

Plains Cree Verbal Derivational Morphology: A Corpus Investigation

Katherine Schmirler, Atticus G. Harrigan, Antti Arppe, and Arok Wolvengrey

The complex morphosyntax of Plains Cree and related languages has long been the focus of Algonquian research. Verbs in particular offer a wealth of morphological issues to be investigated. For Plains Cree, verbal templates have been constructed containing up to seventeen slots for preverbs and inflectional affixes. For preverbs, up to eight slots are allotted for various functions, which may be either grammatical or inflectional. Derivational templates for Plains Cree verb stems also allow for several layers of secondary derivation, and extreme examples of Plains Cree verb stems contain seven or more morphemes (Bakker 2006; Wolfart 1973, 1996; Wolvengrey 2012). Our goal is to investigate the derivational morphological complexity of Plains Cree verbs in actual language use compared to the complexity demonstrated in theoretical templates; furthermore, we consider both lexical preverbs and stem derivational morphemes (i.e., those used in primary and secondary derivation), due to their considerable influence on semantics and lexeme formation, to constitute derivational morphology in the present study.

Previous investigations (e.g., Wolvengrey 2015) of Plains Cree have found restrictions on the number of preverbs before disfluencies (false starts, hesitations, etc.) begin to consistently occur; including both grammatical and lexical preverbs,

up to five are seen to occur before such disfluencies consistently appear. Thus, when lexical preverbs are combined with extreme examples of secondary derivation, we might expect up to fourteen derivational morphemes in the most complex of verbs. However, cursory corpus searches indicate that this kind of complexity is not the norm in Plains Cree. Here, we undertake a quantitative corpus investigation of Plains Cree derivational morphology, in which we investigate the overall derivational complexity of Plains Cree verbs and how stem and preverbal morphology co-occur. Some semantic and functional patterns become evident. Additionally, while extreme complexity might in principle be possible in Plains Cree verbs, actual language use demonstrates considerably fewer derivational morphemes than the theoretical maximum.

Background

Plains Cree

Plains Cree is a member of the Algonquian language family and the westernmost member of the Cree-Montagnais-Naskapi language continuum, spoken across Canada from Alberta to Labrador. Plains Cree has several thousands of speakers, mostly in Alberta and Saskatchewan. Still spoken and written in many contexts today, such as in online communication, radio and television broadcasts, and published books, Plains Cree offers an excellent opportunity for corpus investigations of a native North American language. The corpus used here is further discussed later.

While the focus of this paper is derivational morphology, several features common to Algonquian languages are referenced. These include primarily the noun classification system of animacy and verbal classification by animacy and transitivity. The animacy system includes the grammatical categories of animate and inanimate that are pervasive in Plains Cree through pragmatics, semantics, and morphosyntax. Verbs are categorized by the animacy of their participants and by their transitivity, resulting in four classes: inanimate intransitive, animate intransitive, transitive inanimate, and transitive animate. The derivation of verb stems involves different morphemes for different verb classes, and so these are referenced throughout.

Derivational Morphology

Like all Algonquian languages, Plains Cree is a polysynthetic language with complex derivational and inflectional morphology. There are various templates describing how verbal inflection (person, number, tense) and grammatical and lexical preverbs may occur (see Bakker 2006; Wolfart 1973, 1996; Wolvengrey 2012). Furthermore, recursive derivation within verb stems is also described, with extreme examples containing several layers of secondary derivation. For the purposes of our investigation into derivational morphology, we have included under the umbrella of “derivational morphology” stem derivation (both primary and secondary derivation), the subset of preverbs that are deemed to have more lexical rather than grammatical functions, and reduplicative preverbs—unlike clearly inflectional morphology, these morphological elements contribute to the lexical semantics of a word form and so are considered under one category. Stem derivation and preverbs are discussed in the following subsections.

