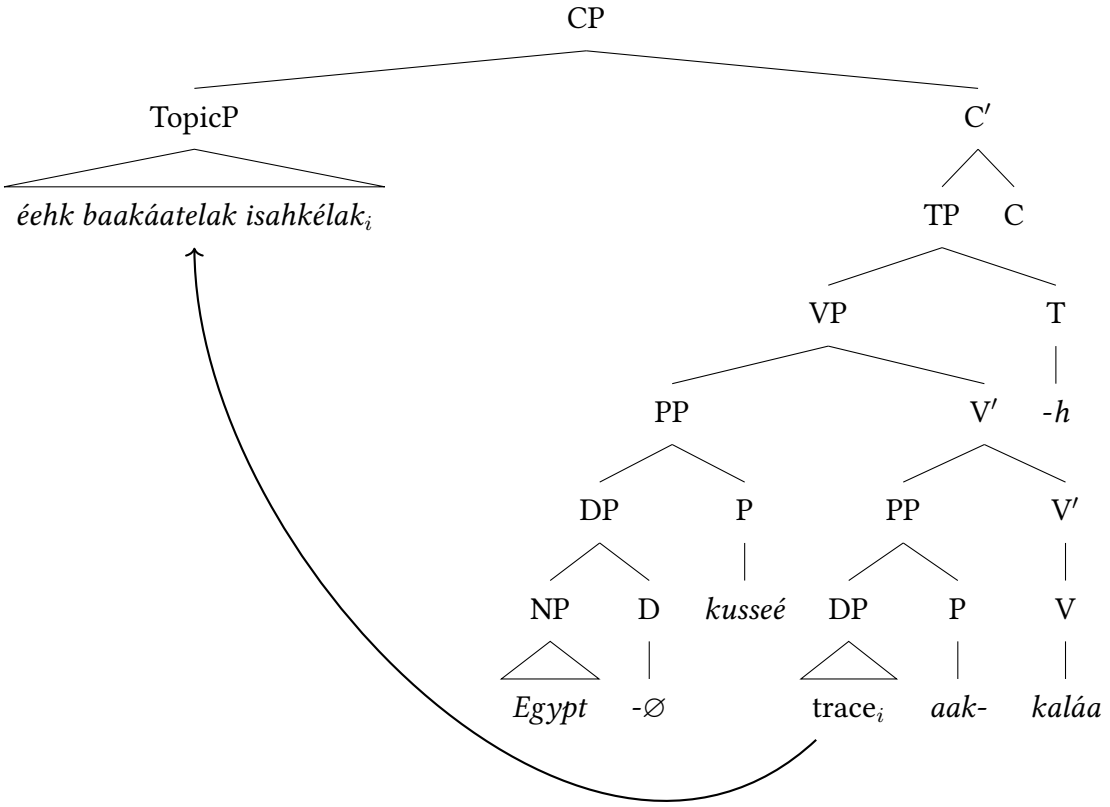
(94) *éehk baakáatelak isahkélak Egypt kussaakkaláah*

a. *éehk baakáate-lak isahké-lak Egypt kuss-aak-kaláa-h*
DET child-and his.mother-and Egypt GOAL-COM-flee-IMP
'Flee to Egypt with that child and his mother!' (Graczyk 2007:388)

b. Underlying Structure



c. Surface Structure (Topicalization)



Example (94a) contains a set of stacked adpositional proclitics on the verb /kaláa/. Based on the lexical semantics of the sentential constituents, it is clear that these two postpositions are not compounding to create a single compositional meaning, as the use of the comitative with 'Egypt' would not make sense. It is obvious that the location (Egypt) must be the GOAL and that the animate actors must be governed by the comitative (COM) adposition. These observations necessitate a syntactic theory capable of explicating the correct underlying postpositional government *and* the stranded postpositions in surface structure. Example (94b) accounts for the government relationships, but is incongruent with the realized surface structure. However, using example (94b) as the underlying structure, we can justify the stranded and stacked adpositions of the surface structure via topicalization-induced movement. This mechanism is illustrated in example (94c).

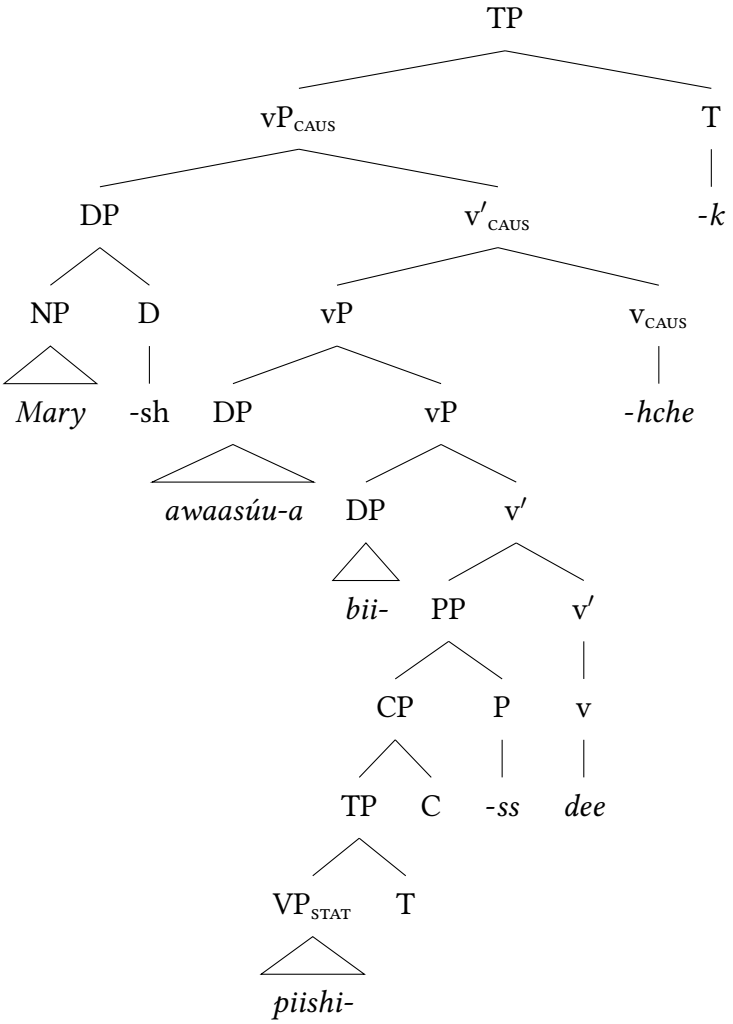
Graczyk (2007:198) gives two examples of topicalization, but does not accompany these with any description or explanation. In Crow, nouns have both a stem form and a citation form. The citation form is able to be used independently. Graczyk (2007:30-33) gives the example of "What is the word for X_{Engl}?" where the response—"X_{Crow}"—would be given in the citation form. I believe there is a connection between the use of this form and topicalized elements.[46] The morphemes for 'child' and 'his mother' found in example (94)—/baakáate/ and /isahké/, respectively—are the citation forms of these nouns (Graczyk n.d.:117, 368). Example (95) provides further data related to this theory of movement in Crow syntax.

[46]Rizzi's (1997) characterization of the left periphery influenced the convention of topicalization movement proposed herein.

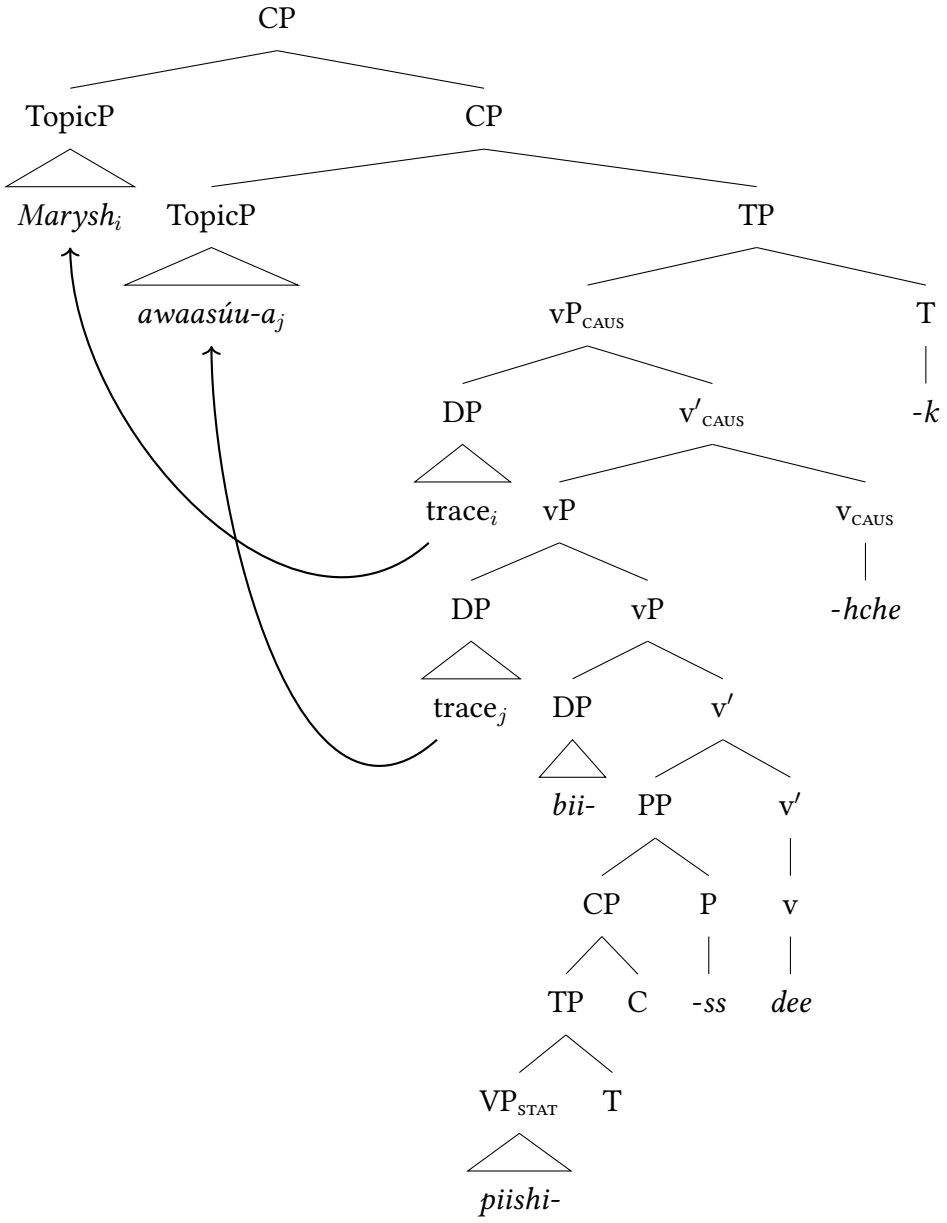
(95) *Marysh awaasúua biipíishissdeehchek*

a. *Mary-sh awaasúu-a bii-píishi-ss-dee-hche-k*
Mary-DEF house-TOP 1SG.OBJ-be.behind.(STAT)-GOAL-go-CAUS-DECL
'Mary sent me to the back of the house.' (Graczyk 1989:6, Graczyk n.d.)

b. Underlying Structure



c. Surface Structure (Topicalization)



Example (95) exhibits two instances of topicalization movement, but the phonetic realization of the utterance remains unchanged because the order of constituents is static. Without movement, we would be able to create a syntactic diagram that correctly reflected the realization of the phonetic form; however, this would be purely coincidental. Both 'Mary' and 'house' support the theory that the use of citation forms is connected to topicalization. For 'house,' /awaasúua/ is the citation form of /awaasúu/ (Graczyk n.d.:82). For 'Mary,' the situation is more complex. The definite marker (/-sh/) typically combines with a citation form, so even though "Mary" is clearly a loanword, we know it is functioning grammatically as a citation form in this instance (Graczyk 2007:32-33).

In example (94), incorporating movement into our working theory of Crow syntax was necessary in order to explain the phonetic form. In example (95), this is not the case. However,

because of examples like (94), it is clear that topicalization movement in Crow occurs before spellout and we thus must include it in examples like (95).

7. Discussion of Crow

The evidence from Crow presented in the section above solidifies the theory suggested by the analysis of Catawba in section 4. Siouan languages exhibit substantial variation in their systems of (para-)adpositional morphosyntax and can differ quite markedly from LDN. Although Catawba is an Eastern Siouan language and is often dismissed because of this, the same cannot be done with Crow, as it is a Western ("core") Siouan language like LDN. Crow, in contrast with LDN, exhibits a surprising degree of flexibility in its adpositional morphosyntax, with one postposition having as many as four distinct forms, each with a unique morphosyntactic locus; this is delineated in section 6.1.

In addition to free postpositions (section 6.4), Crow exhibits left-anchored (section 6.3.1), right-anchored (section 6.3.2), and bidirectionally anchored (section 6.3.3) compounding. Often, a single adposition is capable of inhabiting more than one of these morphosyntactic loci. For example, we see the comitative /aak/ left-anchored in example (79), but right-anchored in example (83) and free-standing in example (89). Moreover, the boundary between applicatives (discussed in section 6.2) and adpositions is not clearly demarcated, making an already intricate picture all the more complex. Some adpositions in Crow appear to be able to undergo conversion, becoming nouns (section 6.5). Furthermore, in addition to governing determiner phrases, postpositions in Crow can govern stative verbs (section 6.6). Crow's adpositional morphosyntax is very much its own, differing as much from LDN as from Catawba.

As mentioned above, some postpositions in Crow can undergo conversion, changing their word class without modifying their phonetic realization. Section 6.6 illustrates this ability, wherein they can be governed by another postposition and reflect a nominal location itself rather than a locational *relationship* between entities. This occurs in English, as well. For instance, "I came from behind" versus "He is behind the car." In addition to postpositional stacking in which one postposition governs the other, example (94) demonstrates that Crow can stack postpositions that each have their own governed term—a form of adposition-stranding. This is achieved through topicalization movement, in which at least one postposition governs a trace in the surface form.

In section 6, /ii/ was realized as both a right-anchored constituent of a compound—section 6.3.2, example (82)—and as an independent postposition—section 6.4, example (85)—in different prosodic environments.[47] This /ii/ is cognate with Crow's applicative /i(i)-/ ('against'), as the /i-/ applicative in LDN (section 2.4) conveys two senses: 'against' and INS. This provides further evidence for the intimate link between adpositions and applicatives in Siouan. This relationship is revisited in section 8.

The superessive applicative in Crow is far more morphosyntactically flexible than a mere fossilized prefix.[48]. In addition to its fossilized form (example (73)), it also appears as a semanti-

[47]Further research is required to determine the precise parameters of alternation. In addition to prosodic features, lexical constraints also likely play a role. Elicitation of more data via fieldwork is necessary.

[48]There are two distinct morphemes that surface as /aak-/, and only one is a cognate of the superessive. The comitative /aak/, like /ii/, can be used as either a proclitic or a free postposition. However, this is *not* a cognate

cally meaningful applicative (example (74)). Moreover, an extended form of the superessive (/aakee/) exhibits both left-anchored and right-anchored compounding; this is shown in examples (75) and (81), respectively. /aakee/ can also attach to the generic locative enclitic /=n/, becoming /aakeen/, and function as an independent postposition (illustrated in example (90)).

The comitative postposition /aak/ is similarly flexible in its morphosyntactic orientation. In example (79), /aak/ is compounded leftward. In example (83), it is compounded to the right. Moreover, /aak/ can appear independently (Graczyk 2007:236). Its alternative form, /áappaa/, is a free-standing postposition (87). In example (94), /aak/ is compounded rightward, while simultaneously having the postposition /kuss/ attached to its left. However, this is not an example of bidirectionally anchored compounding. /kuss-/ does not undergo conversion and is thus not being governed by /aak/. Rather, the complement of /aak/—the DP it governs—underwent topicalization movement, leaving a trace between /kuss-/ and /akk/. As /kuss-/ mandatorily undergoes right-anchored compounding, it attached to /aak/, resulting in the recorded surface structure.

In addition to the example of bidirectionally-anchored compounding in section 6.3.3 (example (84)), example (95) contains an instance of this with /-ss-/. Rightward, /-ss-/ compounds with the verb dominating it; leftward, it compounds with the stative verb /piishi/, which functions adjectivally.[49] The diagrams herein reflect my analysis of this structure as being an adpositional phrase taking a CP complement.

8. Synthesized Discussion

The evidence presented in the sections above strongly suggests that our current understanding of adpositions in LDN, Catawba, and Crow is insufficient. Most previous studies have failed to address the intricacies of (para-)adpositional morphosyntax in these languages beyond phenomena parallel to those already discussed in previous work on LDN. Moreover, no publication to date has provided a comparative analysis of adpositions in Siouan. This paper serves to partially fill that gap, providing such an analysis in the more modest context of the three languages examined herein. In summary, there is more diversity within the Western Siouan branch than the Siouanist literature presents; furthermore, the Eastern Siouan branch is not the decidedly dissimilar outlier that much of the Siouanist literature characterizes it to be. While LDN, Catawba, and Crow all exhibit free-standing, prosodically independent postpositions, the similarities shared by the (para-)adpositional morphosyntax of all three begin and end with this feature.

In both LDN and Crow, stative verbs can immediately precede postpositions. This is illustrated in examples (11) and (95), respectively. I argue that the postpositional phrase takes a complementizer phrase as a complement in both cases.

Catawba and Crow both exhibit right-attaching adpositional phenomena in which an adposition is attached to the verb that immediately dominates it. I analyze this process as proclisis in Catawba and compounding in Crow, but these are nonetheless markedly similar operations.

One similarity shared by LDN and Catawba is the presence of both a clitic form and a free form of some adpositions. For example, LDN has the free postposition /étkiya/ (example (36)) and the corresponding enclitic form /=kiya/ (example (32)). Analogously, Catawba's free

of the superessive applicative. This /aak/ is derived from /eé/ 'to have', which ablauts to /aá/ when preceding the same-subject morpheme /-ak/ (Graczyk 2007:388). The morpheme /áakee/ 'on top' *is* cognate to the superessive.

[49] Crow, as with other Siouan languages, does not contains adjectives. Stative verbs often fill this role.

postposition /hapki/ (example (59)) has the proclitic form /hap=/ (example (64)). Additionally, in both LDN (example (11)) and Catawba (example (70)), there are potential cases of extra-syntactic parentheticals. In both instances, I find other analyses to be more accurate and elucidative of the given data; however, the possibility is still worth mentioning.

LDN and Catawba also share a feature in which the presence of an adposition assigns a [+ambiguous definiteness] feature to the preceding noun, resulting in a null determiner (though this rule is violable in both languages). Ingham (2003) hints at this in LDN by pointing out that the use of a postposition seems to often preclude the presence of a determiner immediately beforehand in Catawba. This was previously unattested.

To some extent, the perception of Catawba as particularly distinct from the Western Siouan languages is fair. Regarding its (para-)adpositional morphosyntax, this is most clearly seen in the absence of applicatives. The semantic work carried out with applicatives in LDN and Crow is performed exclusively by adpositions in Catawba. In addition to elucidating a unique aspect of Catawba, this also illustrates the intimate relationship between adpositions and applicatives in Siouan.

Section 1.1 illustrated and discussed Helmbrecht & Lehmann's (2008)) theory of diachronic applicative fossilization. Table 1 from that section is reproduced below.

