Modelling Plains Cree Negation with Constraint Grammar

Katherine Schmirler
University of Alberta
Department of Linguistics
schmirle@ualberta.ca

Antti Arppe
University of Alberta
Department of Linguistics
arppe@ualberta.ca

Abstract

This paper explores negation in a Plains Cree corpus, using Constraint Grammar to model various aspects of negation (verbal, nominal, particle). Plains Cree, an Algonquian language of North America, displays rich morphological marking on nouns and verbs, but also makes use of a large class of indeclinable particles, including negative markers. Combined with non-configurational syntax and few strict word order patterns, modelling syntactic relationships involving particles is far from straightforward. Using previous grammatical description and corpus observations, we model the relationships between negative particles and other parts of speech in Plains Cree and identify issues for further modelling.

1 Introduction

Previous work on Plains Cree syntactic modelling has aimed to identify basic syntactic relationships between verbs and core arguments (Schmirler et al., 2018), which can, for the most part, be identified on the basis of previously marked lexical features (i.e. noun class or verb class) and morphological features output by a morphological model (cf. Snoek et al., 2014; Harrigan et al., 2017). For further development, considerations beyond simple morphological features need to be made. In the present work, we detail the ongoing process of modelling Plains Cree negative particles, which are identified from a dictionary of Plains Cree (Wolvengrey, 2001) for reference in the CG parser.

2 Background

2.1 Plains Cree

Plains Cree is a member of the Algonquian language family of North America, which ranges across much of the continent, from the Blackfoot and Cree languages spoken as far west as Alberta, Canada and Montana, USA, to the eastern coasts of the both countries. Plains Cree is the westernmost member of the Cree language continuum, which includes Cree dialects across Alberta, Saskatchewan, northern Manitoba and northern Ontario, as well as the closely related Montagnais-Naskapi dialects, spoken further east in Quebec and Labrador (Wolvengrey, 2011).

Algonquian languages are known for their rich morphology and non-configurational syntax, such that syntactic roles are determined by morphological agreement rather than word order. The rich morphology allows for straightforward modelling of core argument relationships while particles, the most frequent word class evidenced in texts, bear essentially no inflectional morphology, and include words with a variety of different functions, which have not yet been given a detailed classification for Plains Cree. Without such a classification, development of particle constraints in the Plains Cree parser is an ongoing process, such as that described for negative particles in section 3.

2.2 A Plains Cree corpus

The texts to which the Plains Cree CG parser is applied are referred to herein as the Ahenakew-Wolfart (A-W) corpus (Arppe et al., in press). The A-W corpus consists of several texts (totalling ∼73,000 words of Plains Cree) collected in the 1980s and 1990s, which have been transcribed, edited, and in some cases translated, then published in several volumes (Ahenakew, 2000; Bear et al., 1998; Kâ-Nîpitêhtêw, 1998; Ma-

suskapoe, 2010; Minde, 1997; Vandall and Douquette, 1987; Whitecalf et al., 1993). Digital versions, which display less editing that the published texts (e.g. more fragments, commas representing pauses, etc.), have been supplied for the digital corpus by H.C. Wolfart. This corpus has been morphosyntactially analysed using a finite-state parser (Snoek et al., 2014; Harrigan et al., 2017), and subsequently tentatively disambiguated and parsed for core arguments (Schmirler et al., 2018) using CG-3 (Bick and Didriksen, 2015). The corpus is available upon request at URL: http://altlab.ualberta.ca/korp. The corpus contains a variety of genres, including historical narratives, personal narratives, funny stories, speeches/lectures, and dialogues; future research is planned to explore the ways in which genre affects morphosyntactic patterns in Plains Cree.

3 Considering negation in the Plains Cree CG parser

3.1 Negative particles

The initial implementation of negative particles in the CG parser was straightforward. First, a LIST of negative particles was created, and negative particle phrases were similarly identified; examples of particles are given in (1) and examples of phrases are given in (2).[1] These were assigned a morphological tag `Neg` for reference in later constraints and in corpus searches. The same morphemes appear repeatedly in these particles: *namôy*, *môya*, and *môy* are reduced forms of *namôya*; *kâwiya*, *êkâya*, *êkây*, *kâya*, and *êkâ* of *êkâwiya*; and *mwâc* of *namwâc*.[2] Even *namôya* is historically derived from *nama* plus a focus particle *wiya*. Similarly, though only a subset of the negative particle phrases are given here, the same phrases also occur with the reduced forms of *namôya* seen in the LIST.