Verb Stem Derivation

Plains Cree verbal derivation involves (1) root or initial, (2) medial, and (3) final morphemes. Primary verb stems are comprised of a root, an optional medial, and a required final, which may be a zero morpheme. A primary stem may then undergo further derivation, again with an optional medial and required final, forming a secondary stem, which may then undergo further derivation, as in the template in Table 1. Verb stems of five morphemes or more, such as those in Table 2 and Table 3, are commonly attested in Cree dictionaries (adapted from Wolfart 1996).¹

In our analysis, we consider the most common pairs of stem derivational morphemes and the patterns they suggest. We then scrutinize common combinations of derivational morphemes with preverbs.

Preverbs


Preverbs in Plains Cree are prefixed to verb stems and are preceded by person prefixes in the independent order. Preverbs may have grammatical functions, such as marking conjunct or relative clauses (*ê-, kâ-*), tense (*kî-, ôh-/ohci-*  *wî-* future intentional), or both (*ta-* conjunct infinitive, *kê-* changed conjunct future). The majority of preverbs, however, have more lexical functions. These occur closer to

TABLE 1. Verb stem derivation template

SECONDARY STEM					
SECONDARY STEM					ETC.
PRIMARY STEM			SECONDARY DERIVATION		
ROOT	(MEDIAL)	FINAL	(MEDIAL)	FINAL	
/wâp-/ 'light, bright'	/-âskw-/ 'wood'	/-(i)kê/ general object (AI)			
/pim-/ 'along'	/-âpisk-/ 'metal, stone'	/-payi/ 'move' (II/AI)			
/mihkw-/ 'red'	/-êk(inw)-/ 'cloth, material'	/-(i)n/ 'by hand' (TA)			
/it-/ 'thus, so'					

TABLE 2. Secondary derivation of verb stems

pimipayihcikêstamâso- 'manage for oneself' (TA)					
pimipayihcikêstamaw- 'manage for s.o.' (TA)					
pimipayihcikê- 'manage things' (AI)					
pimipayihtâ- 'manage, run s.t.' (TI)					
pimipayi- 'work, function' (II)					
ROOT	FINAL	FINAL	FINAL	FINAL	FINAL
/pim-/ 'along'	/-payi/ 'move'	/-htâ/ TI	/-(i)kê/ AI	/-stamaw/ BENEF	/-iso/ REFL

TABLE 3. Secondary derivation of verb stems

kanawêyimiskwêwêski- 'be a habitual wife-watcher' (AI)					
kanawêyimiskwêwê- 'watch one's wife' (AI)					
kanawêyim- 'watch over s.o.' (TA)					
ROOT	MEDIAL	FINAL	MEDIAL	FINAL	FINAL
/kanaw-/ 'watch'	/-êyi-/ 'by mind'	/-m/ TA	/-iskwêw-/ 'woman'	/-ê/ AI	/-ski/ habitual



TABLE 4. Preverb template for Plains Cree

GRAMMATICAL			LEXICAL		
CONJUNCT CLAUSE	TENSE	PERSPECTIVAL ASPECT	PARTICIPANT-ORIENTED MODALITY	PHASAL ASPECT	MANNER/DIRECTION
ê- kâ- ta- kê-	kî-	wî- ô(h)-/ohci-	nôhtê- 'want to' kawkwê- 'try to' nihtâ- 'be good at'	mâci- 'start' pôni- 'stop' kisi- 'finish'	nitawi- 'go and' isi- 'thus' pê- 'come and'

Adapted from Wolvengrey 2012

the verb stem than grammatical preverbs and carry various adverbial meanings that modify the verb stem or other lexical preverbs. Some grammatical and lexical preverbs and their identified functions are demonstrated in the preverbal template in Table 4.

While the set of grammatical preverbs is quite small, and the number of grammatical preverbs that can occur in any verb is restricted, lexical preverbs can occur much more freely and constitute a more open class, with well over 200 listed in Wolvengrey (2001); lexical preverbs may be derived from roots or stems using the particle final *-i*, and existing free particles can be co-opted for use as preverbs as well. Also included in dictionaries are verbal compounds, or lexicalized strings of preverbs and verb stems; they are included as lexical entries when they are common or have unpredictable semantics. For example, *mīyo-* ‘well, good’ compounds with *kīsikāw* ‘it is day’ to form *mīyo-kīsikāw* ‘it is a nice day’, but with *akohpêw* ‘s/he owns blankets’ to form *mīyo-akohpêw* ‘s/he owns good blankets’. Such lexical entries are also taken into account here as a preverb plus the number of stem morphemes.