Table 2: Helmbrecht and Lehmann's Four Stages

	Stage One	**Stage Two**	**Stage Three**	**Stage Four**
Time:	Proto-Siouan	*Not Stated*	*Not Stated*	Present
Status:	Postpositions	Proclitcs	Applicatives	ISCs

The evidence and analysis herein *does* support the underlying notions of Helmbrecht & Lehmann's (2008)) theory. There is a historical process in which free constituents lose their status as prosodic words, becoming proclitics or right-anchored constituents of compounds (though they can also become enclitics or left-anchored constituents of compounds, which is not predicted in their theory). Proclitics and right-anchored constituents can in turn lose their status as morphological words, becoming affixes (such as applicatives), which can subsequently undergo semantic bleaching and fossilization, becoming synchronically unanalyzable. The discussion of LDN in section 3 showed that, although imperfectly, LDN generally follows the paradigm proposed by Helmbrecht & Lehmann (2008:34).

However, contra Helmbrecht & Lehmann (2008), these stages are not mutually exclusive. The relationships between Siouan postpositions, applicatives, and their intermediate forms are far more intricate and entropic than Helmbrecht and Lehmann intimate. The aforementioned constituents are not only related historically, as discussed by Helmbrecht & Lehmann (2008), but are also related synchronically.

One piece of evidence for the intimate synchronic relationship between adpositions and applicatives comes from juxtaposing analyses of LDN and Crow. Recall that in LDN, the instrumental applicative /i-/ has two semantic realizations: the instrumental meaning its name implies and the locative 'against.' In Crow, by contrast, the "instrumental" applicative /i(i)-/ can *only* function as a locative meaning 'against.' However, its related form /ii/—which can appear as a free postposition or compounded rightward onto the verb dominating it—is Crow's instrumental adposition.

Crow's superessive applicative /a(a)-/ appears as both a meaningful, productive applicative and as a semantically null, fossilized element on many verbs. This would place it simultaneously in Helmbrecht and Lehmann's stages three and four. Similarly, Catawba's postpositions have both free forms and procliticized forms, occupying both stage one and stage two. Another example from Crow is its goal adposition, which can appear as /-ss-/, /-ssee/, /kuss-/, or /kusseé/ depending on the environment (Graczyk 1989:8). Thus, the linear development from free constituent to preverb to bound affix suggested by Helmbrecht and Lehmann does not apply exhaustively. While it could be argued that Helmbrecht and Lehmann are purposefully ignoring Catawba, the same cannot be said of Crow.[50] These are a few representative examples of many throughout this paper that evidence synchronic incongruence with Helmbrecht and Lehmann's theory.

The (para-)adpositional morphosyntax of all three languages examined herein exhibits far more complexity and variation than the Siouanist literature to date indicates. In light of these findings, Siouanists (and, more broadly, linguists) should analyze adpositions more closely in future research, recognizing the morphosyntactic diversity of the word-class. Moreover, this study revealed a complex synchronic relationship between adpositions and applicatives. In Crow, the distinction between these grammatical entities is rather opaque. This suggests that—minimally—phenomena relating to Siouan adpositions and applicatives should be analyzed adjacently going forward.

9. Conclusions

The examination herein of the (para-)adpositional morphosyntax of LDN, Catawba, and Crow indicates that adpositions have been largely underanalyzed by Siouanists to date. This has not only affected our understanding of Siouan adpositions, but also our understanding of Siouan applicatives. Siouanists have unconsciously established a de facto description of "Siouan" adpositional morphosyntax that is based primarily on the (para-)adpositional phenomena of LDN, which—as sections 2 and 3 illustrate—is itself inadequate. While data from Catawba may be shrugged off as outside the so-called "core" Siouan languages (the Western branch of the family), the analysis of Crow in sections 6 and 7 demonstrates that the problem of adpositional underanalysis is endemic to the Western branch, as well.

The adpositions and applicatives of the Siouan languages—though their historical relationship is acknowledged—are synchronically treated as entirely distinct phenomena. This analysis works well for LDN, as adpositions are either free or anchored to the left and applicatives are always verbally prefixed. However, despite this dichotomy not extending to the whole Siouan family, the isolated treatment of both phenomena has. As discussed in section 8, the staged chronology of Helmbrecht & Lehmann (2008) presents Siouan adpositions, applicatives, and the relationship between them as far simpler than this research reveals. This is a result of Siouan scholars to date treating adpositions and applicatives as unrelated categories in their formal synchronic analyses—not a fault of Helmbrecht and Lehmann.

My examination of LDN, Catawba, and Crow ultimately reflected the views espoused by Hagège (2010): adpositional systems are underdocumented and underanalyzed. This paper hopes

[50] As discussed in sections 1 and 4, Siouanists tend to treat it as an irrelevant outlier; Helmbrecht (2006) does not mention Catawba and its lack of applicatives in his paper on applicatives in Siouan, for instance.

to serve as a stepping stone towards remedying this deficiency in the Siouanist literature.

Future research on (para-adpositional) morphosyntax in Siouan should include similar surveys on other Siouan languages. It is entirely possible (if improbable) that Catawba and Crow are the only outliers. Figure 1 provides the names and phylogenetic loci of numerous other languages that ought to be explored.

Moreover, much of the research herein is preliminary. In all three languages of inquiry herein, I have proposed morphosyntactic phenomena that scholars in the field thus far have not examined. This marks the beginning of the scientific research process, not the end. All proposals need to be evaluated and tested against novel data to see if they stand up to scrutiny.

For both LDN and Crow, fieldwork is a necessary component of further research. Elicitations with native speakers could easily confirm or reject a number of the hypotheses herein. The Siouan languages and their (para-)adpositional morphosyntax are fertile grounds for further research. This paper is intended to start these conversations, not to end them.

Abbreviations

1	first person
2	second person
3	third person
A	agent or active
ABS	absolutive
ADP	adposition
ADV	adverb
APPROX	approximate
CAUS	causative
COM	comitative
CONT	continuous
DECL	declarative
DEF	definite
DET	determiner
DIFF.SBJ	different subject
DU	dual
EMPH	emphatic
ERG	ergative
GOAL	goal
HAB	habitual
IMP	imperative
INCEP	inceptive
IND	indicative
INDEF	indefinite
INESS	inessive
INF	infinitive
INS	instrumental
INTENSE	intensifier
LOC	locative
N	noun
NARR	narrative
OBJ	object
PAT	patient
PATH	path
PL	plural
PST	past tense
R	reduplication
SAME.SBJ	same subject
SBJ	subject
SG	singular
SOURCE	source
SPEC.LOC	specific locative
STAT	stative
SUPERESS	superessive
SW.REF	switch reference
TOP	topicalizer
V	verb

References

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

Anderson, Stephen. 2005. *Aspects of the theory of clitics*. Oxford, UK: Oxford University Press.

Asbury, Anna. 2008. *The morphosyntax of case and adpositions [De morfosyntaxis van casus en adposities]*. Utrecht, Netherlands: Universiteit Utrecht dissertation.

Boas, Franz & Ella Deloria. 1939. *Dakota grammar*. New York, NY: Columbia University Press.

Booij, Geert. 2005. *The grammar of words*. Oxford, UK: Oxford University Press.

Bowern, Claire. 2018. *The mysterious Taensa Grammar*. Invited talk at 13th annual Whatmough Lecture for the Harvard Department of Linguistics. Cambridge, MA: Harvard University. https://www.youtube.com/watch?v=wDFmnKBVJec. (Accessed on 9 September, 2021).

Broekhuis, Hans. 2015. *Syntax of Dutch*. Amsterdam, Netherlands: University of Amsterdam Press.

Deloria, Ella. 1932. *Dakota texts*. New York, NY: G. E. Stechert & Co.

Dixon, Robert M. W. & Alexandra Y. Aikhenvald. 2003. *Word: a cross-linguistic typology*. Cambridge, UK: Cambridge University Press.

Dorsey, James Owen. 1885. On the comparative phonology of four Siouan languages. In *Smithsonian Institution Annual Report for 1883*. Washington, DC: Smithsonian Institution.

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

Gebhardt, Lewis. 2019. Incorporation in crow? In Ryan M. Kasak (ed.), *Proceedings of the 38th annual Siouan and Caddoan Languages Conference*, 145–154. Chicago, IL: Northeastern Illinois University Linguistics Department.

Graczyk, Randolph. 1989. *Postpositional phrase incorporation in Crow*. Paper presented at the 9th Siouan and Caddoan Languages Conference. Morley, AB: Stoney Nation.

Graczyk, Randolph. 2007. *A grammar of Crow*. Lincoln, NE: University of Nebraska Press.

Graczyk, Randolph. n.d. *Crow dictionary*. Ms. Pryor, MT.

Hagège, Claude. 2010. *Adpositions*. Oxford, UK: Oxford University Press.

Helmbrecht, Johannes. 2006. *Applicatives in Siouan: A study in comparative Siouan grammar*. Paper presented at the Siouan and Caddoan Languages Conferences. Billings, MT: Rocky Mountain College.

Helmbrecht, Johannes & Christian Lehmann. 2008. Hocank's challenge to morphological theory. In K. David Harrison, David S. Rood & Arienne Dwyer (eds.), *Lessons from documented endangered languages* (Typological Studies in Language 78), 271–316. Amsterdam, Netherlands: John Benjamins.

Hewson, John & Vit Bubenik. 2006. *From case to adposition: The development of configurational syntax in Indo-European languages*. Philadelphia, PA: John Benjamins.

Ingham, Bruce. 2001. *English-Lakota dictionary*. Richmond, UK: Curzon.

Ingham, Bruce. 2003. *Lakota sketch*. Munich, Germany: LINCOM Europa.

Kasak, Ryan M. 2016. A distant genetic relationship between Siouan-Catawban and Yuchi. In Catherin Rudin & Binah T. Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 5–39. Berlin, Germany: Language Science Press.

Kasak, Ryan M. 2019. *Affix ordering and templatic morphology in Mandan*. New Haven, CT: Yale University dissertation.

Kasak, Ryan M. 2020a. *A computational approach to Siouan Phylogeny*. Ms. Norman, OK: University of OKlahoma.
Kasak, Ryan M. 2020b. *Personal correspondence with Dr. Ryan Kasak*.
Kurzon, Dennis & Silvia Adler (eds.). 2008. *Adpositions: Pragmatic, semantic, and syntactic perspectives*. Philadelphia, PA: John Benjamins.
LLC. 2021. *New Lakota Dictionary Online*. Lakota Language Consortium. [version 5.0]. https://nldo.lakotadictionary.org/. (Accessed 9 September, 2021).
Mithun, Marianne. 1999. *The languages of Native North America*. Cambridge, UK: Cambridge University Press.
Olthof, Marieke. 2020. *Incorporation: Constraints on variation*. Amsterdam, Netherlands: Universiteit van Amsterdam dissertation.
Park, Indrek. 2012. *A grammar of Hidatsa*. Bloomington, IN: Indiana University dissertation.
Peterson, David. 2007. *Applicative constructions*. Oxford, UK: Oxford University Press.
Pustet, Regina. 2000. Lakota postpositions. *International Journal of American Linguistics* 66(2). 157–180.
Rankin, Robert, John Boyle, Randolph Graczyk & John Koontz. 2003. A synchronic and diachronic perspective on 'word' in Siouan. In Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A cross-linguistic typology*. Cambridge, UK: Cambridge University Press.
Riggs, Stephen. 1895. *Dakota grammar, texts, and ethnography*. Washington, D.C.: Government Printing Office.
Rizzi, Luigi. 1997. *The Fine Structure of the Left Periphery*. Liliane Haegeman (ed.). Dordrecht, Netherlands.
Rood, David & Allan Taylor. 1996. *Sketch of Lakhota, a Siouan language*. Ives Goddard (ed.). Washington, DC: Smithsonian Institution. 440–482.
Rudes, Blair. 2007. *Morphology of Catawba verbs, nouns, and particles*. Ms. Charlotte, NC: University of North Carolina at Charlotte.
Rudin, Catherine & Binah T. Gordon (eds.). 2016. *Advances in the study of Siouan languages and linguistics*. Berlin, Germany: Language Science Press.
Shea, Kathleen. 1984. *A Catawba lexicon*. Lawrence, KS: University of Kansas M.A. thesis.
Speck, Frank. 1934. *Catawba texts*. New York, NY: Columbia University Press.
Speck, Frank G. 1913. Some Catawba texts and folk-lore. *Journal of American Folklore* 26(10). 319–330.
Ullrich, Jan. 2018. *Modification, secondary predication and multi-verb constructions in Lakota*. Düsseldorf, Germany: Heinrich Heine University Düsseldorf dissertation.
Ullrich, Jan. 2020a. Complex predicates with nouns and stative verbs in Lakota: A Role and Reference Grammar analysis. *International Journal of American Linguistics* 86(3).
Ullrich, Jan. 2020b. *Personal correspondence with Dr. Jan Ullrich, October*.
Ullrich, Jan & Jr. Black Bear Ben. 2016. *Lakota grammar handbook*. Bloomington, IN: Lakota Language Consortium.
Voorhis, Paul. 1992. *Catawba vocabulary from the texts recorded by Frank Speck*. Ms. Brandon, MB: Brandon University.
Voorhis, Paul. n.d. *Catawba-English index*. Ms. Brandon, MB: Brandon University.

Topic and focus in Mandan*

Ryan M. Kasak
University of Oklahoma

Abstract: This paper investigates how topics and focus is marked in Mandan, specifically the interface between morphology, syntax, and prosody. Mandan has an enclitic *=na* that has variously been described as a topic marker, a topicalization marker, and a focus marker, and this paper demonstrates that this marker is used by speakers to shift the listener's attention to a new topic or an already established topic, most often a grammatical subject. Topicalization also occurs without *=na* with particular intonation, indicating that there are multiple strategies for marking topics in the discourse structure of Mandan.

Keywords: Mandan, topicalization, focus, prosody, information structure

1. Introduction

The bulk of Siouanist linguistics over the past century has focused on the morphological and syntactic properties of Siouan languages. The earliest attempts at looking at the grammar of Siouan languages were often done through the lens of how segments, formatives, and words interacted with each other, while ignoring matters of the suprasegmental: i.e., phrasal pitch accent, intonation patterns, and sometimes even ignoring stress itself. Early works that delve into aspects of the grammars of Siouan languages are geared towards enabling readers to parse through a collection of transcribed narratives, such as Boas & Deloria's (1941) *Dakota Grammar*, or to add context to why certain lexemes appear in a dictionary, such as in Dorsey & Swanton's (1912) *Dictionary of the Biloxi and Ofo Languages*. Whether intentionally or unintentionally, these grammars and grammars like them are designed to accommodate the understanding of their languages through a writing-centric point of view. This approach has not paid attention to suprasegmental phenomena, which is regrettable, as these languages all have traditionally placed great emphasis on the oral transmission of culture and writing has only been introduced in recent generations. There is rarely any discussion on the topic of prosody, and that lack of discussion causes a deficit in our understanding of the information structure of the language uttered by L1 speakers.

Hirst & Di Cristo (1998:1) remark that intonation systems are "one of the most language specific features of human language." This claim is supported by psycholinguistic studies conducted on neonates, who demonstrate sensitivities to the prosody of their parents' language(s) that are likely due to passive exposure *in utero* (Ramus 2002, May et al. 2011, *inter alios*). One phenomenon that is often associated with intonation systems is that of topic-marking and focus-marking. Mandan possesses an enclitic *=na* that has been varyingly described as a topic marker, topicalizing marker, or a focus marker in existing descriptions of the language. Previous attempts

*Many thanks to Mrs. Delores Sand, Mr. Valerian Three Irons, Mr. Leon Little Owl, and the late Mr. Edwin Benson, who were all consultants who have contributed over the years to the Mandan audio data upon which this work is based. Without their invaluable knowledge of the Mandan language, this work would not have been possible.

to categorize the function of this enclitic have relied on textual corpora, rather than audio data that might help disambiguate what the distribution of *=na* is when factoring in the contextual prosody of utterances where it appears. As such, there is a need to identify and understand what the role of *=na* is and what kinds of intonational cues might coincide with its appearance on a constituent. With *=na* being a marker of information structure and with topic and focus often being associated with particular intonational tunes, this element is a prime entryway into the overall topic of how prosodic elements in Mandan interact with how information is packaged in an utterance.

This paper serves as a preliminary investigation into the interface between information structure and prosody in Mandan. The research herein has the following three goals:

(1) Goals dealing with the information structure–prosody interface in Mandan
   a. To examine previous literature on topic- and focus-marking in Mandan;
   b. To compare and contrast previous literature with audio recordings and their transcriptions;
   c. To provide insight into a subject that can potentially be of use towards Mandan language instruction.

To these ends, I investigated, transcribed, and labeled the first five-minutes of a recording of Mr. Edwin Benson (1931–2016) recounting the narrative "Blackwolf," also called "The Gambler," plus excerpts from Mrs. Mattie Grinnell (1867–1975) recounting the traditional narrative "No Tongue."[1] Utterances were tokenized and labeled in Praat, Version 6.1.27 (Boersma & Weenink 2020). Praat pictures to display the pitch (F0) track of each utterance were created using a Praat script by Elvira-García (2017) that I slightly modified.

The work herein initially explores what is meant by "topic" and "focus" in §2, as both these terms occur frequently in grammars of Siouan languages—and of other language families as well—without explicitly stating the role of the constituent designated as such In §3, I explore the structural manifestations of topic and focus, namely the role that a constituent bearing the enclitic *=na* plays within the utterance. Prosodic manifestations of topic and focus are discussed in §4, wherein we can observe the behavior of pitch on topicalized and focused element. I discuss the morpho-syntactic and prosodic manifestations of topic and focus in Mandan and how they interact in §5, and then conclude with some generalizations about topic marking and focus in Mandan, along with some avenues for future research in §6.

2. Topic and focus

Various terms relating to some kind of prominence within the information structure of an utterance are used in many Siouan languages, typically without any explanation by the describer for employing one term over another. In §2.1, I address some of the formatives and terms employed in other Siouan languages relating to this prominence in the information structure that is morphologically marked, and I then provide theoretical context behind terms like topic and focus in

[1]The recording of "Blackwolf" was done by Dr. Sarah Trechter along with Mr. Corey Spotted Bear between 2007 and 2010 in Twin Buttes, North Dakota. The recording of "No Tongue" was conducted by the late Dr. Robert Hollow in Twin Buttes, North Dakota, sometime during his doctoral fieldwork between 1966 and 1968.

§2.2. With these delineations between what is a topic versus what is a focused element, I examine whether either of these pragmatic notions have a specific morphological manifestation, or if syntactic structures and/or prosody play a major role in indicating these emphasized elements in the information structure in §2.3.

2.1. Information structure marking in Siouan

In other Siouan languages, there are morphological elements that mark emphasis, focus, topic, or topicalization. Rankin et al. (2015) posit several reconstructions for elements that are associated with topics or focus in Proto-Siouan: e.g., *-ya 'emphatic topic,' *-ri 'focus,' and *-šV 'emphatic, contrastive.' We can see reflexes of these elements in the data below in bold.

(2) Examples of topic, focus, and emphasis marking in Siouan languages

a. Nakoda[2]

*zitkána žé nína jústina-**ȟtįyą***
bird DEM very small-**INT**

'this bird is the smallest' (Collette 2019:81)

b. Lakota

Missionta mníŋ kteló. Níš?
Mission-ta w-yA ktA-lo ni-**š**
Mission-LOC 1A-go POT-IND.M 2S-**EMPH**

'I am going to Mission. What about you?' (Ingham 2003:51)

c. Tutelo

*hąksí-k-**ya** wį-ki:tǫ́*
stick-DEF-**EMPH** 1SG.DAT-belong

'my stick' [lit. 'the stick belongs to me'] (Oliverio 1996:130)

d. Hidatsa[3]

*nuxbáaga iháhdaa**ri** wiiguxdáabag*
ruxbaaga ihaa-hdaa-**ri** wii-guxdi-aaba-g
people different-GOAL-**TOP** 1S-help-PL.COLL-SS

'the people from the other clans helped all of us' (Boyle 2007:70)

e. Biloxi

*ątatka-**yą** khu-ni ǫni e-tu xa*
child-**TOP** give-NEG PST say-PL HAB

'she did not give him the child, they say' (Kaufman 2008:155)

[2]The so-called intensive marker in Nakoda (a.k.a. Assiniboine) appears to be a combination of the pan-Siouan augmentative marker *xtE and the emphatic topic *ya. Nasalization of the topic element looks to be progressive nasal harmony stemming from the preceding syllable. This is an expected process in Lakota, as mentioned in Kasak & Lundquist (2019:103*ff*). If the nasality on the second syllable of the intensifier here is allophonic rather than phonemic, then it is possible that progressive nasal harmony is a feature in Nakoda as well.