(1) Negative particles

```
LIST NEG = "namôya" "namôy"
"môya" "môy" "êkâwiya"
"kâwiya" "êkâya" "êkây" "kâya"
"êkâ" "nama" "namwâc" "mwâc" ;
```

(2) Negative particle phrases

`"namôya_wîhkâc" Iph Neg` 'never'
`"namôya_ahpô" Iph Neg` 'not even'
`"namôya_cî" Iph Neg` 'it is not so?'

3.2 Constraint development

An initial constraint then assigned the function tag `@Neg>` when any of these words or phrases appeared immediately before a verb. While this does produce adequate results, it does not fully capture negation in Plains Cree. First, there can be some intervening material between the negative particle or particle phrase and the verb, some of these negative particles can modify nouns or other particles as well as verbs, and some of these particles modify only particular elements. For example, variations of *êkâwiya* modify particular conjugation patterns, and *nama* only occurs in particle phrases in the A-W corpus.

A cursory exploration of the A-W corpus begins to demonstrate the complexities of negation: of the ∼1510 negative particles (excluding particle phrases), ∼580 occur immediately before verbs, which are identified by the original simple constraint. Of the remaining particles, ∼25 occur immediately before nouns, ∼100 before pronouns, ∼550 before particles, ∼160 before punctuation, and the remainder before as-yet-unlabelled elements. While an exploration of how negatives interact with each of these types is beyond the scope of this paper, those before nouns can be explored readily. Several occur before a noun (or noun phrase) which is followed by punctuation, and so the negative particle does appear to modify the noun. In other cases, however, the noun is followed by a verb, without intervening punctuation, suggesting that the verb is negated. One sentence displays a series of negated elements, including nouns and a verb; a portion of this is given in (3).

(3) *...namôya nipiy,*
not water
namôya sâkahikana k-âtâmitân,
not lakes I buy from you
namôya kinosêw; ...
not fish
'...I do not buy the water, nor the lakes, from you, nor the fish...' (Kâ-Nîpitêhtêw, 1998, p. 110-13)

In example (3), where *namôya* occurs before *nipiy* 'water' or *kinosêw* 'fish', it is straightfor-

[1] The LIST approach was implemented for the initial development stage described herein; in future development, tags for particle functions can be added in the morphological model.

[2] All of these are glossed as 'no, not'.

ward to analyse each negative as dependent on the following noun. Before *sâkahikana* 'lakes', however, the verb is negated (cf. the English translation, where the verb is negated once and the remaining nouns are preceded by 'nor'). With this observation in mind, a constraint was written for nouns that makes use of clause boundaries, and the verbal constraint was modified to allow for intervening nominals.

A similar constraint was also written for particles, and the verbal constraint modified once again to allow for any intervening material, excluding clause boundaries. The current three negation constraints are given in (4). While these three constraints allow for many of the negative particles to receive a dependency tag, considerable further examination of the corpus is required to determine their accuracy.

(4) Negation constraints for Plains Cree

```
MAP:NegV @Neg-V> TARGET Neg
IF (*1 V BARRIER CLB) ;
MAP:NegN @Neg-N> TARGET Neg
IF (1 N) ;
MAP:NegIpc @Neg-Ipc> TARGET
Neg IF (1 Ipc) ;
```

This empirical approach quickly reveals issues of scope: often, a negative particle appears to negate a verb later in a sentence, rather than the immediately following noun or particle, regardless of punctuation. While these constraints can begin give some idea of how negation works in Plains Cree, and can be used to further develop more accurate constraints, additional research into the scope of negation is also required.

4 Negation in a Plains Cree corpus

4.1 Negative particles and other word classes

The parser labels ∼1495 negative particles with an `@Neg` tag. Of these, ∼1260 are dependent on a verb, ∼10 on a noun, and ∼140 on another particle. While the counts for nouns and particles are relatively similar to those given above for the parts of speech that occur immediately after particles, considerably more negatives are dependent on verbs, as the relevant constraint allows for any intervening material, including nouns and particles.

For these negative particles, one can also begin to explore how different subtypes behave, even without fully exploring the features of the verbs and nouns they are dependent upon. Over 1100 of the particles that receive an `@Neg` tag are *namôya*-type while only ∼300 are *êkâwiya*-type. For the *namôya*-type particles, ∼940 are dependent on verbs, ∼10 on nouns, and ∼110 on particles. For the *êkâwiya*-type, ∼290 are dependent on verbs, ∼5 on nouns, and ∼10 on particles.