While different slots for lexical preverbs of different functions are recognized, multiple preverbs of each lexical function may occur, so more than four lexical preverbs can theoretically be present in a verb form. Furthermore, reduplicative preverbal elements may also occur before lexical preverbs and the stem itself, and lend adverbial meaning: *Ca-* indicates an ongoing action and *Câh-* a repeated action. We have also counted these as lexical preverbs as they lend semantic information to the verb form. Wolvengrey (2015) has found in a corpus investigation that the maximum number of preverbs that can occur is five; after five preverbs, disfluencies appear consistently, such as false starts, hesitation markers (e.g., *ê-kî-pê-is[i]-âya-* . . .), or intrusive syntactic particles (e.g., *kâ-kî-î-hci-mâci-na-* . . .). While false starts or other repairs occur within verb stems, *this is*  rarer than disfluencies in the preverb complex; this suggests that preverbal derivation falls more toward the syntactic side of the morphosyntactic spectrum of the Plains Cree verbal complex.

The Corpus

The corpus used for the present study is comprised of Plains Cree texts edited by Freda Ahenakew and H. C. Wolfart (Bear et al. 1992; *Kâ-Nîpitêhtêw* 1998; Masuskapoe 2010; Minde 1997; Vandall and Douquette 1987; *Whitecalf* 1993). These texts add up to 108,413 tokens (18,649 types), of which 73,189 tokens (15,994 types) are

Plains Cree words (the difference being punctuation and English, French, and other foreign names and words), which, though dwarfed by available corpora for many majority languages, is a sizable corpus for an indigenous language and is sufficient for the quantitative analyses reported here. The corpus has been morphologically analyzed by means of finite state transducer (FST) tools (e.g., Beesley and Karttunen 2003) that provide for each word form its word class, subclass, and features such as number, obviation, person, direction, verbal order, possession, and diminution, and, key to the present investigation, identifies preverbs and presents the lemma form of the word based on Wolvengrey (2001). Furthermore, this corpus has been manually verified and corrected as necessary by two researchers, improving its coverage and accuracy considerably. The majority of errors or missing analyses in the output of the morphological analyzer involved orthographic differences such as vowel length mismatches, missing lemmata in the database, and morphological elements that have not yet been modeled. As the morphological analyzer is still in the development stage (currently described in Harrigan et al. 2017; Snoek et al. 2014), this manual verification not only improves the analyzed corpus but also identifies issues for further development.

After the manual verification, where we did not take sentential context into account, a degree of ambiguity remained among the analyses, with 22.9 percent of the Cree word form types having two or more possible analyses (of which the majority, 16.7 percent, were two-way ambiguities).² We applied a simple but straightforward heuristic implementing a preference for morphological simplicity that selected the analysis with the smallest number of morphemes, and these being equal, the smallest number of preverbs. As a result of this, only one morphological analysis remained for each Cree word form in the corpus. The majority of verbal ambiguities involve inverse VTA inflections (e.g., the suffix sequence *-ikoyâhk*, which may be a proximate, obviative, or inanimate actor with a first person plural exclusive goal), forms that may be grammatical preverbs or reduplication (e.g., *ka-*, *tâh-*), and preverbs that are also found in lexicalized preverb-stem entries in the database (e.g., the morphological analyzer returns two analyses for *mîyo-atoskêw* 'she works well', as this could be the preverb *mîyo-* 'good, well' plus the verb *atoskêw* 's/he works', but it is also a lexicalized compound found in dictionaries). These three patterns represent over one-third of the ambiguity in the corpus, and less frequent patterns are still predominantly inflectional. The latter two of these patterns do involve the verb stem and lexical preverbs and as such affect the present study; we opted to select lexicalized compounds as they are especially representative


of derivation. Additionally, though our disambiguation allowed for any potential reduplicative suffix to be represented as such, these were infrequent enough that they do not occur in our results.