[3]This datum comes from Boyle's (2007) dissertation, but the orthographic representation here has been altered to reflect my personal interpretation of the phonological representation of these words.

f. Hoocąk[4]

Nąąnį,	*nįį**ži***	*toįkewehige*	*waa'ų*	*hakiriną*
nąąnį	nįį-**ži**	too-hį-kewehi-ge	wa-ha-'ų	ha-kiri-ną
mother.VOC	1SG.PRO-**FOC**	PV-1S-hungry-CAUS	UNSP-1A-be	1A-return-DECL

'Mother, it is I, returning, extremely hungry' (Helmbrecht & Lehmann 2006:74)

The data above all display some reflex of the aforementioned Proto-Siouan formatives that have some connection to topic, focus, and emphasis. We likewise see an overlap between how these different elements are treated by scholars. For example, the suffix *-ya* in Tutelo is considered to be an emphasis marker in (2c), but it is cognate with the topic marker in Biloxi in (2e). There are additional instances of semantic differences, the emphasis marker *-š* in Lakota that we see in (2b) has a cognate in Hoocąk *-ži* that is typically referred to as a focus marker in (2f).

The goal of the present work is not to uncover the topic- or focus-marking elements in Proto-Siouan, but to investigate the synchronic systems for doing so in one Siouan language: Mandan. Mandan possesses one formative that is referred to throughout Kasak (2019) as a topic marker, the enclitic *=na* (underlyingly /=rą/). We can see examples of this element in the data below.

(3) Examples of *=na* marking in Mandan

a.

máahsi	*máakahe rátsee**na***	*káherekto'sh,*
wąąh#si	wąąkahe r'-at=s=ee=**rą**	ka'#hrE=kt=o'sh
arrow#feather these	2POSS-father=DEF=DEM.DIST=**TOP**	possess#CAUS=POT=IND.M

ínuma'ktaa
i-ruwą'k=taa
PV.DIR-man=LOC

'these feathers, your father gave them away, to the men' (Hollow 1973a:226)

b.

*ínuupshashka**na***	*húpe*	*ké'ka'rak*	*kú'kerek*
i-rųųp-sha-shka=**rą**	huup=E	ke'#ka'=ak	kų'=krE=ak
PV.ORD-two-COLL-INTS.COLL=**TOP**	shoe=SV	keep#have=DS	give=3PL=DS

'both of them kept shoes for him' (Hollow 1973b:109)

c.

óo	*úųpa**na***	*nurúsanaahini*	*nuhúuro'sh*
oo	ųųpa=**rą**	rų-ru-srąąh=rį	rų-huu=o'sh
DEM.MID	elk=**TOP**	1A.PL-INS.HAND-leave.behind=SS	1A.PL-come.here=IND.M

'we left an elk here and came' (Hollow 1973a:180)

We can see *=na* on nominal elements in (3) above, though not all of them are necessarily the first constituent within their respective utterances. Not appearing first suggests that there may not be a syntactically privileged position for elements bearing topic marking in Mandan: i.e., topics with *=na* are not inherently the leftmost element within the domain of a clause. Therefore, I argue that elements bearing *=na* represent morphological manifestations of information structure that are not inherently conveyed by their presence in a particular position within the syntax. That

[4]My thanks to Sarah Lundquist for helping me find and analyze this Hoocąk datum.

is to say, the =*na* is not obligatory for any particular element, as we see =*na* on both subjects, like in (3a) and (3b), as well as direct objects, like in (3c). The distribution of =*na* will be discussed in greater detail in §3.

2.2. Theoretical notions of topic and focus

The morphological elements discussed above in §2.1 are used by authors to indicate that there is some kind of prominence placed upon the words to which said elements adjoin. To consider which, if any, of the terms invoked above apply to Mandan =*na*, we should establish a baseline for evaluation. To this end, we must define what is meant by terms like "topic" or "focus" so that we do not perpetuate the use of *ad hoc* labels for this formative in Mandan. Adherence to some theoretical notion of what constitutes a topic, a focused element, or a topicalized element is important not only to the greater typology of how these elements manifest in human language, but it has a practical value to revitalization efforts.

2.2.1. Topic

A topic is sometimes referred to as a "theme" in discourse analysis, and it stands in opposition to what is being said about the topic, also called the "rheme" (Baker & Ellece 2011:151). Topics are some central piece of information within a selected stretch of discourse, meaning there may be topics that are clause-level topics or discourse-level topics. In (4) below, I have provided paraphrased translation of Mrs. Annie Eagle talking to Dr. Robert Hollow in Mandan about her garden. The topic at the discourse level below is "the garden" (Hollow 1973b:55*f*). The rheme is, ergo, what follows the theme throughout the rest of the discourse below.

(4) I want to tell you about **the garden** I have. When it is spring, I am not able to do it every spring, but I always manage to have myself a garden. Now, this summer, there was no one to plow it for me, though barely a month now, the one that just past, they did plow out there for me.

For those authors who argue for a systematic correlation between discourse roles and formal properties of topics, topics occupy a specific position within an utterance. Rizzi (1997, 2001) proposes that topics are uniquely positioned in the left periphery of a clause. In Frascarelli & Hinterhölzl (2007) and Frascarelli (2007, 2012), the authors connect the formal properties of topics to differing types of information conveyed by said topic. We can break these topics down into three types of topics. An aboutness-shift topic references Reinhart's (1981) notion of *aboutness*, where the topic is newly introduced, reintroduced, or something to which the theme of an utterance is changed. Contrastive topics mark some alternative entity in the discourse that is not the focus and serves as a counterpoint to other topics. A familiar topic is one that conveys some given information within the context of the discourse that has already been established or is a resumption of background information.

We can see in the example situation below where three different topics can be present at once in the underlined portion.

(5) This is the situation: I asked two of my professors, my Phonology professor and my Syntax professor, to spread the word about a Linguistics Club meeting this Friday afternoon to the

students in both of my classes. My Phonology professor told everyone about it at the top of the class, but **the meeting**, *my Syntax professor* TO OUR CLASS did not bring up at all.

This hypothetical speaker is a student who is talking about requesting that their two professors mention an upcoming Linguistics Club meeting that week. The student then relates how the Phonology professor followed through with that promise to let the students in that class know about the meeting, while the Syntax professor did not, much to the frustration of the student narrator. In the underlined portion, we see an aboutness-change topic in bold (**the meeting**), as the speaker is reintroducing what the utterance is about. This aboutness-change topic also happens to be the direct object of this proposition. In italics, we see the subject of the proposition as a contrast topic (*my Syntax professor*), indicating that the speaker is juxtaposing this topic with a counterpoint: the Phonology professor. The familiar topic is displayed in small caps (TO OUR CLASS), which marks the indirect object as background information that was mentioned earlier in the text. The underlined element is certainly not the default configuration for an English sentence and may initially seem to be quite arcane, but it is an acceptable one when delivered with the appropriate intonation and when given sufficient narrative context leading up to this series of topics presented sequentially.

2.2.2. Topicalization

Another related term that is discussed in Siouanist literature related to the markers found in (2) and (3) is topicalization. Ross (1967, 1986) introduces the term topicalization to refer to a constituent that has moved to the left peripherery of a clause. This constituent is then coindexed with an empty category left behind within the clause. Topicalized elements below are shown in bold.

(6) Example of topicalization with empty categories
Mushrooms, *I love* _ *on pizza, but* ***anchovies,*** *I hate* _.

Much of the discussion of topicalization in linguistic literature assumes an analysis that is consistent with the systemic correlation between discourse roles and formal properties of topics as outlined in §2.2.1 above. The primary assumption in much of the literature is that topicalization is a process whereby a topic is moved to the left periphery of an utterance.

Far less attention has been paid to matters of the right periphery. Growing research in recent years suggests a certain degree of overlap between the functions of topicalization and right dislocation (Beeching & Detges 2014; Fant, Bartning & Österberg 2021). However, pragmatic differences are purported to exist. Beeching & Detges (2014:11) suggest a functional assymetry between the left periphery and right periphery of a clause. The left periphery is strongly associated with serving as a landing spot for the grammaticalization of phenomena related to information structure or arguments structure. This periphery tends towards expressing subjectivity in the sense of Lyons (1982:102), where the speaker is expressing their own attitudes and beliefs. The right periphery, by contrast, is associated with modal or polarity phenomena. This periphery tends towards expressing intersubjectivity as defined by Traugott & Dasher (2002:11): the speaker's attention to the needs and self-image of the addressee.

For the purposes of the present work, we can assume an intersection between the notion topic as laid out in the systematic correlative sense in §2.2.1 and the movement of some topic to

the left periphery as topicalization. Furthermore, phenomena involving the right periphery will also be lumped under the term topicalization for the time being in order to investigate what kind of overlap might exist between the classical notion of topicalization and elements that have been right dislocated.

2.2.3. Focus

The notions of topic and focus in Siouanist literature can sometimes not seem to be very distinct ones. An element that is described as a topic marker by one scholar is described as a focus marker by another. It is not uncommon for there to be little to no explanation about why a scholar has elected to use one description of a morphological marker over another, and there has been virtually no discussion about any prosody associated with these formatives. Topics and topicalization have been addressed in §2.2.1 and §2.2.2, respectively, but this subsection addresses what is meant by focus within the confines of this paper.

Lee (2015:1) questions what is meant by focus by presenting the reader with a brief dialog, repeated below:

(7) Example of focus in English

a. *What does John drink?*

b. *John drinks beer.*

A question is posed in (7a), and an answer is provided in (7b). We can decompose (7b) further by noting that *John drinks* is now old information, given the fact that this information is simply repackaged from the original question. The new information, *beer*, is the true answer to the question. Within the information structure of this exchange, there is an informative part (*beer*) that acts as the focus, and there is an accompanying prosodic prominence to this element. Focus in this sense indicates a discourse function where a constituent informational item is emphasized (Ladd 1984, 2008; Xu, Xu & Sun 2005, *inter alios*).

There are multiple competing notions throughout the literature regarding the nature of focus, but for the purposes of this paper, I adhere to notion of focus as described above: i.e., there is a prosodic realization of emphasis that is grounded in the pragmatics or discourse structure. Xu, Xu & Sun (2005:81) remark that the consensus on focus is largely that it is expressed through variations in the fundamental frequency (F_0), though amplitude and duration can also play a role in focus marking. Focus is not inherently expressed structurally in the same way that topicalization is: i.e., through the movement of a constituent to some periphery.

Much like topics, focus is not a singular monolithic category. Lee (2015:2*ff*) outlines a typology of three classes of focus: discourse-new focus, contrastive focus, and corrective focus. Discourse-new focus involves information that is only just being introduced into the discourse and has no possible retrievability from prior context. Contrastive focus indicates some exhaustive choice from among some set of spoken or unspoken alternatives. Corrective focus is one that corrects information deemed false by the speaker. We can see examples of these foci below in (8) with each focus represented in brackets.

In (8a), we see that one person is posing a question to another about the price of a cup of coffee. The answer, *two dollars*, is novel information, making it discourse-new. Likewise, in (8b), the questioner provides two options for the responder, and the responder utilizes contrastive

focus in their answer, *turkey*, to highlight which kind of sandwich they prefer. Lastly, a questioner in (8c) inquires as to whether they have Jenny's phone number right, but they seem to be slightly off. The responder employs corrective focus on the one digit that the questioner had wrong, *9*, to ensure that the questioner understood which particular number was incorrect. All three of these foci can be omitted and these examples would still be grammatically correct, but there would be something slightly pragmatically deficient about the responses to the questions below.

(8) Examples of focus types

a. Discourse-new focus

i. *How much did that coffee cost?*

ii. *It cost* $[\textit{two dollars}]_{\text{F}}$.

b. Contrastive focus

i. *Between bologna and turkey, what kid of sandwich do you prefer?*

ii. *I* [*prefer* $[\textit{turkey}]_{\text{F}}$ $]_{\text{VP}}$.

c. Corrective focus

i. *Jenny's phone number is 867-5308, right?*

ii. *No, Jenny's phone number is 867-530*$[\textit{9}]_{\text{F}}$.

The examples in (8) above all involve foci that appear *in situ*: i.e., there is no change to the ordering of words within each clause. Rizzi (1997, 2001) notes that focus can also be present in a kind of topicalization within the left periphery. This analysis synthesizes the prosodic and pragmatic nature of focus with the syntactic structure of a clause. Under this proposal, there exists a specific functional projection in the left periphery of the clause above the tense phrase (TP) layer. A single focus may appear between two topics, but this analysis holds that each clause can contain a maximum of one focus.

2.2.4. Summary

There exists a whole swath of literature beyond that of Rizzi (1997, 2001) on the notion of topicalization versus focus, as well as how topics are marked in a language. We can see instances of both topicalization and focus in the example below. In (9a), the use of topicalization implies that the notes should be given to Parker as opposed to Parker receiving something else. We can contrast this usage with what we see in (9b), where focus on the direct object implies that the listener should give Parker their own notes as opposed to someone else's notes.

(9) Left-dislocation and information structure in English

a. Topicalization
Your notes, *you should give* _ *to Parker* (but not anything else).

b. Focus fronting
YOUR NOTES *you should give* _ *to Parker* (but not mine).

The word order for both topicalized and focused elements above can appear in similar positions within a sentence, but the pragmatic motivation for using one versus the other necessitates some

distinction between them. Speakers do not freely interchange topicalization with focus fronting because there are different motivations for doing one over the other.

Most discussions of the notion of "topic" in Siouanist literature have referenced the particular class of topics that equates some entity that is salient to the discourse with a syntactically prominent position in the left periphery as we saw in (5) and (6). Namely, the use of the term "topic" has been conflated with topicalized elements. Likewise, items designated as focus markers are typically found on nominals that are in the left periphery of a clause, rather than on focused elements appear in *in situ* along the lines of those seen in (8). Therefore, the locus of investigation should center around the notions of topic and focus as outlined above, where some constituent is moved to the left periphery for the purposes of conveying emphasis. Likewise, the work herein will delineate the differences between topic and focus in Mandan moving forward.

2.3. Towards an analysis for Mandan

The question of how topic and focus are realized in Mandan is not as straightforward than it seems at face value. All languages have intonation patterns, even those with lexical tone or pitch accent (Yang 2016). It is therefore a foregone conclusion that Mandan possesses a system of intonation that conveys attitudinal information or discourse functions. The question is, however, whether we are able to make meaningful analyses of intonation patterns in Mandan as they relate to topicalization and focus marking in the absence of L1 speakers who can provide judgments. One research goal of this paper is to establish that examining prosody and pragmatics on languages that no longer have speakers can be a meaningful topic of inquiry. Subsequent sections of this paper outline ways in which we can infer certain aspects of the interface between morpho-syntax and pragamatics, plus the prosodic system of Mandan.

3. Structural manifestations of topic and focus

The most obvious place to start in our discussion of the structural manifestations of topic and focus in Mandan is the enclitic *=na*. There is no unambiguously direct analog of *=na* in Proto-Siouan per the Comparative Siouan Dictionary, but we do see nasalized reflexes of the so-called emphatic topic marker *ya in other Siouan languages Rankin et al. (2015). Kaufman (2008:150) describes *=yą* as familiar topic clitic. This formative appears to be a cognate with Mandan *=na*, given that PSi *y merged with *r in Pre-Mandan, and /r/ becomes [n] before nasal vowels in the synchronic grammar of Mandan (Kasak 2019:130). As such, one possibility is that Proto-Siouan had competing forms for this formative, *ya∼*yą. Another possibility is that nasality on *ya could originate from the loss of some other morphological material bearing an underlying nasal that has since been lost except for the nasality assimilated onto the vowel in *ya.

3.1. Previous analyses of *=na*

Different authors have provided alternative analyses of the role of this enclitic within Mandan discourse structure. The first documented explanation for what this element is can be found in Kennard's (1936:26) *Mandan grammar*, wherein he states that "if the speaker wishes to designate either the subject or the object as the important element of a sentence, the suffix *-na* is used." This

formative is described as an emphasis marker and there is no further elaboration as to when one might expect to use it versus when one would never use it. One addendum to this description is that Kennard notes that this element frequently occurs with the demonstrative *ée*. He provides several examples of this formative:

(10) Examples of emphatic =*na* from (Kennard 1936:26)[5]

a. *sísohsii**na***
si#soh#sii=**rą**
feather#be.rounded.point#be.yellow=**TOP**
'a yellow hawk'

b. *súkee**na*** *ée**na***
suk=ee=**rą** ee=**rą**
child=DEM.DIST=**TOP** DEM.DIST=**TOP**
'that boy'

c. *kíishekasee**na***
kV-i-sek=ka=s=ee=**rą**
AGT-PV.INS-make=HAB=DEF=DEM.DIST=**TOP**
'the maker'

In all the examples in (10) above, we can see that this =*na* is not overtly changing the semantics of the nominal upon which it is found. Of particular note is the use of =*na* with the unbound version of the distal demonstrative *ée*, while the overt nominal *súk* features the bound version of the same demonstrative in (10b). This behavior or =*na* doubling will be discussed further in §3.2.

Hollow's (1970) dictionary of the Mandan language is the first to provide some kind of lexicon for the language that includes a morphological breakdown of lexemes. As such, this dictionary contains a list of morphological items and a brief description of how they are used, but there is no mention of this formative in the dictionary. This is surprising because there are ample instances of =*na* throughout his transcribed narratives, as evidenced by the fact that most examples are drawn from the two boxes of transcribed narratives that are part of the Robert C. Hollow Collection at the North Dakota State Historical Society archives.

Building on Hollow's (1970) dictionary, Coberly (1979) produces a grammatical sketch of Mandan using narratives originally transcribed by Kennard (1934) that were later re-elicited and transcribed by Hollow (1973a). She continues to call =*na* an emphatic marker per Kennard (1936), though she notes that Kennard's (1936) description of the so-called emphatic typically is accompanied by a vowel he transcribes as <e>, which he assumes is an indefinite article.[6] Coberly (1979:57) discounts =*na* being associated with indefiniteness as Kennard (1936) does, especially given the fact that there are numerous instances of what are transcribed as <e> plus <na> that occur on stems bearing overt definite marking, as we have seen earlier in (10c).

Wolvengrey (1991:585) likewise observes that many instances of the emphatic marker in Kennard (1934, 1936) may not be single formatives. To examine the nature of this formative, Wolvengrey conducts a corpus study of the distribution of =*na* in the transcribed narrative

[5] I employ glossing conventions for Mandan used in Kasak (2019) for the sake of consistency through the text.

[6] There is no indefinite article in Mandan, though there is a definite article. The nature of this word-final vowel in Mandan is addressed more thoroughly in Kasak (2019:317*ff*).

in Kennard's (1936) Mandan grammar, as well as three other narratives originally transcribed by Kennard (1934) that were later re-elicited and transcribed by Hollow (1973a). Wolvengrey (1991:588) analyzes this formative as a cleft-focus marker. This description as marking some kind of cleft seems to be consistent with the data. The use of "focus" to describe the motivation for this kind of cleft, however, is inconsistent with the terminology laid out in §2.2.3, where focus is primarily a prosodic phenomenon rather than a syntactic one.