4.2 Future considerations

Future research is planned to explore how negative particles interact with different inflectional patterns on verbs, as well as more detailed looks at nouns and particles. This also requires a deeper investigation of particle phrases and their functions; of particular interest are negative particle phrases with *kîkway* 'thing', as these often function as nominals. Many phrases also seem to negate clauses, and thus verbs, rather than nouns or particles—an impressionistic observation that requires further consideration.

5 Conclusions

Despite their morphological simplicity, negative particles in Plains Cree have presented an interesting exercise in modelling their relationships with nouns, verbs, and other particles. The combinatory freedom of particles and flexible word order of Plains Cree present an ongoing challenge for the development of a parser, and the identification of broad functions within the particle class, such as negation, has been but one step towards more detailed and accurate syntactic function tags for Plains Cree.

References

Alice Ahenakew. 2000. *âh-âyîtaw isi ê-kî-kiskêyihtahkik maskihkiy / They Knew Both Sides of Medicine: Cree Tales of Curing and Cursing Told by Alice Ahenakew*. University of Manitoba Press, Winnipeg.

Antti Arppe, Katherine Schmirler, Atticus G. Harrigan, and Arok Wolvengrey. in press. A morphosyntactically tagged corpus for Plains Cree. In *Papers of the 49th Algonquian Conference*.

Glecia Bear, Minne Fraser, Mary Wells, Alpha Lafond, and Rosa Longneck. 1998. *Our Grandmothers' Lives: As Told in Their Own Words*, bilingual edition edition. University of Regina Press, Regina.

Eckhard Bick and Tino Didriksen. 2015. Cg-3—beyond classical constraint grammar. In *Proceedings of the 20th nordic conference of com-*

putational linguistics, NODALIDA 2015, may 11-13, 2015, vilnius, lithuania, 109, pages 31–39. Linköping University Electronic Press.

Atticus G. Harrigan, Katherine Schmirler, Antti Arppe, Lene Antonsen, Trond Trosterud, and Arok Wolvengrey. 2017. Learning from the computational modelling of Plains Cree verbs. *Morphology*, 27(4):565–598.

Jim Kâ-Nîpitêhtêw. 1998. *ana kâ-pimwêwêhahk okakêskihkêmowina / The Counselling Speeches of Jim Kâ-Nîpitêhtêw*. University of Manitoba Press, Winnipeg.

Cecilia Masuskapoe. 2010. *piko kîkway ê-nakacihtât: kêkêk otâcimowina ê-nêhiyawastêki*. Algonquian and Iroquoian Linguistics, Winnipeg.

Emma Minde. 1997. *kwayask ê-kî-pê-kiskinowâpatihicik / Their Example Showed Me the Way: A Cree Woman's Life Shaped by Two Cultures*. University of Alberta Press, Edmonton.

Katherine Schmirler, Antti Arppe, Trond Trosterud, and Lene Antonsen. 2018. Building a Constraint Grammar Parser for Plains Cree Verbs and Arguments. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Conor Snoek, Dorothy Thunder, Kaidi Lõo, Antti Arppe, Jordan Lachler, Sjur Moshagen, and Trond Trosterud. 2014. Modeling the Noun Morphology of Plains Cree. In *Proceedings of the 2014 Workshop on the Use of Computational Methods in the Study of Endangered Languages*, pages 34–42, Baltimore, Maryland, USA. Association for Computational Linguistics.

Peter Vandall and Joe Douquette. 1987. *wâskahikaniwiyiniw-âcimowina / Stories of the House People, Told by Peter Vandall and Joe Douquette*. University of Manitoba Press, Winnipeg.

Sarah Whitecalf, H. C. Wolfart, and Freda Ahenakew. 1993. *kinêhiyawiwiniwaw nêhiyawêwin / The Cree Language is Our Identity: The La Ronge Lectures of Sarah Whitecalf*. University of Manitoba Press, Winnipeg.

Arok Wolvengrey. 2001. *nêhiyawêwin: itwêwina // Cree: Words*, bilingual edition edition. University of Regina Press, Regina.

Arok Elessar Wolvengrey. 2011. *Semantic and pragmatic functions in Plains Cree syntax*. Netherlands Graduate School of Linguistics.