Method

For the purposes of our investigation, we have made use of the output of the analyzed corpus in two key ways. First, we do not consider grammatical preverbs. These include preverbs that mark conjunct clauses and tense preverbs. Some preverbs have multiple functions, some grammatical and some lexical; for instance, *ohci-* can indicate past tense in a negative clause or indicate the source or origin of an action. While these were not distinguished before the analyses, where *ohci-* occurs as the first member of a frequent preverb pair, it is apparent from corpus searches that these clauses are negative; these pairs are then ignored in the analysis.

Second, we have taken the lemma form supplied by the analysis and drawn its morphological composition from the database underlying Wolvengrey's (2001) dictionary of Plains Cree. This database gives the root, medial, and final morphemes of each lemma, where they can be identified through synchronic and diachronic analysis. The lexical preverbs identified by the analyzer and the stem morphology drawn from the database are used in the following analyses. For instance, a verb form such as (1) found in the corpus was identified to have four preverbs, *kâ-*, *kî-*, *isi-*, and *miyo-* (of which the last two are lexical, namely *miyo-* and *isi-*), and the lemma *waskawîstamâsow* 's/he works for him/herself' has five derivational morphemes, given in (2).

- (1) *kâ-kî-isi-miyo-waskawîstamâsot*
 CNJ-PST-thus-well/good-work.for.oneself.3SG
 'thus s/he worked well for him/herself'

- (2) *waskawîstamâso—*  *kaw-* /-î/ /-st/ /-amaw/ /-iso/
 turn, move go, move TRZ BENEF REFL
 root VAI final final VTA final AI final
 'to work for oneself'

Results

Altogether, the corpus contained 9,983 unique verb forms types (representing 16,538 tokens); thus, 62.4 percent of all the Cree word form types only 22.5 percent of the word tokens in the corpus are verbs. The verb forms in the corpus exhibited 149 preverbs (of which 18 were grammatical ones) and 1,171 distinct stem morphemes. The most common lexical preverbs, in terms of their occurrences in various verb form types, were *pê-* ‘come’ ($n = 541$), *isi-* ‘thus’ ($n = 277$), *nitawi-* ‘go’ ($n = 229$), *ati-* ‘begin to’ ($n = 146$), and *kakwê-* ‘try’ ($n = 124$), and the most common verb stem morphemes were */-ê/* ($n = 860$), */-â/* ($n = 836$), */-h/* ($n = 808$), */-ht/* ($n = 740$), */-m/* ($n = 716$), */-i/* ($n = 687$), and */iT-/³* ($n = 628$).

Next, co-occurrence frequencies and statistics were calculated, retaining co-occurrence order as apparent in the word forms, though that could in principle vary, yielding in all 303 unique preverb pairings, 4,027 stem morpheme pairings, and 4,788 preverb-stem morpheme pairings. Among the lexical preverbs, the ones that combine the most others are *pê-* (24), *isi-* (13), and *nitawi-* (12). Of the many co-occurrence statistics, MUTUAL INFORMATION (MI) was selected, since it is fairly easy to understand in representing the strength of a joint occurrence of two features in comparison to the individual frequencies of these two features; the higher the MI value is, the stronger the joint occurrence of a feature pair is in comparison to the frequencies of the individual features. Nevertheless, with a small corpus such as the one at our disposal, the downside of the MI measure is that it is relatively useless when the individual and joint feature frequencies are relatively low (cf. Evert 2004), and thus we do not make much use of it but provide it simply for reference.

In Tables 5 and 6, the last two columns give the morphemes (preverbs or stem derivation elements) in question and their glosses. The first columns indicate the MI value, the second and third columns indicate the individual frequency of the first and second elements respectively, and the fourth column the joint frequency of a feature pair.