Mixco (1997:41) describes =*na* as a topicalizing enclitic in his grammar. Topics marked by this enclitic may be nominals, including demonstratives that refer to a topic. Mixco identifies the <e> that often co-occurs with =*na* in Kennard (1936) and Coberly (1979) as the distal demonstrative *ée*.[7] This demonstrative can appear as a free word or an enclitic, and the topicalizing enclitic likewise encliticizes onto either a nominal or demonstrative. Mixco (1997) does not include any further discussion of the behavior of this formative beyond what has been stated above.

Kasak (2019) provides a partial grammar of Mandan, but does not discuss =*na* beyond labeling it as a topic marker. This lack of description is due to the fact that this work focuses on verbal morphology rather than nominal morphology. However, there is some discussion of the interaction between =*na* and verbal morphology when discussing unbound manifestations of the unspecified argument marker (Kasak 2019:243). The treatment of =*na* as a topic marker, sometimes referred to as a topizalizer, follows on Mixco's (1997) analysis of =*na*. There is no discussion of the kinds of topics represented by this formative in the senses established in §2.2.1, but mention of the treatment of =*na* in Kasak (2019) is mentioned here only because it represents the most recent description of this enclitic.

3.2. Distribution of =*na* in the corpus

One notable contribution by Wolvengrey (1991) is his discussion that =*na* occurs on more than just subjects and direct objects. He examines four transcribed narratives and finds 111 instances of =*na*. Table 1 below is a modified reproduction of the one in Wolvengrey (1991:586), where he lists the number of occurrences of =*na* and what role that element is playing in a clause. He classifies each occurrence of =*na* by what role the nominal it modifies plays: active subject (A. Subj.), stative subject (S. Subj.), direct object (Dir. Obj.), oblique object governed by postpositions (Prep. Obj.), direct reference to quoted speech (Quot.), possessor (Poss.), and adverbial (Adv.).[8] Wolvengrey likewise divides these roles by whether the =*na* occurs on a vowel-final stem (StemV), a stem with the distal demonstrative =*ee* (Stem =*ee*), a definite nominal with the distal demonstrative (Def. =*s*=*ee*) or alone as an unbound word (*ée*=*na*).[9]

[7]Previous descriptions of Mandan have irregular marking of long vowels (e.g., Kennard 1936) or deny the existence of long vowels altogether (e.g., Hollow 1970). See Kasak (2019:103*ff*) for further discussion of this issue in Mandan description.

[8]Wolvengrey (1991:586) refers to oblique objects as postpositional objects and to direct references to quoted speech as object complements. An example of direct reference to quoted speech would be "'Now," I said, so that is when I expect it to be ready.'

[9]Wolvengrey (1991:595*f*) hypothesizes that <-ená> is a single formative that differs from the <-ną>. This analysis comes from working with only textual sources and not from engaging in fieldwork with L1 Mandan speakers who can explain that what is being written as <-ená> is really two different formatives: a distal demonstrative =*ee* and the topic marker =*na*. This morphological structure is first proposed in Mixco (1997) and later confirmed in my own fieldwork with the late Mr. Edwin Benson.

We can see on Table 1 below that active subjects represent the overwhelming majority of instances where =*na* appears in the set of four narratives. Stative subjects and direct objects also occur with some frequency, but not to the degree of active subjects. Contrary to descriptions by Kennard (1936) or Coberly (1979), nominals that are not subjects or direct objects appear to take =*na*, though such occurrences are certainly rarer.

Table 1: Frequency and manifestation of =*na* in Wolvengrey's (1991) study

Role	StemV =***na***	Stem =***ee*=*na***	Def. =***s*=*ee*=*na***	***ée*=*na***	Totals
A. Subj.	20	15	39	5	79
S. Subj.	4	4	2	1	11
Dir. Obj.	5	5	1		11
Prep. Obj.	2	2			4
Quot.		1		1	2
Poss.	1		1		2
Adv.	1			1	2
Totals	33	27	43	8	111

For the present study of topic and focus marking in Mandan, Wolvengrey's (1991) most valuable contribution is that he demonstrates that the the full range of nominals in Mandan can potentially be marked by =*na*. Upon further inspection of the corpus, I have identified instances of indirect objects that can take =*na* marking, expanding the possible grammatical roles that can be marked by =*na*. There is no mention of indirect objects in Wolvengrey's (1991) study. We can see examples of these diverse set of instances of =*na* on nominals with differing grammatical roles below along with its corresponding θ-role in parentheses.

(11) Examples of =*na* on nominals with differing roles

a. Active subject (agent)

kowóorooreena	*máah íseksoomaksịh*
ko-wooroo=ee=rą	wąąh i-sek=s=oowąk=sịh
3POSS.PERS-husband=DEF=DEM.DIST=TOP	arrow PV.INS-make=DEF=NARR=INTS

'**her husband** made an arrow' (Hollow 1973b:86)

b. Stative subject (experiencer)

súknuma'k shínasheena	*ó'roomako'sh*
suk#ruwą'k shi=rąsh=ee=rą	o'=oowąk=o'sh
child#man be.good=ATT=DEM.DIST=TOP	be=NARR=IND.M

'it was **a nice young man**' (Hollow 1973b:125)

c. Direct object (patient)[10]

Kóoxą'te Míihs	***tasúkseena***	*íratax*ak
kooxą'tE#wịịh=s	**ta-suk=ee=rą**	i-ra-tax=ak
corn#woman=DEF	**AL-child=DEM.DIST=TOP**	PV.INS-INS.MTH-make.loud.noise=DS

[10]This verb 'weep for' is not a transitive verb in English, but *íratax* is transitive in Mandan.

'Corn Woman was crying **for her child**' (Hollow 1973b:112)

d. Indirect object (goal/recipient)

Wáaratookaxi'heena		*"hiré, rapéhini*	
waa-ratoo=ka#xi'h=ee=rą		hire	ra-peh=rį
NOM-be.mature=HAB#be.old=DEM.DIST=TOP		now	2A-announce=SS
raréehto'sh,	*mí'ti*	*nata,"*	*éeheekereroomako'sh*
ra-rEEh=t=o'sh	wį'#ti	rąt=E=∅	ee-hEE=krE=oowąk=o'sh
2A-go.there=IND.M	stone#house	be.in.middle=SV=CONT	PV-say=3PL=NARR=IND.M

'they said **to the old man**, 'now, you should go announce it while in the middle of the village.'" (Hollow 1973a:208)

e. Oblique object of a postposition (instrument)

Rá'puseena	***mí'***	***réxeena***	***ó'hara***	*pá*
ra'-pus=ee=rą	**wį'**	**rex=ee=rą**	**o'hrE**=∅	pa
INS.HEAT-be.spotted=DEM.DIST=TOP	**stone**	**glisten=DEM.DIST=TOP**	**with=CONT**	head
róotkika'ehe				
rootki=ka'ehe				
hit=QUOT				

'Charred-in-Streaks hit her head **with a translucent rock**, it is said' (Kennard 1936:36)

f. Direct reference of a quoted speech

"Manákiniireena,"	*éepeso'sh*
wrą#krį̨=ee=rą	ee-pe=s=o'sh
water#be.stacked=DEM.DIST=TOP	PV-say.1SG=DEF=IND.M

'"**An embankment**," I said' (Kennard 1936:37)

g. Adverbial adjunct (temporal)

Konúuke	*túk,*	***éena***	*háni*	*tashíxteroomako'sh*
ko-rųųkE	tu=ak	**ee=rą**	ha=rį	ta-shi-xtE=oowąk=o'sh
3POSS.PERS-sister	be.some=DS	**DEM.DIST=TOP**	see=SS	AL-be.good-AUG=NARR=IND.M

'he had a sister and she saw him **then** and really liked him' (Hollow 1973b:134)

The data above in (11) reflect that nominals can play any kind of grammatical role while featuring the *=na* enclitic. Thus, the distribution of this morphological marking of prominence is not restricted to core arguments like subject or direct objects; all nominals can bear *=na*. The majority of instances of *=na*, however, tend to be found on active subjects. One reason for this preponderance of *=na* marking on active subjects is that direct quotations in Mandan are often introduced by adding *=na* to whomever said the quotation, followed by the quotation and the verb. We can see examples of this behavior below in (12).

Another common tendency in the corpus is for the one being quoted to be marked with *=na* followed by the quote, but with an elided verb. We see examples of this use of *=na* in (12c) and

(12d) below, where the element marked with *=na* is used to indicate the person who is producing the quoted speech.[11]

(12) Use of *=na* in direct quotations

a. ***Kinúma'kshiseena*** *"ípashahąkt*
ki-ruwą'k#shi=s=ee=rą i-pasha#hąk=t
MID-man#be.good=DEF=DEM.DIST=TOP PV.DIR-north#POS.STND=LOC
náaketaa *máa'ak íwasehki,* *ní'shak* *í'aakahąkt*
rąąkE=taa waa'ąk i-wa-sek=ki r'̃ishak i-aaka#hąk=t
be.sitting=LOC land PV.INS-make=COND 2POSS-pro PV.DIR-south#POS.STND=LOC
írasekto'sh," *éeheka'ehe*
i-ra-sek=t=o'sh ee-hE=ka'ehe
PV.INS-2A-make=POT=IND.M PV-say=QUOT

'It is said that **Royal Chief**[12] said 'if I make land that way to the north, you should make it to the south' (Hollow 1973b:9)

b. *Karóotiki* ***súknuma'kseena*** *"húų, waréehto're,"*
ka=ooti=ki **suk#ruwą'k=s=ee=rą** hųų wa-rEEh=t=o're
PV=EVID=COND **child#man**=DEF=DEM.DIST=TOP yes 1A-go.there=IND.F
éeheeroomako'sh
ee-hEE=oowąk=o'sh
PV-say=NARR=IND.M

'And then the young man said, "yes, I will go."' (Hollow 1973a:234)

c. *Háktek* ***Kinúma'kshiseena*** *"súkinite,* *káare*
ha=kte=ak **ki-ruwą'k#shi=s=ee=rą** suk=rįt=E kaare
PROV=POT=DS **MID-man#be.good=DEF=DEM.DIST=TOP** child=2PL=SV IMP.NEG
ptáhinista! *Kú'hinista!* *Wáa'ąskaharaxi'sh,*
ptEh=rįt=ta ku'h=rįt=ta waa-ąska#hrE=xi=o'sh
run=2PL=IMP.M come.back.here=2PL=IMP.M NEG-be.a.certain.way#CAUS=NEG=IND.M
kotáwaratoore *húuk."*
ko-ta-wa-ratoo=E huu=ak
3POSS.PERS-AL-UNSP-be.mature=SV come.here=COND

'So **Royal Chief** [was like,] "children! Do not run away! Come back! You do not act that way when his uncle comes." (Hollow 1973b:28)

[11]It is the case that more instance of verb elision with direct quotations are present in Hollow's (1973a) re-elicitation of Kennard's (1934) narratives. One possible reason for this is that certain speakers favored the elision in direct quotations, though it is not clear if there is some stylistic choice of when to preserve the quotative verb and when to elide it. I have glossed the elided instances as having a less formalized way of expressing reported speech, using 'was like' versus 'so and so said' to depict this potential difference in style. It is not possible to say conclusively if this dichotomy is entirely accurate, given the absence of L1 speaker judgments.

[12]This figure's name in English is often rendered as 'Old Man Coyote' or 'Trickster' by non-Native sources like Hollow (1973a,b) and Kennard (1934), but Mrs. Mattie Grinnell would always say his name in English as 'Royal Chief.' For this reason, I use the latter when translating his name into English.

d. ***súkmiihseena*** *"waawateerehereki,*
suk#wįįh=s=ee=rą waa-wa-tee#re-hrE=ki
child#woman=DEF=DEM.DIST=TOP some-UNSP-be.dead#2A-CAUS=COND
íshųųhe *ímaare* *ąąwe, íshųųhe* *ąąwe*
i-shųųh=E i-wąą=E ąąwe i-shųųh=E ąąwe
PV.POSS-sinew=SV PV.POSS-body=SV all PV.POSS-sinew INS.HAND-grasp
rushá *makų́'ta"*
ru-shE wą-kų'=ta
1S-give=IMP.M

'**The young woman** [was like,] "if you happen to kill some, all the sinew of the carcass, take all the sinew for me."

In the examples involving *=na* marking so far, we have seen *=na* on a single element in a clause. However, there does not appear to be a firm maximum number of elements that can bear *=na* marking. There are instances where multiple nominal constructions bear *=na* within the same clause within the corpus. One such example has already been seen in (11e) above, which has been reproduced below in (13b). We can see other such examples in (13) below. For each instance of multiple *=na* marking in a single clause, we see some kind of subject with *=na*, plus another nominal element bearing *=na*. The examples in (13a) and (13c) both involve the secondary element with *=na* being coindexed. For (13a), the coindexed element is a parenthetical describing the subject. This examples contrasts with (13c), where the subject itself appears again, as a kind of resumptive element that has been right dislocated at the end of the utterance to reinforce who it is that said the reported speech in question. In (13b), we see a subject that bears *=na*, plus the semantic instrument 'translucent rock.' These examples demonstrate that multiple *=na* marking is permitted on other elements within the clause.

(13) Double *=na* marking

a. Stative subject and parenthetical description of said stative subject

Áa Hąshkana, ***súknuma'keena,*** *xópinini* *wáa'okaraahe*
aa#hąshka=rą **suk#ruwą'k=ee=rą** xoprį=rį waa-o-kraah=E
arm#be.long=TOP child#man=DEM=TOP be.holy=SS NOM-PV.IRR-be.afraid=SV
míkak
wįk=ak
be.none=DS

'**Long Arms**, **a young man**, was holy and had no fear' (Hollow 1973b:151)

b. Active subject and oblique object of a postposition used as an instrument

Rá'puseena ***mí'*** ***réxeena*** ***ó'hara***
ra'-pus=ee=rą **wį'** **rex=ee=rą** **o'hrE=∅**
INS.HEAT-be.spotted=DEM.DIST=TOP stone glisten=DEM.DIST=TOP with=CONT
pá *róotkika'ehe*
pa rootki=ka'ehe
head hit=QUOT

'**Charred-in-Streaks** hit her head **with a translucent rock**, it is said' (Kennard 1936:36)

c. Active subject and resumptive right dislocated active subject

íko'tseena			*"mí'shak,*
i-ko-at=s=ee=rą			w˜-ishak
PV.POSS-3POSS.PERS-father=DEF=DEM.DIST=TOP			1POSS-pro
ą'skarahara'shka	*éewaharani*	*minikímaxani,"*	*éeheeroomako'sh,*
ą'ska#ra-hrE=ą'shka	ee-wa-hrE=rį	w-rį-kiwąxE=rį	ee-hEE=oowąk=o'sh
be.that.way#2A-CAUS=PSBL	PV-1A-CAUS=SS	1A-2S-ask=SS	PV-say=NARR=IND.M

maní'o'na	
wa-rį-o'=rą	
UNSP-2S-be=TOP	

kó'tseena
ko-at=s=ee=rą
3POSS.PERS-father=DEF=DEM.DIST=TOP

'**that father of hers** said, "me, I thought that you were the one who maybe did something so I asked you,' **her father** did.' (Hollow 1973a:238)

The presence of multiple =*na* marking in different sources (i.e., Kennard 1936, Hollow 1970, Hollow 1973a) indicates that multi-=*na* structures are possible across a range of speakers, regardless of generation, and that such constructions are not idiosyncratic.

3.3. Summary

Looking at the contexts within the corpus where =*na* appears, there are two general situations where we see it: when =*na* is encliticized onto old information that is being brought up again or if there is new information that the speaker wishes to bring to the forefront of the listener's attention. Previous analyses of focus in §2.2.3 allow for different kinds of focus to manifest within an utterance, but focus is generally described as a culminative prosodic feature: i.e., there can be a maximum of one focused element within a clause. Therefore, we can eliminate the hypothesis that =*na* is a focus marker.

Under the analysis discussed in §2.2.2, topicalized elements are constituents that are found at the left periphery (or perhaps also at the right periphery) of a clause through the involvement of some syntactic operation: i.e., movement. We have seen examples throughout this section where the nominals bearing =*na* have not undergone any movement to some peripheral position. For example, (13b) features both a subject with =*na* and an oblique with =*na*, where both elements appear in their canonical word order for Mandan. This same pattern applies to (11c), where the subject does not bear any topic marking, but the direct object does. This datum likewise features canonical word order for Mandan. The fact that =*na* does not necessitate the movement of a constituent from its position lower in the syntactic structure into a peripheral position within the clause eliminates =*na* as a marker of a topicalized element.

After eliminating the possibility of =*na* being a focus marker or an indicator of a topicalized element, we are left with the hypothesis that =*na* is a topic marker, as described in §2.2.1. The novelty of a topic is not inherent to whether it will bear =*na*, but the pragmatic choice of the speaker to mark some kind of aboutness-change, a contrast, or a familiar topic. Throughout the text, I have followed the convention from Kasak (2019) to gloss =*na* as TOP for 'topic marker,' and that impression is borne out by the distribution of =*na* in the corpus.

One question that Wolvengrey (1991) raises is what is the pragmatic difference between instances involving just =*na* versus =*na* plus the distal demonstrative =*ee* and the definite article

=*s*? This question appears linked to the overall pragmatic difference in whether overt definiteness marking is present in Mandan or not. While preparing my dissertation (Kasak 2019), I observed that definiteness is not obligatorily marked on Mandan nouns and that definiteness can be implied by context. Given the focus on verbal morphology and not nominal morphology, this tendency was not expressly written down in that work. However, I suggest that the motivation for whether to overtly mark a nominal as being definite or not in Mandan is done for some pragmatic reason. Further investigation of how definiteness is expressed in Mandan is a topic worth futher research but is beyond the scope of the present paper.

4. Prosodic manifestations of topic and focus

Discussions of topic and focus in Siouan have generally relied on textual data rather than audio data. There is nothing inherently problematic about this approach, though ignoring the prosodic aspect of how topic and focus are expressed in these languages misses the opportunity to describe an additional layer of information structure that transcribed data can easily miss.

4.1. Previous work on prosody in Siouan

To my knowledge, Larson's (2009) SCLC presentation on Umo^{n}ho^{n} prosody is among the first instances to attempt to shift attention to suprasegmental aspects of a Siouan language. In his dissertation, Mirzayan (2010) engages in a massive multi-year undertaking to focus on a prosodic analysis of a Siouan language in his study on intonation and prosody in Lakota. This study is groundbreaking in Siouan linguistics, given the fact that so much of what we have historically examined has been done with data transcribed by ear rather than with the assistance of instrumentation. His analysis of Lakota reveals that, while lexical stress is generally associated with higher F0, higher intensitive, and longer duration, there are disassociations between the presence of phrasal pitch and and the pitch typically observed in lexical stress (Mirzayan 2010:3). Prosodic data throughout his work is represented using a ToBI coding protocol per Pierrehumbert (1980), Beckman & Pierrehumbert (1986a), Pierrehumbert & Hirschberg (1990), *inter alios*.

Gordon (2016) likewise examines the interaction between information structure in a set of Siouan languages with their prosody. She provides examples of prosodic data from four Siouan languages (Umo^{n}ho^{n}, Baxoje, Hidatsa, and Hoocąk) that incorporate a ToBI-style analysis of the tunes found in audio samples from these languages. Both Mirzayan (2010) and Gordon (2016) make use of Praat by Boersma & Weenink (2020) to assist in the analysis of pitch and intensity in the data analyzed in their works, rather than relying on trying to analyze the data by ear.