Preverbs

In Table 5, we present the twelve most common lexical preverb pair combinations in the corpus. While the most common pair *pê- . . . isi-* occurs in thirty-eight word forms in the corpus,⁴ the rest of these most frequent pairs drop to less than a

TABLE 5. Most frequent preverb pairs

MI	N ₁	N ₂	N ₁₊₂	PREVERB PAIR	PAIR GLOSSES
0.898483	541	277	38	pê- isi-	'come', 'thus, towards'
2.566461	541	11	8	pê- kîwê-	'come', 'back, homeward'
0.120329	124	277	4	<i>kakwê- isi-</i>	'try', 'thus, towards'
1.251731	277	40	4	isi- miyo-	'thus, toward', 'good, well'
2.038673	52	97	4	nipahi- misi-	'really, extremely', 'large, great'
-0.591184	541	97	3	pê- misi-	'come', 'large, great'
-0.110818	541	40	2	<i>pê- nihtâ-</i>	'come', 'be good at'
-0.136952	229	97	2	nitawi- misi-	'go', 'large, great'
-0.736143	146	277	2	<i>ati- isi-</i>	'begin to', 'thus, towards'
-1.242220	124	541	2	<i>kakwê- pê-</i>	'try', 'come'
-1.710206	541	198	2	pê- ohci-	'come', 'from'
1.093155	541	12	2	pê- sipwê-	'come', 'leave'

dozen. Compared to the individual counts for each preverb, we see that the pairs are not particularly common; while common combinations do exist, they are much less frequent than their individual members. This further reinforces the idea that preverbs combine quite freely and behave more as syntactic elements, with greater combinatory freedom, than the morphological elements within verb stems.

However, among these common preverb pairs, some function combination patterns do arise. First, in black boldface, are combinations of manner/direction preverbs; *pê-* 'come', *isi-* 'thus, thither', *ohci-* 'by means of, thence', *kîwê-* 'back, homeward', *sipwê-* 'leaving' are seen here. Second, in italicized boldface, are combinations of manner/direction preverbs and participant-oriented modality preverbs (*kakwê-* 'try', *nihtâ-* 'be good at'). However, due to the order of manner followed by modality, which contradicts the order given in the template in Table 2, further investigation into the context of the forms is required. Finally, there is one instance of a phasal aspect preverb *ati-* 'begin to' followed by a manner/direction preverb (italicized), which is in line with the aforementioned templates. As there is only one instance here, however, no further comment can be made. The remaining preverbs, *misi-* 'big', *nipahi-* 'extremely', and *miyo-* 'good, well', are qualitative adverbs and are not included in the available templates.

TABLE 6. Most frequent stem morpheme pairs

MI	N ₁	N ₂	N ₁₊₂	STEM COMBINATIONS	PAIR GLOSSES
1.922021	540	740	282	/-êyi-/ /-ht/	'think', TI
1.434042	740	836	268	/-ht/ /-â/	causative, TI
2.170965	325	860	253	/-oht-/ /-ê/	'walk', AI
2.423391	256	808	241	/wîht-/ /-h/	'fellow', causative TA
1.781144	540	716	237	/-êyi-/ /-m/	'think', TA
2.225492	234	836	187	/ay-/ /-â/	'thing', TI
1.013962	687	716	140	/-i/ /-m/	AI, TA
2.947364	148	443	129	/wîht-/ /-amaw/	'tell about', benefactive TA
1.654565	716	331	128	/-m/ /-o/	TA, AI
2.645999	114	687	114	/w-/ /-i/	3, AI
1.670467	628	325	112	/iT-/ /-oht-/	'thus, towards', 'walk'
2.453104	128	716	110	/âh-/ /-m/	'tell about', TA
2.645999	108	687	108	/-îk-/ /-i/	'dwelling', AI
4.432833	114	108	107	/w-/ /-îk-/	3, 'dwelling'
0.471248	687	836	95	/-i/ /-â/	AI, TI

Stem Morphology

In Table 6, we present the fifteen most common pairs of derivational morphemes found in stems in the corpus. Compared to the preverbs above, we see much more frequent pairs occurring throughout the words found in the corpus, though the individual occurrences still number in the hundreds. While these indicate pairs that co-occur in a stem regardless of whether they are adjacent, most still represent sequences of morphemes. For instance, **-êyi-ht** and **-êyi-m** which derive 'thought processes' (VTI and VTA respectively, e.g., *itêyih tam* 'he thinks thus (of s.t.)', *itêyimêw* 'he thinks thus of s.o.'). are often considered extended finals and written without hyphens (Wolfart 1973:74).