4.2. Prosody of topic and focus in Mandan

The interface between prosody and information structure has been discussed in different Siouan languages to varying degrees of depth. There has been some discussion of suprasegmental features in Mandan in with respect to the interaction between F0 and lexical stress (Kasak 2019:136*ff*), but no discussion beyond the level of the word and into the realm of the phrase or the utterance.

What follows below thus represents the first attempt at examining patterns of prosodic behaviors in Mandan with respect to topics and focus. The phonetic data herein come from two

different Mandan recordings. One recording features Mrs. Mattie Grinnell (1867–1975), relaying the traditional narrative "No Tongue." The other recording is of the narrative "Blackwolf" or "The Gambler," as told by Mr. Edwin Benson (1931–2016). Each narrative is approximately 30 minutes of continuous speech in Mandan. While there are transcriptions of both narratives, I have only analyzed the first 5 minutes of each narrative using Praat.

Rather than go through the data using a ToBI protocol, I examined instances involving *=na* by creating a TextGrid for each token and then running a the Praat script "create-pictures-selected-sound-and-textgrid" by Elvira-García (2017) to create a Praat picture of each token that featured the spectrogram and TextGrid with an F0 curve superimposed over the waveform. This F0 curve serves to illustrate the pitch track of each token to show how the pitch accent manifests on each element within these sentences.

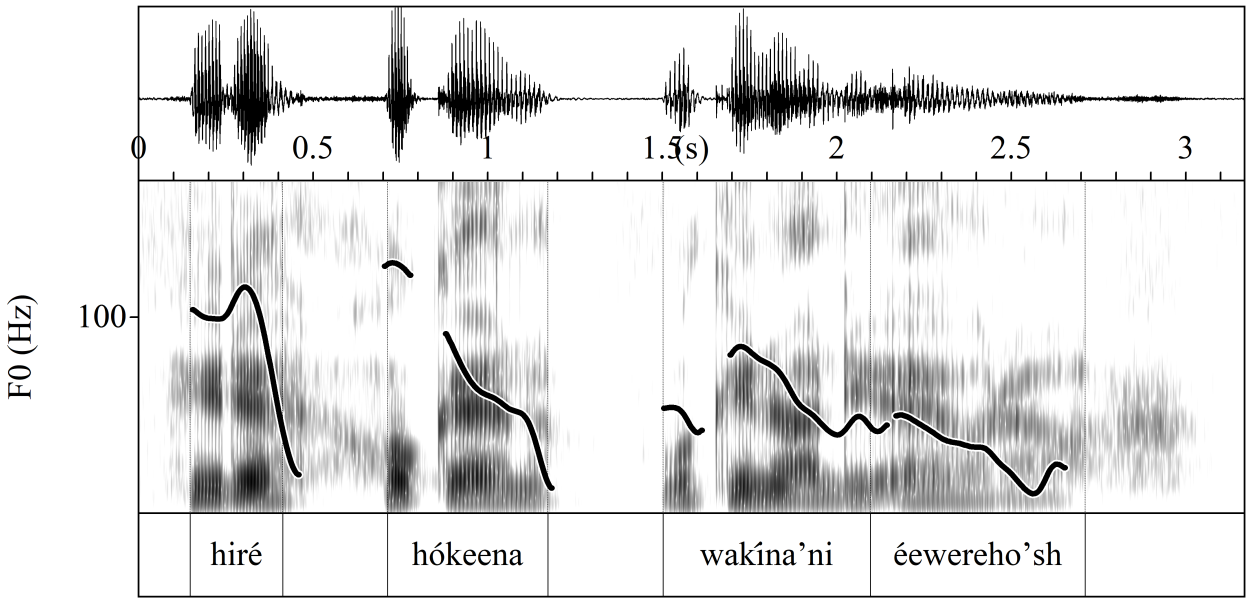
Let us begin by looking at some examples from Mr. Benson's telling of "Blackwolf." In (14) below, we see the direct object *hók* 'story' bear the distal demonstrative *=ee* plus the topic marker *=na*. The direct object is not the leftmost element in the sentence below. The leftmost element, instead, is the adverbial *hiré* 'now.' Figure 1 shows the F0 curve throughout the course of this utterance. We can see that the two initial elements, the adverbial and the direct object, both have a much higher pitch than the verbal complex *wakína'ni éewereho'sh* 'I want to tell.'

(14) "Blackwolf" excerpt 1[13]

Hiré ***hókeena***	*wakína'ni*	*éewereho'sh*
hire **hok=ee=rą**	wa-kirą'=rį	ee-we-reh=o'sh
now **story=DEM.DIST=TOP**	1A-tell=SS	PV-1A-want=IND.M

'Now, I want to tell a story.' (Trechter 2012:11)

Figure 1: Excerpt 1 from "Blackwolf"



[13]This .wav file is available here: https://github.com/ryankasak/SCLC41/blob/main/BW1.wav.

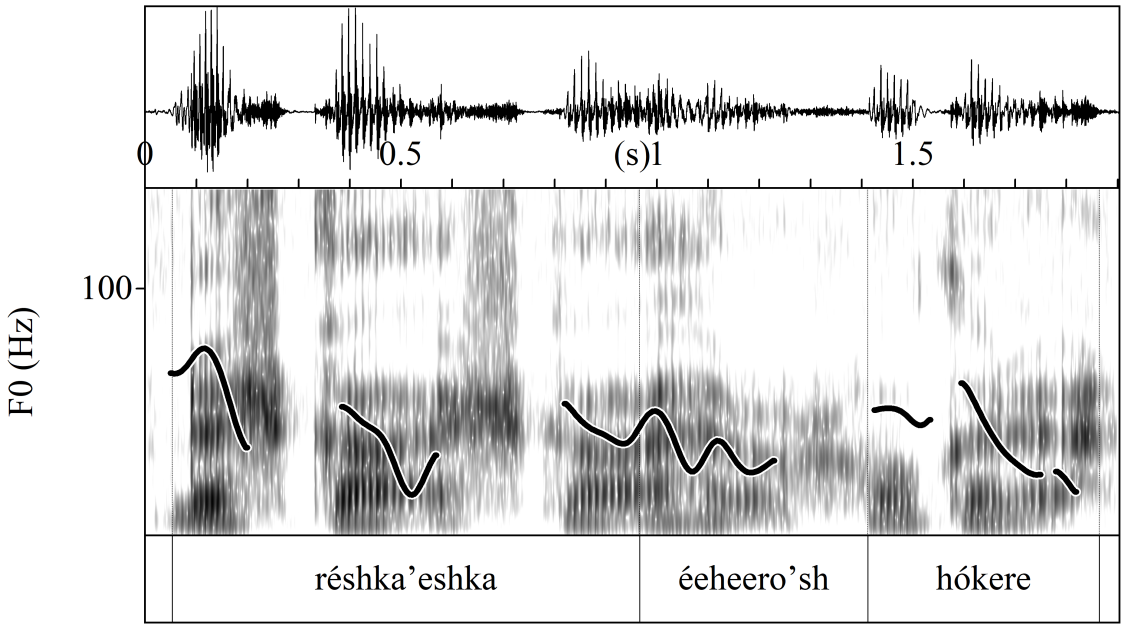
We can compare the data in (14) with what we see in (15) below, which is also taken from "Blackwolf." There is no element bearing *=na*, but we do see a right dislocated direct object, *hókere* 'this story.' This direct object appears to be a postposed familiar topic, and yet it does not bear a topic marker. Typically, we see strong declination in Mandan, where the final word in an utterance has a drastically lower pitch than the initial word. However, *hókere* begins with a pitch that is approximately the same as the preceding verb *éeheero'sh*, which under normal circumstances in an SOV language like Mandan would be the final element in the utterance. Instead, we see an uptick in pitch after the end of *éeheero'sh*, followed by a steep drop-off at the end off the utterance.

(15) "Blackwolf" excerpt 2[14]

Réshka'eshka éeheero'sh, hókere.
reshka-eshka ee-hee=o'sh hok=re
this.way-SIM PV-say=IND.M story=DEM.PROX

'One says it like this, this story.' (Trechter 2012:11)

Figure 2: Excerpt 2 from "Blackwolf"



In (16), we again see an instance where there is an utterance-initial adverbial *máxha* 'one time, once,' followed by a nominal bearing *=na*, *numá'keena* 'a man.' This time, the element bearing the topic marker is the subject. Again, given the fact that Mandan has historically been described as a language with prototypical SOV word order, it is worth noting that subjects are not always the initial element to occur in a sentence. We can see that the F0 curve starts out high utterance-initially, as is expected, but then it drastically increases even farther up to align

[14]This .wav file is available here: https://github.com/ryankasak/SCLC41/blob/main/BW2.wav.

with the lexical stress in the word *numá'keena.* Declination is on full display within the rest of the clause, as the F0 peaks within *numá'keena* 'a man,' which is the subject, and promptly drops down as we get to the final verb of the clause, *ó'rak* 'be.'

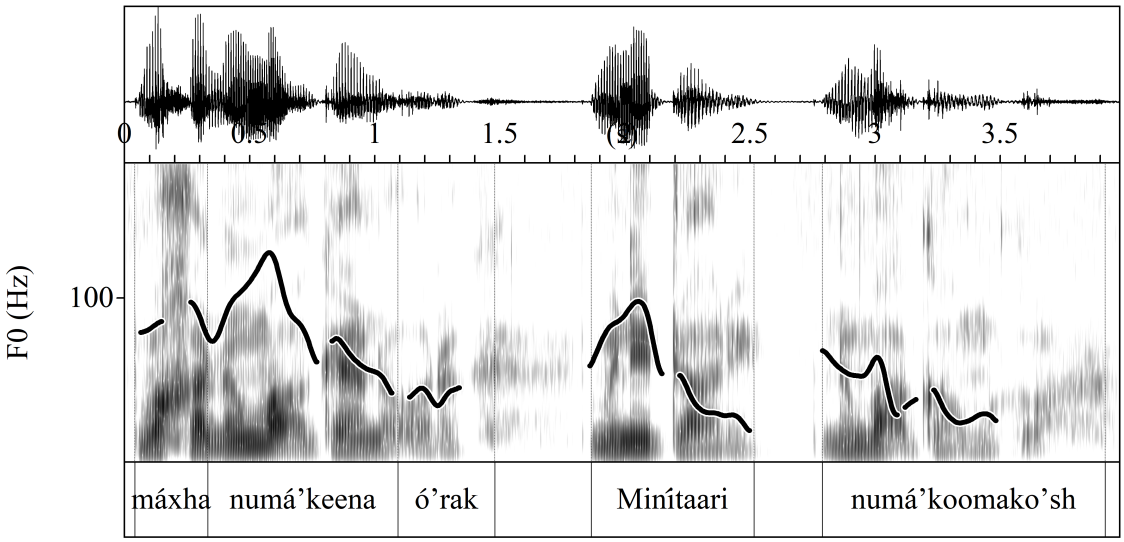
At the onset of the following clause, we see another nominal element, *Minítaari* 'Hidatsa,' feature a high F0 which then decreases and goes back up to a mid tone for the utterance final verb *numá'koomako'sh* 'was a man,' where the declination is even more stark than what we saw in the first clause. For both nominals that appear before a verb, we see a rapid fall from high to low pitch, though the verb *ó'rak* started out with low pitch and ended with low pitch, while the utterance-final verb *numá'koomako'sh* began with a mid pitch before dropping to a low pitch that was even lower than the low pitch in *ó'rak.*

(16) "Blackwolf" excerpt 3[15]

Máxha	***numá'keena***	*ó'rak*	*Minítaari*	*numá'koomako'sh.*
wąx#ha	**ruwą'k=ee=rą**	o'=ak	wrį#taari	ruwą'k=oowąk=o'sh
one#TIMES	**man=DEM.DIST=TOP**	be=DS	water#cross	man=NARR=IND.M

'One time, there was **a man** and he was a Hidatsa man.' (Trechter 2012:11)

Figure 3: Excerpt 3 from "Blackwolf"



The final excerpt from "Blackwolf" appears below in (17). Like (15), there are no elements that bear the *=na* topic marker. However, there is a perceptual prosodic prominence on the verb *kíikini'sįh* 'to really gamble.' When running Elvira-García's (2017) Praat script to analyze this example, there is a disconnect between how the pitch track is interpreted by the basic loadout of Praat in Figure 4 (pitch track shown as a blue line superimposed over the spectrogram versus the intensity shown as a yellow line) versus how the F0 track appears in Figure 5 (i.e., the version produced with the Praat script).

[15]This .wav file is available here: https://github.com/ryankasak/SCLC41/blob/main/BW3.wav.

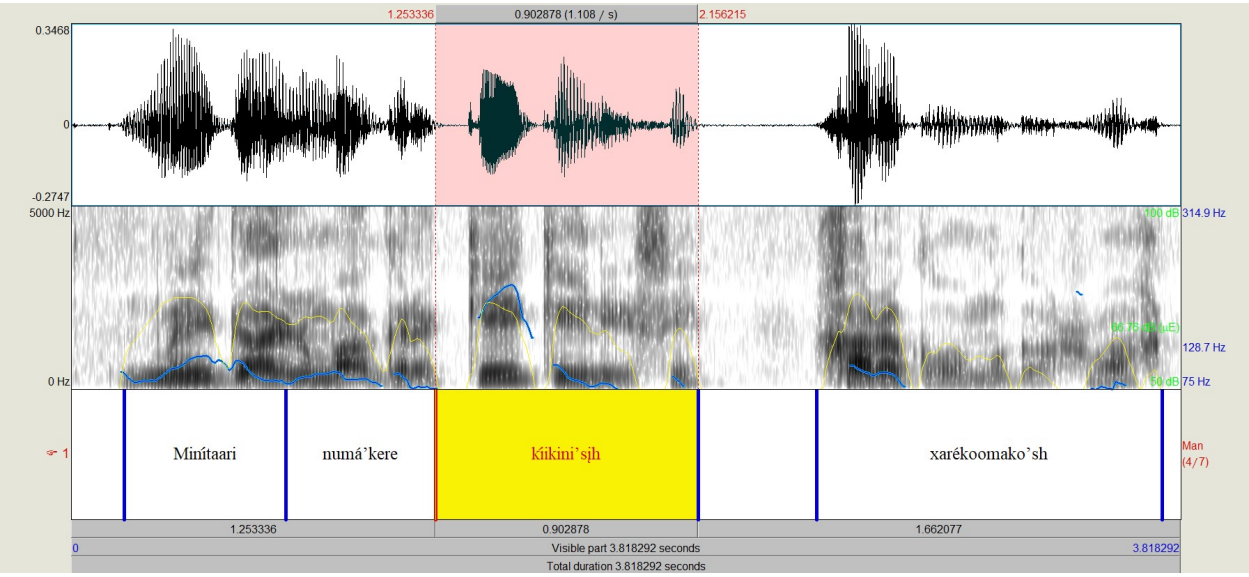
(17) "Blackwolf" excerpt 4[16]

Minítaari	*numá'kere*	*kíikini'sįh*	*xarékoomako'sh*
wrį#taari	ruwą'k=re	kiikrį'=sįh	xarek=oowąk=o'sh
water#cross	man=DEM.PROX	gamble=INTS	be.brave=NARR=IND.M

'The Hidatsa man loved to gamble.' (Trechter 2012:11)

Praat's built-in pitch tracking shows the pitch in *kíikini'sįh* to be drastically higher than in any other word. This pitch is represented in Figure 4 below as a blue line on the waveform. To my ear, this word stands out as being more prosodically prominent than the other words in the sentence, but the F0 analysis undertaken by the Praat script does not confirm Praat's own pitch tracking. The Praat picture in Figure 5 shows *kíikini'sįh* as having noticeably lower pitch than the high pitch on the word *Minítaari*, contrary to what we see above in Figure 4.

Figure 4: Excerpt 4 from "Blackwolf" (without Praat script)



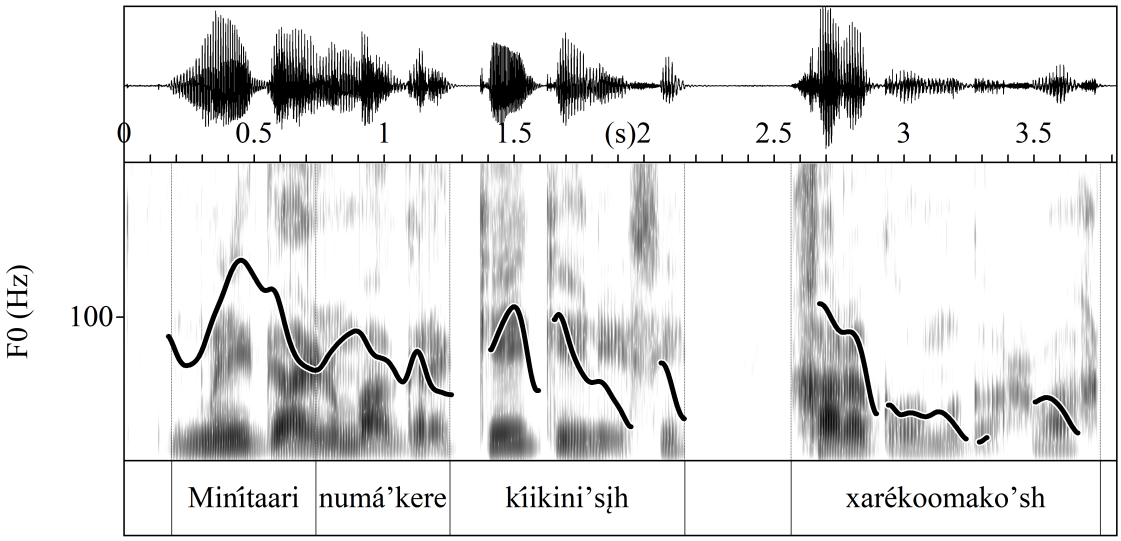
Even if this word does not have the F0 cue that I had initially expected, we do still see the expected high F0 for the first word in the utterance, followed by a very low F0 at the end of the utterance.[17]

The previous four examples come from Mr. Edwin Benson's telling of "Blackwolf," but he is not alone in making use of overt *=na* to mark topics or in employing other prosodic or syntactic strategies to indicate some kind of topic or focus. The other speaker whose Mandan is

[16]This .wav file is available here: https://github.com/ryankasak/SCLC41/blob/main/BW4.wav.

[17]One possible explanation for this mismatch between the pitch track in Praat and the Praat picture derived therefrom could be that the perceived sharp increase and subsequent drop in pitch is so drastic that the "Octave Jump cost" value of 0.35 could be high enough to affect the algorithm's decision about whether a jump in F0 is reasonable and negate the magnitude of the rise in F0 here. Rather than cherry-picking the data, I have elected to maintain the same parameters of this Praat script throughout this paper. Further examination of this overcompensation for such a drastic pitch jump in Mandan prosody warrants future attention that is beyond the scope of the present work.

Figure 5: Excerpt 4 from "Blackwolf" (with Praat script)



discussed here is Mrs. Mattie Grinnell, who was already a centenarian by the time Dr. Robert Hollow interviewed her for his study of the Mandan language. The following example is from her telling of the narrative "No Tongue."

In (18), we can see a stative subject *mí'ti xténa* 'a big village' bearing the topic marker =*na*. In this situation, it is not the noun that bears the topic marker, but the stative verb acting in an adjectival capacity. The presence of the topic marker on the stative verb indicates that the entire noun phrase is being treated as the topic, rather than just an element within the noun phrase being the topic.

(18) "No Tongue" excerpt 1[18]

Mí'ti	***xténa***	*téroomako'sh.*
wį'#ti	**xtE=rą**	tE=oowąk=o'sh
stone#home	**be.big=TOP**	stand=NARR=IND.M

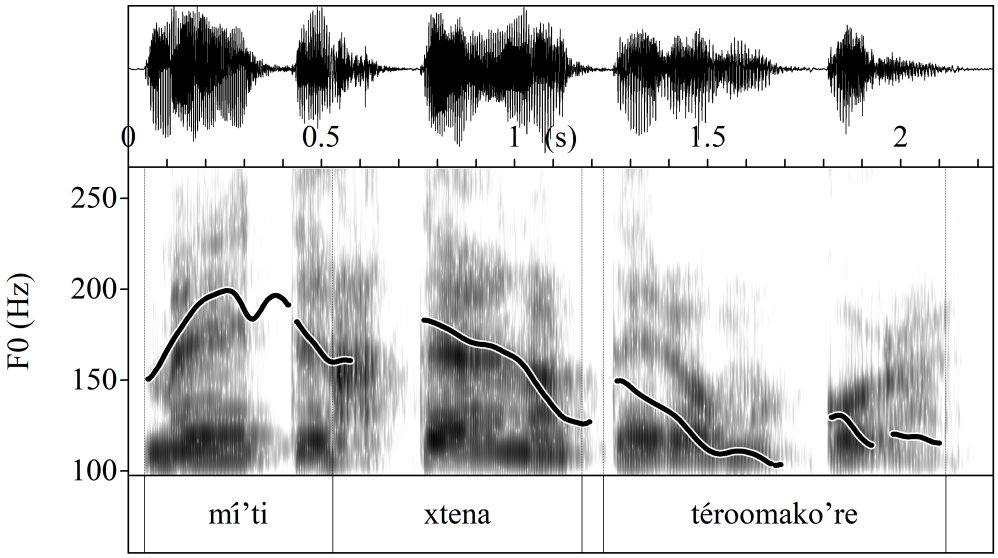
'A big village was there.' (Hollow 1973b:176)

Looking at the F0 track in Figure 6 below, we see that the highest F0 values are on the head of the noun phrase, *mí'ti* 'village,' rather than the adjunct, *xténa* 'big,' to which the topic marker is encliticized. Once again, like we have seen in the examples from Mr. Edwin Benson, there is a sharp declination as Mrs. Mattie Grinnell reaches the end of an utterance. This consistent drastic declination suggests that these drops in pitch in sentence-final environments are not idiosyncratic, but a strong tendency of Mandan prosodic patterns.

The final piece of data analyzed here is also from "No Tongue." In (19), we see the subject *minísweeruteena* 'the dog,' postposed after the verb in a case of right disjunction or tacking it onto the end of the sentence as an afterthought or parenthetical. The F0 peak is highest on the verb *íhekoomako'sh* 'he knew it,' and there is a low F0 plateau after the initial lexical stress within this

[18]This .wav file is available here: https://github.com/ryankasak/SCLC41/blob/main/NT1.wav.

Figure 6: Excerpt 1 from "No Tongue"



word. The F0 rises again on the primary and secondary stresses of the word *minísweeruteena* 'dog' [ˈmⁱnis.ˌweː.ɾu.ˌteː.nã].[19] As we can see for *minísweeruteena* in Figure 7, the F0 falls sharply halfway through the syllable as Mrs. Grinnell ceases phonation and the final syllable becomes almost whispered.

(19) "No Tongue" excerpt 2[20]

Íhekoomako'sh,	***minísweeruteena.***
i-hek=oowąk=o'sh	**wrį́s#wee#rut=ee=rą**
PV.INS-know=NARR=IND.M	**horse#feces#eat=DEM.DIST=TOP**

'He knew it, **the dog**[21] did." (Hollow 1973b:181)

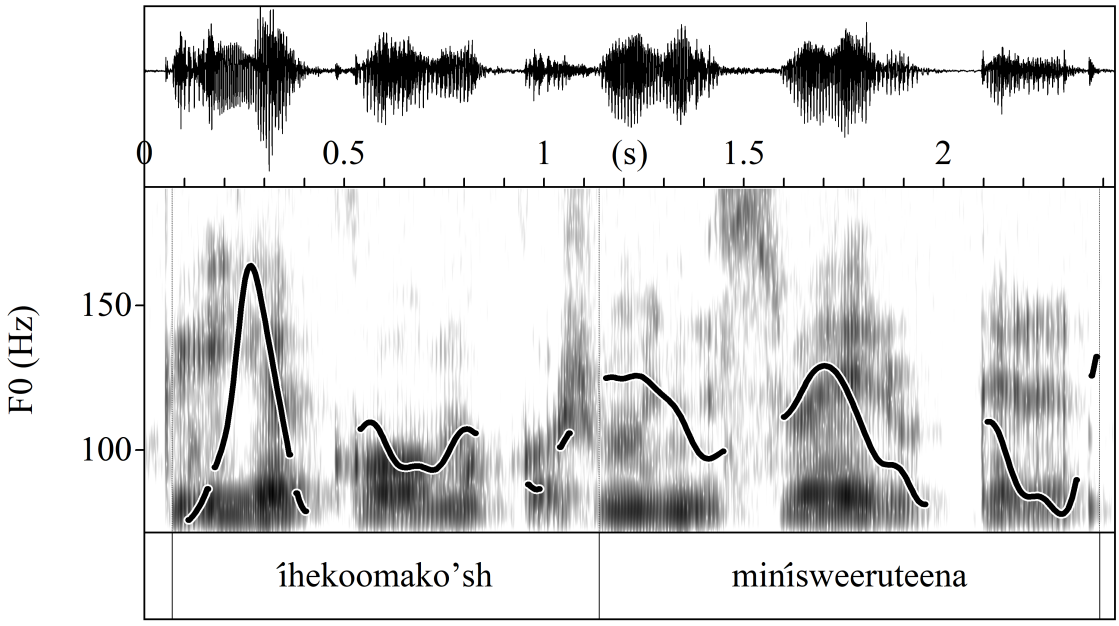
This datum in (19) further illustrates that there is no unique prosodic characteristic for nominal elements bearing *=na*. The expected high left boundary tone is visible in Figure 7 and the peak of this boundary tone in *íhekoomako'sh* 'he knew it' is dramatically higher than any of the pitch peaks within *minísweeruteena* 'the dog.' Again, this lack of prosodic prominence on the element bearing *=na* reinforces the argument herein that this formative is not a focus marker.

[19]The superscript [ⁱ] here indicates an excrescent vowel. As discussed in Kasak (2019:78*ff*), these vowels, often called Dorsey's Law vowels, are a phonetic phenomenon found in consonant clusters where the second element is a sonorant. These vowels are merely an extension of the following vowel, which is the real nucleus of the syllable. While these vowels are generally perceptible by Mandan speakers as can be determined from home orthographies, they are extraphonological and are really tautosyllabic with the following vowel. See Hall (2006) for further discussion on the phonological treatment of excrescent vowels across the world's languages.

[20]This .wav file is available here: https://github.com/ryankasak/SCLC41/blob/main/NT2.wav.

[21]The Mandan word for dog literally means 'horse that eats feces.' Many other Siouan languages likewise have some lexical connection between dogs and horses: e.g., Lakota *šúŋkawakȟáŋ* 'horse' (lit. 'sacred dog').

Figure 7: Excerpt 2 from “No Tongue”



4.3. Summary

Previous descriptions of Siouan grammars have rarely addressed how prosody is treated or the nuances of how the information structure of a language is treated by the prosody or the syntax of said language. The most extensive treatment of prosody and intonation is found in Mirzayan’s (2010) dissertation on these aspects of Lakota grammar. Other discussions of information structure have largely been relegated to describing various morphological formatives as focus markers, topic markers, or topicalization markers with little to no description of the conditions under which a speaker uses them or whether there are situations where their use would be obligatory or even proscribed.

All previous studies on Mandan that involve instrumentation like in Kasak & Lundquist (2019) or Kasak (2019) focus on word-level phonological phenomena rather than the interface of pragmatics and prosody. The six examples of Mandan above focus on looking at the behavior of F0 as an attempt to investigate whether there are certain prosodic markers to indicate focus or topics. This section has involved description of what the data show, rather than the implications of the data. A discussion of these findings appears in §5 below.

5. Discussion of prominence marking

Six different excerpts from the first five minutes of two Mandan narratives have been examined above in §4. Several patterns regarding the prosodic treatment of prominence marking with respect to information structure emerge. First and foremost, there is a strong tendency to move topicalized elements to the left periphery of an utterance, as well as to shift an element that acts as a topic of reminding the listener about a fact to the right edge of an utterance. These topics did not necessarily bear the topic marker =*na*, either.

When we look at the kinds of topics that appear in the left periphery, we see several patterns as they pertain to the prosody of Mandan. Firstly, adverbial and nominal elements in the left periphery of an utterance have much higher F0 peaks than the matrix verb or the matrix verb plus a subordinated verb. In cases where there are no adverbial or nominal elements that precede the verb, as we see in (19), the verb takes on a high F0 peak that is comparable to the non-verbal elements heretofore discussed. We can therefore surmise that there must be some kind of left boundary tone for certain intonations in Mandan that are causing this tendency to have very high F0 at the onset of an utterance. This boundary tone seems to be associated with intonational phrases rather than the overall utterance itself, as we have two clauses in a single utterance in (16), and we see the expected high F0 peaks on the adverbial and nominal elements in the left periphery of the first clause, as well as a high F0 peak again on the nominal in the left periphery of the second clause.[22]

When multiple elements appear in the left periphery, like adverbials or topicalized nominals, both elements have similar high F0 peaks. This F0 behavior reinforces the notion that both of these elements are kinds of topicalized elements, since both *máxha* 'once' and *numá'keena* 'a man' in (16) with Figure 3 experience high F0 peaks while the verb *ó'rak* 'be' has the expected low F0 for the end of a clause or utterance. We see a version of this F0 distribution in (14) with Figure 1, where again we have multiple elements in the left periphery: *hiré* 'now' and *hókeena* 'a story.' The adverbials discussed here each have their own intonational phrase, which accounts for why they have such high F0 values when compared to following verbs. However, when compared to the nominal that follows these adverbials, the nominal elements bearing the topic marker *=na* both appear to have marginally higher peak F0 values than the adverbials do.

Elements that have been shifted to the right periphery have a similar F0 behavior, such as the postposed direct object *hókere* 'this story' in (15) with Figure 2 and the postposed subject *minísweeruteena* 'the dog' in (19) with Figure 7. In both of these examples, there is a nominal element that follows what should be an utterance-final verb, given the fact that Mandan has a default SOV word order. We have seen in §4.2 that there is a strong process of declination in Mandan, and that utterance-final verbs universally have very low F0 throughout. In these two instances, however, there is a nominal element that serves as a reminder of a familiar topic that appears in utterance-final position. Contrary to the strong drop in F0 we have seen on verbs, we see a resumption of high F0 values that overlap with the lexical stress of the postposed word. There is no discernable difference in the manifestation of F0 for a word bearing the topic marker *=na* like in (19) versus the word that lacks the *=na* marker in (15).

One of the examples examined in §4.2 has an utterance-initial nominal that appears with the topic marker *=na*: (18) in Figure 6. The subject *mí'ti xténa* 'a big village' exhibits the expected high F0 peak, though not on the word bearing the *=na*, but on the head of the overall noun phrase, *mí'ti* 'village.' There does not seem to be a correlation between *=na* and whether a lexical item takes a focus intonation. Likewise, it is not obvious whether a construction like the one in (18) is prosodically different from a sentence that begins with a nominal lacking the topic marker *=na*.

[22] I adopt the notion of the prosodic hierarchy per Nespor & Vogel (1986), Beckman & Pierrehumbert (1986b), Selkirk (1986, 2011), *inter alios*, where the domain of the utterance (Utt) can consist of one or more intonational phrases (ιP), which in turn are made up of phonological phrases (φP), which contain prosodic words (ω), which are made up of feet (Ft), which are divided up into syllables (σ), which can be decomposed into morae (μ). Different phonological phenomena can have a sensitivity to one or more kinds of domain boundaries within this prosodic hierarchy. See the aforementioned authors for a more thorough explanation of this notion.

This topic is introducing new information into the discourse, which makes it categorizable as a kind of aboutness-shift topic as discussed in §2.2.1. There does not seem to be a difference in the prosody of an aboutness-shift topic versus any other element that happens to be the first element of an utterance or intonational phrase.

The sole instance of a possible focused element in the data analyzed for this study is found in (17), represented in Figures 4 and 5. While going through the audio for this narrative, I was immediately struck by how much different the word *kíikini'sįh* 'to really gamble' was to my ears and how it initially looked to have a drastically higher F0 than anything else I have observed in going through the data for either narrative. However, there is conflicting information being given by Praat itself and the Praat script used in this study. As mentioned above in §4.2 the very fact that there was such an extreme octave jump involved in this word could actually be nullified by the algorithm used by the Praat script to track the course of the F0 over the course of the token. The monumentally high formant values that shot up and quickly came back down on the first syllable of the word *kíikini'sįh* could have been treated as a set of outliers and smoothed out by the algorithm of the script.

6. Conclusion

Throughout the work presented here, I have had three over-arching goals stated in (1). The first of these goals was to examine previous literature on topic- and focus-marking in Mandan. I have done so by looking at how the issue of topic- and focus-marking has been treated in other Siouan languages in §2 and then looking at what those who have worked on Mandan have said about this question in §3. Overall, the only discussion of information structure in Mandan has revolved around the mention of the formative *=na*. This formative is referred to by different nomenclature by different authors, but I establish here that it must truly be a topic marker and not a focus marker or a topicalization marker, given the fact that there are no pitch patterns that would equate to focus intonation and that the topic marker can appear on elements that are not in the left periphery of a clause (i.e., they have not been topicalized). We do not consistently see high or some other kind of pitch contour associated with any element bearing *=na* in the data presented here. Therefore, we are left with the conclusion that *=na* is definitely a topic marker, as we have excluded the other two possibilities.

The secondary goal of the work herein is to compare and contrast previous literature on topic- and focus-marking in Mandan with actual audio recordings of L1 Mandan speakers to evaluate the veracity of previous claims as to the status of the topic marker *=na*. Using Praat, I employed five instances that involve topics and one that possibly involves focus. The Praat pictures generated by Elvira-García's (2017) Praat script yielded readily analyzable figures that show the behavior of F0 with respect to words in the periphery of a sentence, both left and right.[23] These figures illustrate the fact that elements bearing *=na* are associated with higher F0 peaks, even when appearing utterance-finally when declination would otherwise depress the F0 values as phonation began to cease. Clearly, some kind of prosodic prominence is being bestowed upon elements bearing *=na*, but there is not sufficient evidence from this study to demonstrate that there are unique tunes for different kinds of topics in Mandan.

[23] An earlier version of this script was also employed in the analysis of Hidatsa not having lexical pitch accent in Boyle et al. (2016).

In Frascarelli & Hinterhölzl (2007), the authors note that both Italian and German have different tunes for different topics, and as such, it is worth exploring the Mandan corpus further to see if similar observations can be made. These authors argue further that the three different intonational contours are reflective of three different syntactic positions occupied in the left periphery by these topic. More work is needed on Mandan to determine if the same conclusion can be drawn about it. There is a marginally higher peak F0 in constructions involving an adverbial that is followed by a nominal bearing the topic marker *=na*, but this study is not conclusive in its findings about whether there is a generalized H*+L topic intonation used across the different topic types. Furthermore, it is not altogether obvious that the adverbials that we have seen in the left periphery in §4.2 are not themselves a kind of topic. If this is the case, then perhaps the temporal adverbials act as contrastive topics while the topics bearing *=na* that follow them are aboutness-shift topics. The right dislocated topics discussed here all are familiar topics. If this analysis holds, then we have at least some empirical basis to state that there is some order in which topics can co-occur, where contrast topics must precede aboutness-shift topics. It is not clear yet how familiar topics fit into this ordering, and further work is needed to resolve this issue.

The final goal of the work here as laid out in 1 is to provide insight into how an understanding of this aspect of Mandan prosody can be used in Mandan language instruction. The loss of Mr. Edwin Benson in 2016 means that there are no longer any remaining L1 speakers of this language. There are still those on Fort Berthold who are working to preserve and promote the Mandan language, and while learning vocabulary and figuring out verb paradigms are all vital components of the language learing process, learning to put words together in a "Mandan way," having prosody that reflects speech that *sounds* like how an L1 Mandan speaker would say it is likewise one of the many goals towards fluency. This need to acquire the prosody of an L2 is no different for those learning any other language. More work is needed on this issue, but what I have discovered in this study is that Mandan syntax and information structure are deeply intertwined and that word order is not as rigid as described elsewhere (Mixco 1997:46). Mandan is far more flexible in its word order than previously thought, and additional work is needed to see if the generalizations uncovered here are borne out throughout the corpus in a meaningful way, or if there are other patterns of topic-marking yet to be described.

Some future avenues of research on this matter will undoubtedly need to involve a larger amount of analyzed audio. This process is already underway for the narratives "Blackwolf" and "No Tongue," but I have many hours of Mandan speech that needs to be transcribed or retranscribed plus annotated in Praat. There are also, no doubt, countless hours of recordings in the possession of the Nueta Language Initiative in Twin Buttes or archived at Nueta Hidatsa Sahnish College in New Town, plus recordings of spoken or sung Mandan in the possession of family members in and around Fort Berthold. Other issues that have yet to be addressed involve other information structure morphology attested in Mandan, such as the formative *=nu*, which Mixco (1997:42) glosses as 'the aforementioned' or 'the former'.

(20) Use of 'the aforementioned' in the corpus

róo	*numá'kaakikereseena*	***súkeenus***
roo	ruwą'k-aaki=krE=s=ee=rą	**suk=ee=rų=s**
DEM.MID	person-COLL=3PL=DEF=DEM.DIST=TOP	**child=DEM.DIST=ANF=DEF**

tanúma'kshikereroomako'sh
ta-ruwą'k#shi=krE=oowąk=o'sh
AL-man#be.good=3PL=NARR=IND.M

'these people had **the aforementioned child** as their chief' (Hollow 1973b:209)

The *=nu* enclitic is very rare in the corpus, so it is not obvious if there is some stylistic choice for when to use it versus when to treat an element as a familiar topic. Further work is need to compare the prosody of elements bearing *=nu* versus those bearing *=na* or even topics that have no morphological topic marking. In particular, more attention is needed on the ordering of enclitics in constructions featuring *=nu*. The definite marker *=s* is almost universally found preceding the distal demonstrative *=ee* when *=nu* is present, but we see an inversion of this order in the example above in (20).

The unfortunate truth about the work presented here is that each "answer" I have for some issue I investigate, multiple additional avenues of inquiry spring forth. Ultimately, this paper has aimed to address this question: "what can we tell about the prosody of a language that has no L1 speakers using only older audio recordings?" To that question, I can provide the pithy response: "quite a bit." The data examined here were sufficient to formulate an understanding of the purpose of the formative *=na* in Mandan. The audio recordings, even the limited analysis I have been able to accomplish, reveal that there are likely intonational patterns as they relate to multiple topics present in a sentence versus non-topic subjects or direct objects. Finally, I find that the correlates of focus marking in Mandan are not clear and may be marked by more than mere pitch, and that additional analysis must happen that perhaps takes intensity and other cues into consideration. This study was never meant to provide a conclusive, all-encompassing answer to the issue of topic and focus in Mandan, but it does demonstrate that the well is deeper than we originally thought and that there is more about prosody and the interaction between pragmatics and morphology in Mandan to be described.

References

Baker, Paul & Sibonile Ellece. 2011. *Key terms in discourse analysis*. New York, NY: Continuum.

Beckman, Mary & Janet B Pierrehumbert. 1986a. Japanese prosodic phrasing and intonation synthesis. In Alan W. Biermann (ed.), *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, 136–144. New York, NY: Association for Computational Linguistics.

Beckman, Mary E. & Janet B. Pierrehumbert. 1986b. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255–309.

Beeching, Kate & Ulrich Detges. 2014. Introduction. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change* (Studies in Pragmatics 12), 1–23.

Boas, Franz & Ella Deloria. 1941. *Dakota grammar*. Washington, DC: U.S. Government Printing Office.

Boersma, Paul & David Weenink. 2020. *Praat: Doing phonetics by computer*. [Computer program]. Version 6.1.27, retrieved 13 October 2020 from http://praat.org.

Boyle, John P. 2007. *Hidatsa morpho-syntax and clause structure*. Chicago, IL: University of Chicago dissertation.

Boyle, John P., Ryan Kasak, Sarah Lundquist, Armik Mirzayan, Jonnia Torres & Brittany Williams. 2016. *A preliminary study on accentuation in Hidatsa*. Paper presented at the annual meeting of Society for the Studies of Indigenous Languages of the Americas at the Linguistic Society of America annual meeting. Washington, DC.

Coberly, Mary. 1979. A text analysis and brief grammatical sketch based on 'Trickster challenges the Buffalo': A Mandan text collected by Edward Kennard. *Colorado Research in Linguistics* 8. 19–94.

Collette, Vincent. 2019. Nakoda "intensifier" *-ȟ*. In Ryan M. Kasak (ed.), *Proceedings of the 38th Siouan and Caddoan Languages Conference*, 79–97. Chicago, IL: Northeastern Illinois University Linguistics Department.

Dorsey, James Owen & John R. Swanton. 1912. *A dictionary of the Biloxi and Ofo languages* (Smithsonian Institution, Bureau of American Ethnology, Bulletin 47). Washington, DC: United States Government Printing Office.

Elvira-García, Wendy. 2017. *Create pictures with tiers*. [Praat script]. Version 4.4, retrieved 15 May 2017 from https://github.com/wendyelviragarcia/create_pictures.

Fant, Lars, Inge Bartning & Rakel Österberg. 2021. The left and right periphery in native and non-native speech – A comparative study between French L1/L2, Spanish L1/L2 and Swedish L1. *International Review of Applied Linguistics in Language Teaching* 59(1). 87–120.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*: An interface approach to the linking of (null) pronouns. *Natural Language & Linguistic Theory* 25(4). 691–734.

Frascarelli, Mara. 2012. The interpretation of discourse categories: Cartography for a crash-proof syntax. In Valentina Bianchi & Cristiano Chesi (eds.), *Enjoy linguistics! Papers offered to Luigi Rizzi on the occassion of his 60th birthday*, 180–191. Siena, Italy: Centro Interdipartimentale de Studi Cognitivi sul Linguaggio (CISCL) Press.

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form: Generalizations across languages*, 87–116.

Gordon, Binah T. 2016. Information-structural variations in Siouan languages. In Catherine Rudin & Binah T. Gordon (eds.), *Advances in the study of Siouan languages and linguistics* (Studies in Diversity Linguistics 10), 393–423. Berlin, Germany: Language Science Press.

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23(3). 387–429.

Helmbrecht, Johannes & Christian Lehmann (eds.). 2006. *Hocąk–English/English–Hocąk learner's dictionary*. 2nd edn. Erfurt, Germany: Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt.

Hirst, Daniel & Albert Di Cristo (eds.). 1998. *Intonation systems: A survey of twenty languages*. Cambridge, UK: Cambridge University Press.

Hollow, Robert C. 1970. *A Mandan dictionary*. Berkeley, CA: University of California, Berkeley dissertation.

Hollow, Robert C. 1973a. Mandan texts. Box 5: Robert C. Hollow materials. (A set of Edward Kennard's (1933-1934) 22 texts re-elicited, transcribed, and translated by Robert C. Hollow held at the North Dakota State Historical Society.)

Hollow, Robert C. 1973b. Mandan texts. Box 3: Robert C. Hollow materials. (A set of 22 texts recorded, transcribed, and translated by Robert C. Hollow held at the North Dakota State Historical Society.)

Ingham, Bruce. 2003. *Lakota*. Munich, Germany: LINCOM Europa.

Kasak, Ryan M. 2019. *Affix ordering and templatic morphology in Mandan*. New Haven, CT: Yale University dissertation.

Kasak, Ryan M. & Sarah Lundquist. 2019. Nasal harmony in Hoocąk and Mandan. In Ryan M. Kasak (ed.), *Proceedings of the 38th Siouan and Caddoan Languages Conference*, 99–122. Chicago, IL: Northeastern Illinois University Linguistics Department.

Kaufman, David V. 2008. Focality and topicality marking in Biloxi. *Kansas Working Papers in Linguistics* 30. 150–158.

Kennard, Edward. 1934. *Mandan folkloristic texts*. (Collection of 28 narratives elicited and transcribed by Edward Kennard held at the American Philosophical Society.)

Kennard, Edward. 1936. Mandan grammar. *International Journal of American Linguistics* 9(1). 1–43.

Ladd, D. Robert. 1984. English compound stress. In Dafydd Gibbon & Helmut Richter (eds.), *Intonation, accent and rhythm: Studies in discourse phonology*, 253–266. New York, NY: Walter de Gruyter.

Ladd, D. Robert. 2008. *Intonational phonology*. 2nd edn. Cambridge, UK: Cambridge University Press.

Larson, Rory. 2009. *Pronunciation and prosody in Omaha*. Paper presented at the 26th Siouan and Caddoan Languages Conference. Lincon, NE: University of Nebraska–Lincoln.

Lee, Yong-cheol. 2015. *Prosodic focus within and across languages*. Philadelphia, PA: University of Pennsylvania dissertation.

Lyons, John. 1982. Deixis and subjectivity: *loquor, ergo sum?* In Robert J. Jarvella & Wolfgang Klein (eds.), *Speech, place, and action: Studies in deixis and related topics*, 101–124. New York, NY: John Wiley & Sons.

May, Lillian, Krista Byers-Heilein, Judit Gervain & Janet F. Werker. 2011. Language and the newborn brian: Does prenatal language experience shape the neonate neural response to speech? *Frontiers in Psychology* 2(222). 1–9.

Mirzayan, Armik. 2010. *Lakota intonation and prosody*. Boulder, CO: University of Colorado, Boulder dissertation.

Mixco, Mauricio. 1997. *Mandan*. Munich, Germany: LINCOM Europa.

Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht, Netherlands: Foris.

Oliverio, Giulia R. M. 1996. *A grammar of Tutelo*. Lawrence, KS: University of Kansas dissertation.

Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation*. Cambridge, MA: Massachussets Institute of Technology 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 communications*, 271–311. Cambridge, MA: MIT Press.

Ramus, Franck. 2002. Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition* 2(1). 85–115.

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

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27(1). 53–94.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar: Handbook of Generative Syntax*, 281–337. Dordrecht, Netherlands: Kluwer.

Rizzi, Luigi. 2001. On the position "Int(errogative)" in the left periphery of the clause. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi* (North-Holland Linguistic Series: Linguistics Variations 59), 287–296. New York, NY: Elsevier.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT dissertation.

Ross, John Robert. 1986. *Infinite syntax!* Norwood, NJ: Abelex Publishing.

Selkirk, Elizabeth O. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3. 371–405.

Selkirk, Elizabeth O. 2011. The syntax-phonology interface. In John Goldsmith, Jason Riggle & Alan C. Yu (eds.), *The handbook of phonological theory*, 2nd edn., 435–484. Hoboken, NJ: Wiley Blackwell.

Traugott, Elizabeth C. & Richard B. Dasher. 2002. *Regularity in semantic change*. Cambridge, UK: Cambridge University Press.

Trechter, Sarah. 2012. In the words of our ancestors: The Mandan language and oral traditions preservation project. (Transcriptions that accompany two DVDs of Edwin Benson relaying traditional Mandan narratives.)

Wolvengrey, Arok. 1991. A marker of focus in Mandan discourse. In Frances Ingemann (ed.), *The proceedings of the 1990 Mid-America Linguistics Conference and Conference of Siouan-Caddoan Languages*, 583–598. Lawrence, KS: University of Kansas.

Xu, Yi, Ching X. Xu & Xuejing Sun. 2005. On the temporal domain of focus. In *Proceedings of the International Conference on Speech Prosody 2004*, 81–84. Nara, Japan: ISCA.

Yang, Chunsheng. 2016. *The acquisition of L2 Mandarin prosody* (Bilingual Processing and Acquisition 1). Philadelphia, PA: John Benjamins.

Relative clauses in Omaha-Ponca (Umo^{n}ho^{n})*

Julie Marsault
Université Sorbonne Nouvelle

Catherine Rudin
Wayne State College

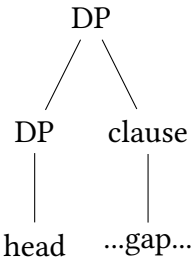
Abstract: Relative clauses in Omaha-Ponca (Umo^{n}ho^{n}) are internally headed. They are identical in form to non-relative clauses except for the presence of a clause-final article which serves to nominalize (or relativize) the clause. The head of the relative clause can have any function and can be any type of nominal (noun, pronominal, or null). The clause-final article also has other functions, so the distinction between relative clauses and other types of clauses or nominal phrases is not always sharp.

Keywords: relative clause, internally headed relative clause, definite article, relativizer

1. Introduction

Relative clauses (RCs) in Umo^{n}ho^{n} are internally headed; that is, the head noun is part of the clause, occupying the same syntactic position it would hold in a "normal", non-relative clause, instead of being a sister to a clause which contains a coreferential gap. Compare the externally headed relative clause in (1) with the internally headed one in (2). The rough tree diagrams below each example correspond to the two types. (Throughout the paper the head noun of each RC is underlined, the associated article is double-underlined, and the verb stem is boldfaced.)

(1) Externally headed relative clause (English)
*[DP [DP the <u>book</u>] [CP (that/which) I **read** ___]]*

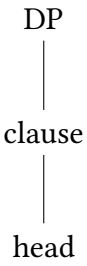


(2) Internally-headed relative clause (Umo^{n}ho^{n})

*[DP [CP <u>tanúka</u> **thizé ithá**=i <u>k^{h}e</u>]]*
meat get promise.PX ART.HORIZ

'the meat that he promised to get'

*Sincere thanks to Johannes Helmbrecht for comments on an earlier version of this article.



Internally headed relative clauses (IHRCs) are found in most if not all Siouan languages, as documented in a number of works: Lakhota (Williamson 1987, Ullrich & Black Bear 2016, Ullrich 2018), Osage (Quintero 2004), Hoocąk (Helmbrecht in progress), Hidatsa (Boyle 2016) and Crow (Graczyk 1991). Not all scholars agree that all RCs in Siouan languages are IHRCs, but IHRCs are certainly the norm for Siouan languages. Umo^{n}ho^{n} is thus typical of Siouan languages in having IHRCs. Umo^{n}ho^{n}, along with at least some other Dhegiha languages, differs from other branches of Siouan in the specifics of the construction, including especially the fact that there is no dedicated relativizer morpheme in the language: instead RCs are simply nominalized clauses marked (usually) with one of the many definite articles.

In this paper we will first demonstrate, in section 2, that Umo^{n}ho^{n} relative clauses are in fact internally headed. This section also covers the characteristic features of IHRCs in this language and crosslinguistically, including the fact that the head is always indefinite: the accompanying article is associated with the entire IHRC, not the head. Section 3 digs deeper into the range of different types of IHRCs in Umo^{n}ho^{n}, showing that (1) RC heads can be any nominal function, that is, any type of argument or adjunct; (2) RC heads can also be any nominal form: overt noun, null or pronominal; and (3) the typical relativizing article can be absent. In section 4 we briefly mention some issues for future research.

2. Characteristics of RCs in Umo^{n}ho^{n}

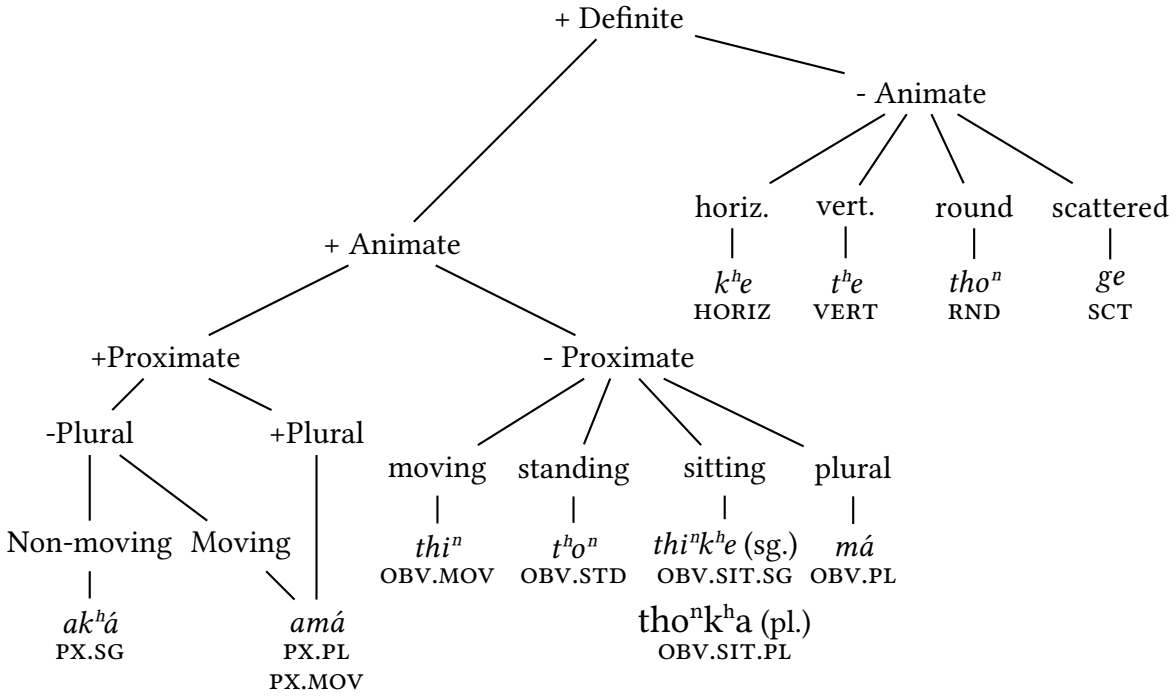
Umo^{n}ho^{n} IHRCs have been described by Koontz (1984), Rudin (1991), Rudin & Shea (2005), as well as Marsault (2021), which is the source of much of this paper. In the earliest of these, Koontz (1984:171) writes: "the technique for forming relative clauses involves substituting for the head [noun] in the matrix clause the entire modifying clause, with its own version of the head [noun] intact". This corresponds to the definition of internally headed relative clauses that we find in Creissels (2006:244), for example. Later works have all agreed with this basic description. Reduced to its simplest form, an Umo^{n}ho^{n} RC is as in (3) - a complete clause, normally followed by one of the group of words usually called articles. The article acts as a nominalizer, turning the CP (complementizer phrase) into a DP. It functions in a similar way to relativizers in other languages. We will sometimes refer to this item as a "relativizer" for convenience.

(3) [[clause] article]

The status of the articles, their meaning, usage, and whether in fact they are articles or some other type of particle, is a complex and much-discussed topic in Dhegiha languages. (See for instance Eschenberg 2005.) A discussion of the article system would take us too far afield here. For purposes of this paper just notice in Figure 8 that the articles carry information not only of definiteness, but also animacy, position, obviation, and other features.

Figure 8: The article system of Umó^{n}ho^{n}

(adapted from Koontz 1984 and Eschenberg 2005)



Looking in more detail at the Umo^{n}ho^{n} relative clause given in (2) and repeated in its full-sentence context in (4a), we see that the RC (bracketed) serves as the object of the main verb, *thizé* 'get'. The head noun, *tanúka* 'meat' is in the position it would normally occupy if "he promised he would get meat" was an independent clause. The only difference in form from the corresponding independent clause in (4b) is that the noun cannot be followed directly by an article; instead, an article, k^he in this case, marks the end of the clause. The demonstrative phrase *thé* k^he 'this thing' is an appositive; appositive phrases commonly occur with both IHRCs and other types of nominals in Umo^{n}ho^{n}.

(4) a. *Thé* k^he {*tanúka* ***thizé*** ***ithá-i*** $\underline{\underline{k^he}}$} *thizá=i* t^he.
this ART.HORIZ {meat he.gets.it he.promised-PX ART.HORIZ} he.got-it=PX EVID
'He got the (piece of) meat (that) he promised he would get.' (More literally: 'He got this thing, the meat he promised he would get.') (Rudin et al. 1989-1992 / speaker : Mary Clay)

b. *Tanúka* k^he *thizé ithá-i.*
meat ART.HORIZ he.gets.it he.promised-PX
'He promised to get the (piece of) meat.' (constructed example)

The clause-initial position of the head might be taken to suggest that it is separate from the clause, i.e. that this is actually an externally-headed RC like the English example in (1). There are two reasons to believe that this is incorrect and that Umo^{n}ho^{n} RCs are in fact internally headed. The first of these is that initial position is not absolute. Although the head noun almost always appears at the beginning of the RC, it can be preceded by another element. Rudin (1991) pro-

vides example (5) from elicitation (Rudin et al. 1989-92); both sentences in (5) were provided as a translation of "The boy wants the man who lives here to leave". Clifford Wolfe proposed (5a) first, and then (5b) when asked to repeat. In (5a), the head noun is preceded by the adverbial adjunct *théthudi* 'here', and we see once again that the RC has the same structure as an independent clause, with considerable word-order freedom.

(5) a. *Núzhi^{n}ga ak^{h}a* {*théthudi* *níkashi^{n}ga* ***gthí^{n}*** *thi^{n}k^{h}e*} *thé gi-gó^{n}tha.*
boy ART.PX {here person sit ART.OBV} go POSS-want

b. *Núzhi^{n}ga ak^{h}á* {*níkashi^{n}ga* *théthudi* ***gthí^{n}*** *thi^{n}k^{h}e*} *éshti thé gi-gó^{n}tha.*
boy ART.PX {person here sit ART.OBV} too go POSS-want

'The boy wants {the man who lives here} to leave.' (Rudin et al. 1989-92:Tape 3, 35'26 / Clifford Wolfe)

The second argument is that head nouns are never followed by articles, as noted by both Koontz (1984) and Rudin (1991). It is a well known characteristic of IHRCs across languages that they always have a morphologically indefinite head (see for instance Williamson 1987, Platero 1974). This "Indefiniteness Restriction" has been attributed to compositional semantic effects which are beyond the scope of this paper; for our purposes the point is just that a restriction to indefinite heads in Umo^{n}ho^{n} RCs is expected if they are IHRCs but would be unexpected for external-headed relatives. Since all articles in Umo^{n}ho^{n} are definite, the head of an IHRC cannot be determined by any article. Instead, the RC as a whole is almost always determined by a single article, which functions as a clause nominalizer, marking the clause as a DP. In (6), the head noun *wa'ú* 'woman' has no article, and the whole clause is at the same time both relativized and determined (marked as definite) by the article *ak^{h}á*.

(6) {*wa'ú* *Ø dúda* ***a****-i* *ak^{h}á*} *í-ko^{n}* *wiwíta.*
{woman here come-PX ART.PX} POSS:3-grandmother POSS:1SG

{'The woman who's coming over here} is my grandmother.' (Rudin et al. 1989-92:Tape 17 / Clifford Wolfe)

In a simple sentence the noun *wa'ú* would be expected to have the article, as in (7), where the DP *wa'ú ak^{h}á* 'the woman' is the subject of the sentence. Compare (6), where the notation'Ø' indicates that no article is possible in that position.

(7) *Wa'ú* *ak^{h}á* *dúda a-i*
woman ART.PX here come-PX
'The woman is coming over here' (constructed example).

The articles used as relativizers mark the role of the RC in the matrix clause, not that of the head noun in the relative clause. For instance, the choice of proximate versus obviative article depends on the function of the RC within the matrix clause, not the function of the head noun within the RC. So, in (8), where the head noun is the subject of the predicate inside the RC, but the RC itself is the object in the matrix clause, the relativizer is an obviative article, generally used for object rather than subject.

(8) Relative clause as object of matrix clause

shó^{n} wabágtheze uxthé-xchi thé, {wabáxu théthu thi^{n}k^{h}é}, éskana tha-'í
and letter soon-INTENS go {writer here ART:OBV} OPT A2-give
gthí-tha-the ko^{n}bthégo^{n}.
arrive.back-A2-CAUS A1SG.hope

'I hope that, as this letter goes very soon, you will give one and send it to {the writer who is here}.' (Dorsey 1890:509.2-3 / Dúba-mo^{n}thi^{n})

3. Diversity of RCs: range of functions and structures

3.1. Relativized arguments and adjuncts

All types of arguments can be relativized in Umo^{n}ho^{n}. Example (5) above illustrates the relativization of the subject of an intransitive active verb. (To repeat, the relativized syntactic role is visible within the RC: 'the man lives here'. It must not be mistaken for the syntactic role of the RC in the matrix clause.) Other relativized syntactic roles are exemplified below. In each case, the RC is between brackets, the head noun is single-underlined and the verb stem is boldfaced, as usual. (Note that in (9), *théthu* 'here' is not a verb stem, but it is used predicatively. It corresponds to an intransitive stative construction.)

(9) Relativization of the subject of a stative predicate

shó^{n} wabágtheze uxthé-xchi thé, {wabáxu ***théthu*** *thi^{n}k^{h}é}, éskana tha-'í*
and letter soon-INTENS go {writer here ART:OBV} OPT A2-give
gthí-tha-the ko^{n}bthégo^{n}.
arrive.back-A2-CAUS A1SG.hope

'I hope that, as this letter goes very soon, you will give one and send it to {the writer who is here}.' (Dorsey 1890:509.2-3 / Dúba-mo^{n}thi^{n})

Since property words are verbs in this language, any noun modified by a property stative verb is a RC in Umo^{n}ho^{n}. This is the analysis followed by Koontz (1984: 175), who illustrates it with (10), another example of relativization of the subject of a stative predicate.

(10) *{Mó^{n}ze ná-****zhide*** *thé-k^{h}e} ú t^{h}e í-pistásta ki (...)*
{iron INS:TEMP-red DEM-ART.HORIZ} wound ART.VERT AP-A1SG.press.down when
'When I press {these heated irons} repeatedly against the wounds, (...)' (Dorsey 1890:231.19 / Páthi^{n}-no^{n}pázhi)
Literally: When I press against the wounds those {irons which are heated by fire}, (...).

In a transitive clause, either the subject or the object can be relativized, as shown in (11) and (12) respectively:

(11) Relativization of the subject of a transitive verb

"..." á-biamá {níashi^{n}ga shínudo^{n} ***thixé ahí*** *ak^{h}á.}*
... say-REPORT {person dog pursue arrive ART:PX}

'said {the men who had pursued the dog and arrived at the Orphan's}.' (Dorsey 1890:113.4 / Frank La Flesche)

(12) Relativization of the object of a transitive verb

{*Waxí^{n}ha* ***gthí****-tha-the* *thó^{n}*} *bthíze.*
{paper arrive.back-A2-CAUS ART.RND} A1SG.take

'I have received {the letter which you sent home}.' (Dorsey 1890:511.1 / Ishtáthabi)

Example (13) shows the relativization of the theme of *'í* 'to give x to y'. Example (14) shows the relativization of the recipient of the same verb, but the RC is in apposition to the noun phrase headed by *mí^{n}zhi^{n}ga* 'girl'; note that the appositive phrases, *mí^{n}zhi^{n}ga thi^{n}k^{h}é* 'the girl' and *úwi a-'i thi^{n}k^{h}é* '(the one) I gave the earrings to,' have the same final article. The RC here is a headless relative clause; that is, its head is null as indicated by "Ø".

(13) Relativization of the theme in a ditransitive clause

{*shó^{n}ge shéna* *o^{n}-thá-****'i*** *tho^{n}k^{h}á*} *t'a=í*
{horse no.longer P1SG-A2-give ART.OBV.PL} die=PL

'{The horses which you gave me} have died.' (Dorsey 1890:480.2 / Pí-zi-thí^{n}ge)

(14) Relativization of the recipient in a ditransitive clause

Mí^{n}zhi^{n}ga thi^{n}k^{h}é *tó^{n}be,* {*Ø úwi* *a-****'í*** *thi^{n}k^{h}e.*}
girl ART.OBV A1SG.see { earring A1SG-give ART.OBV }
OBJ SBJ+VERB APP = RC

'I saw the girl, {(the one) I gave the earrings to.}' (Rudin et al. 1989-92:Tape 17 / Mary Clay)

Example (15) illustrates the relativization of an applicative object indicating a location (inessive locative 'in', introduced by the applicative prefix *u-*). The applicative verb is *utí* 'to camp in x'.

(15) Relativization of an applicative object

Gó^{n} {*wach^{h}íshka zhí^{n}ga o^{n}g-ú-****ti****=i* *k^{h}e*} *híde-ata shó^{n}ge ma*
and {stream small A1PL-AP:INESS-camp=PL ART.HORIZ} base-ALL horse OBV.PL
thé-o^{n}-wo^{n}-tha-í.
go-A1PL-O3PL-CAUS=PL

'We sent the horses towards the mouth of {the small stream by which we camped}.' (Dorsey 1890:438.3 / Páthi^{n}-no^{n}pázhi)

Adjuncts of time and place can also be relativized, as shown in (16) and (17).

(16) Relativization of an adjunct of time

ki {*ó^{n}ba wi-****tó^{n}be*** *t^{h}e*} *ét^{h}o^{n}do^{n}* *wa-shtó^{n}be* *tat^{h}é ebthégo^{n}.*
and {day A1SG/P2.see ART.VERT} by.that.time ANTIP-A2.see IRR A1SG.think

'I think that you shall see it by {the day that I see you}.' (Dorsey 1890:741.7 / Fred Merrick)

Example (17) shows a RC corresponding to an adjunct of place. We assume that the RC is headless, and precedes the phrase *pamú ámusta* 'right above the descent of the hill,' and that both together form the adjunct. The literal translation would be: "When he goes, where they surrounded the herd, on the hill on top of it, I will lie looking at you." The RC can be recognized from the article that nominalizes it.

(17) Relativization of an adjunct of place

*thé t^{h}e {Ø wá-na~**náse** tho^{n}} pamú ámusta wi-tó^{n}be*
go when { ANTIP-REDUP~surround ART.RND} downhill top A1SG/P2-A1SG.see
a-zhó^{n} tá mi^{n}k^{h}e, á=biamá.
A1SG-lie IRR 1SG.AUX say=PX.REPORT

'When he goes, I will lie looking at you, right above the descent of the hill {where they have surrounded the herd from time to time}.' (Dorsey 1890:45.10 / Nudó^{n}-axa)

3.2. Headless relative clauses

In Umo^{n}ho^{n}, numerous RCs lack an overt head noun, as already noted for (14) and (17). Another example is (18). The absence of a head noun is highlighted by "Ø". This kind of RC is very common in Umo^{n}ho^{n}.

(18) *{ Ø wa-**mó^{n}tho^{n}**=no^{n} thi^{n}k^{h}e } íbaho^{n}*
{ ANTIP-steal=HAB ART:OBV} know
'They know {the one who steals}.' (Rudin et al. 1989-92:Tape 3 / Mary Clay)

Note that when the head noun is null, it sometimes implicitly is taken to refer to a specific entity, as in (18), or it can be non-specific, and refer to any entity that would match the description of the RC. In the latter case, the head noun can be expressed by a generic term like *i^{n}dádo^{n}* 'what' in (19), with the same semantic result. These headless RCs, or RCs with generic heads, correspond to free relative clauses in English (Rudin 1991; see also Creissels 2006:208).

(19) *{I^{n}dádo^{n} i^{n}-wí^{n}-**go^{n}za**=i t^{h}e} gáxe go^{n}tha=bazhi=no^{n}.*
{what A1PL-D3PL-show=PL ART.VERT} make want=PL.NEG=HAB
'They don't want to do {what we teach them}.' (Rudin et al. 1989-92:Tape 15 / Bertha Wolfe)

The fact that the relativizers and definite determiners are the same set of morphemes (the "articles"), and the head noun is regularly missing, explains why in many contexts nouns are difficult to distinguish from verbs and relative clauses can be confused with simple DPs, especially when the supposedly relativized verb takes the ambiguous prefix wa- (which has both antipassive and nominalizing functions), as in (17) and (18). The RC *wa-mó^{n}tho^{n}=no^{n} thi^{n}k^{h}e* 'the one who steals' could alternatively be analyzed as a noun phrase 'the thief'.

3.3. Some RC lack a nominalizing/relativizing article

Umo^{n}ho^{n} relative clauses can also lack a relativizer/nominalizer (article), though this occurs infrequently. This is the case in (20) for the relativization of an applicative object (an instrument), and probably in (21) for an object.

(20) {*edádo^{n}* *í-**shko^{n}~shko^{n}*** *Ø*} *thi^{n}gé.*
{what AP:INS-REDUP~act (ART)} lack
'He has nothing {by means of which he can act often} (?)'[1]. (Dorsey 1891:75.4 / To^{n}wo^{n}-gaxe-zhi^{n}ga)

(21) *shé-ama* {*níkashi^{n}ga* *a-wá-**to^{n}be*** *Ø*} *úzhawa XXX*
that-PX.PL {person A1SG-O3PL-A1SG.see (ART)} enjoy
'These people I see are having a good time' (Rudin et al. 1989-92:Tape 19 / Coolidge Stabler; transcribed with Octa Keen)
(The end of the sentence is inaudible. The RC seems to be an apposition to shé-ama'these ones' and would be expected to end with the same article. The translation is mine (Marsault). Octa gave a word-for-word translation.)

Example (22) shows two relative clauses, the first lacking an article, and the second lacking a head-noun (another good example of the tricky noun-verb distinction).

(22) {*shó^{n}ge* *wa-**'í^{n}**-k^{h}ithé* *Ø* } *wébthi^{n}=hno^{n}=mó^{n}* {*Ø* *wa-**náse***
{horse ANTIP-carry-DAT.CAUS ART } D3PL.have=HAB=1SG.AUX { ANTIP-surround
amá }.
ART:PX.PL }
'But I used to take care of {the packhorses} for {those who surrounded the herd}.' (Dorsey 1890:466.2 / Frank La Flesche)

The reason why some RCs lack an article remains to be understood. Non-clausal DPs also lack an article in some cases, for instance when the referent is easily accessible or when it is indefinite or generic (see Gordon 2016; Marsault 2021:§8.4). Inasmuch as RCs externally act as DPs, it seems logical that they are subject to the same kind of article deletion. This topic requires further investigation.

At the typological level, Rudin (1991) notes that internally-headed relative clauses are a feature that often appears cross-linguistically with other types of nominalized clauses (see Culy 1990). Creissels (2006:246) provides a list of features often found in languages that have IHRCs, among which are the clause-final position of the verb, that the relativizer is typically positioned at the right edge of the clause, and that it is not uncommon to find constructions without any relativizer (which supports our point here).

4. Conclusion: theoretical issues, open questions

Within Umo^{n}ho^{n}, the most interesting issue raised by IHRCs is how they contribute to an understanding of the articles. As mentioned earlier, the Dhegiha article system is unusual in encoding

[1]The question mark in parentheses is in Dorsey's original translation.

a number of semantic features which are not typical of definite articles across languages, such as position, animacy, and obviation. Furthermore, this group of words has functions other than the typical article function of nominal determiner; all (or nearly all) of the "articles" also function as clause-final markers of some sort too: complementizer, evidential, or various types of auxiliaries (Eschenberg 2005). In the case of the IHRC, the article seems almost balanced between clausal and nominal functions; it is clause-final, but it also marks the end of a DP and functions as a nominalizing element.

Widening the focus beyond Umo^{n}ho^{n} and Siouan, IHRCs are cross-linguistically rare (occurring in less than 3 percent of Dryer's 2013 sample of 824 languages), and although they are found in a number of language families, they are still quite under-studied. In addition to Siouan languages, IHRCs have been described in Quechua (Cole 1987, Hastings 2004), Navajo (Basilico 1996, Willie 1989), Korean (Chung & Kim 2003), at least one Tai language (Moroney 2017), several African languages (Culy 1990), and have been noted to occur in various others.

A number of formal syntactic issues raised by IHRCs remain open. Deeper study of IHRCs in Umo^{n}ho^{n} and other Siouan languages could contribute to the general linguistic conversation over these issues. Rudin (1991) listed several questions, at least two of which remain unanswered: (a) Is there an external (empty) head N or NP (at any level)? (b) Is there movement of the head N or NP or an abstract wh- element to some position such as Spec or Comp? (at LF)? Even asking these questions obviously depends on a particular theoretical framework (and may be of little interest to most Siouanists). We will not go into the arguments here except to note that, while some works (e.g. Moroney 2017) have suggested both answers may be "yes", the issue is far from settled.

Abbreviations

1 first person; 2 second person; 3 third person; A agent; ALL allative; ANTIP antipassive; AP applicative; APP appositive; ART article; AUX auxiliary; CAUS causative; CP complementizer phrase; D dative (pronominal); DAT dative; DP determiner phrase; EVID evidential; HAB habitual; HORIZ horizontal; INESS inessive; INS instrument; INTENS intensifier; IRR irrealis; MOV moving; NEG negative; O object; OPT optative; OBV obviative; P patient; PL plural; POSS possessive; PX proximate; REDUP reduplication; REPORT reportative; RND round; SBJ subject; SCT scattered; SG singular; SIT sitting; STD standing; TEMP temperature; VERT vertical.

References

Basilico, David. 1996. The Navajo relative clause. *International Journal of American Linguistics* 40. 202–246.

Boyle, John P. 2016. The syntax and semantics of internally headed relative clauses in Hidatsa. In Catherine Rudin & Binah T. Gordon (eds.). Berlin, Germany: Language Science Press.

Chung, Chan & Jong-Bok Kim. 2003. Differences between externally and internally headed relative clause constructions. In Jong-Bok Kim & Stephen Wechsler (eds.), *The proceedings of the 9th International Conference on HPSG*. Stanford, CA: Stanford University.

Cole, Peter. 1987. The structure of internally headed relative clauses. *Natural Language and Linguistic Theory* 5(3). 277–302.

Creissels, Denis. 2006. *Syntaxe générale, une introduction typologique, tome 2, La phrase*. Paris, France: Hermes Science Publications.

Culy, Christopher. 1990. *The syntax and semantics of internally headed relative clauses*. Stanford, CA: Stanford University dissertation.

Dorsey, James Owen. 1890. *The Cegiha language* (Contributions to North American Ethnology, Vol 6.). Washington, DC: U.S. Government Printing Office.

Dorsey, James Owen. 1891. *Omaha and Ponka letters.* (Bulletin 11). Washington, DC: Bureau of American Ethnology.

Dryer, Matthew S. 2013. Order of relative clause and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. (Available online at https://wals.info/chapter/90). Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology.

Eschenberg, Ardis. 2005. *The article system of Umóⁿhoⁿ (Omaha)*. Buffalo, NY: State University of New York at Buffalo dissertation.

Gordon, Binah T. 2016. Information-structural variations in Siouan languages. In Catherine Rudin & Binah T. Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 393–422. Berlin, Germany: Language Science Press.

Graczyk, Randolph. 1991. *Incorporation and cliticization in Crow morphosyntax*. Chicago, Illinois: University of Chicago dissertation.

Hastings, Rachel Elizabeth. 2004. *The syntax and semantics of relativization and quantification: the case of Quechua*. Ithaca, NY: Cornell University dissertation.

Helmbrecht, Johannes. in progress. *A corpus-based reference grammar of Hoocąk*. Ms. Regensburg, Germany.

Koontz, John. 1984. *Preliminary sketch of the Omaha-Ponca language*. Ms. University of Colorado Boulder.

Marsault, Julie. 2021. *Valency-changing operations in Umóⁿhoⁿ: affixation, incorporation, and syntactic constructions*. Paris, France: Université Sorbonne Nouvelle - Paris 3 dissertation.

Moroney, Mary. 2017. Internally-headed relative clauses in Shan. Paper presented at Chulalongkorn International Student Symposium on Southeast Asian Linguistics.

Platero, Paul R. 1974. Head position and internally headed relative clauses. *Language* 72(3). 498–533.

Quintero, Carolyn. 2004. *Osage grammar*. Lincoln, NE: University of Nebraska Press.

Rudin, Catherine. 1991. *Relative clauses in Omaha-Ponca (Siouan)*. Presented at the American Anthropological Association annual meeting, Conference on American Indigenous Languages sponsored session.

Rudin, Catherine, Mary Clay, Clifford Sr. Wolfe & Bertha Wolfe. 1989-92. *Catherine Rudin's field tapes*. Ms. Transcriptions and audio files.

Rudin, Catherine & Kathleen Shea. 2005. Omaha-Ponca. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., vol. 9, 25–29. Boston, MA: Elsevier Publishers.

Ullrich, Jan. 2018. *Modification, secondary predication and multi-verb constructions in Lakota*. Düsseldorf, Germany: Heinrich Heine University dissertation.

Ullrich, Jan & Jr. Black Bear Ben. 2016. *Lakota grammar handbook*. Bloomington, IN: Lakota Language Consortium.

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.

Willie, MaryAnn. 1989. Why there is nothing missing in Navajo relative clauses. In Eung-Do Cook & Keren Rice (eds.), *Athapaskan linguistics*, 407–437. Berlin, Germany: Mouton de Gruyter.

About the Proceedings

The *Proceedings of the Siouan and Caddoan Languages Conference* (SCLC) is an annual periodical where any presenters at the SCLC are welcome to submit papers for publication. The scope of the papers that appear here range from highly theoretical linguistics papers to papers dealing with philology, language education, and cultural topics relating to Siouan- or Caddoan-speaking peoples. The aim of the *Proceedings* is to be inclusive towards scholars from a variety of backgrounds and provide a venue for work that scholars, community members, and/or educators can utilize to promulgate their work relating to these two language families.

These papers are reviewed by each volume's editor(s) and other volunteers from the Siouanist and Caddoanist communities. These volunteers are typically sought through the Siouan Linguistics Listserv. While the Listserv may have "Siouan" in the title, it is also a community for those working on Caddoan languages, given the historical overlap between scholars who worked with communities from both cultural and linguistic groups.

Anyone interested in Siouan and/or Caddoan languages is welcome to sign up for the Siouan Linguistics Listserv, which is operated through the University of Nebraska. The link to sign up for the Listserv appears below:

https://listserv.unl.edu/signup-anon?LISTNAME=siouan&LOCKTYPE=LIST&REALNAME=TRUE

Volumes of the *Proceedings* from the SCLC 38 and onward are available at http://www.siouan.org, a website dedicated to archiving information related to the SCLC, as well as compiling bibliographies of published worked on Siouan and Caddoan languages.