While many of the pairs in this table contain more abstract grammatical morphemes for which concrete meanings are not discernable (e.g., *-i*, *-a*, *-m*), some patterns are seen in finals. For instance, we see considerable number of occurrences of the 'thought process' morphemes mentioned above (*-êyih*, *-êyim*) which indicate the use of many VTI and VTA verbs formed with these morphemes. Similarly, the sequence *-ht-â*, which commonly forms VAI+O verb stems (Wolfart 1973:74), indicates that many frequent VTIs have been formed using these morphemes. Forms with more specific semantics are also evident, given in italicized boldface.

The sequence *-oht-ê* forms an extended VAI final meaning ‘walk’ (e.g., *pimohhtë-* ‘to walk’, from the root *pim-* ‘along’), and *-oht-* is also seen in the pair *iT-oht*, which occurs in the frequent VAI *itohtëw* ‘he goes (there)’.⁵ Finally, common root-final combinations are seen here, given in italics in Table 6. The frequency of /wîht-/ /-amaw/ matches the occurrences of the VTA stem *wîhtamaw-* ‘to tell someone about it/him’ in the corpus, indicating that this is not only a common sequence but in fact is only realized in one common verb stem. The pair /wîht-/ /-h/ is seen in the similarly common VTA stem *wîcih-* ‘to help someone’, but is also seen in secondary derivations based on this stem, such as the VAI *wîcihiwê-* ‘to help people’.

Unlike the preverb pairs above, these morpheme pairs not only occur more commonly together but also form common verb stems or extended finals. Though the MI measures may not fully reliable in such a small data set, a cursory comparison of Tables 5 and 6 reveals a stark difference between the numbers: the stem morpheme pairs occur together more often than preverb pairs.

Co-occurrences

In Table 7, we present the fifteen most common pairs of preverbs and stem morphemes. The common manner and direction preverbs are once again apparent, especially *pê-*, though these generally occur with short, abstract verb finals. There are two pairs that suggest a semantic pattern and one that is indicative of a common preverb-verb stem combination, all given in boldface. The pairs *pê- -oht* and *pê- iT* indicate combinations of manner, direction, and motion, similar to preverb combinations; they are combined in *pê-itohtëw* ‘s/he comes walking’, a lexicalized compound found in dictionaries of Cree (e.g., Wolvengrey 2001). The root *pimât-*, found in the frequent VAI stems *pimâtisi-* ‘to be alive’ and *pimâciho-* ‘to live, travel, make living’, frequently occurs with *isi-* ‘thus’, indicating frequent comment on the ways in which people live or make a living in the corpus. Though these are not yet included in dictionaries as lexicalized compounds, this is the most strongly correlated preverb-stem morpheme in the corpus.

Overall Derivational Complexity

While preverbs and stem morphemes can co-occur quite freely with little discernable patterning, of greater interest is the overall numbers of derivational morphemes, both preverbs and stems, that may occur. In our corpus, we have found

TABLE 7. Most frequent preverb-stem morpheme combinations

MI	N ₁	N ₂	N ₁₊₂	PREVERB-STEM COMBINATIONS	PAIR GLOSSES
0.302369	541	860	65	pê- /-ê/	'come', AI
0.317487	541	808	62	pê- /-h/	'come', causative TA
0.902805	277	808	57	isi- /-h/	'thus, towards', TA
0.907752	541	325	45	pê- /-oht-/	'come', 'walk'
0.053292	541	628	37	pê- /iT-/	'come', 'thus, towards'
0.660967	229	808	37	nitawi- /-h/	'go', TA
-0.150912	541	687	33	pê- /-i/	'come', AI
0.773724	541	256	31	pê- /wît-/	'come', 'fellow'
-0.139707	541	597	29	pê- /-n/	'come', 'by hand'
0.158640	541	443	29	pê- /-amaw/	'come', benefactive TA
1.030786	124	808	29	kakwê- /-h/	'try', causative TA
-0.356561	541	716	28	pê- /-m/	'come', TA
-0.511510	541	836	28	pê- /-â/	'come', TI
1.606040	541	97	27	pê- /takw-/	'come', 'join'
2.033289	277	119	26	isi- /pimât-/	'thus, towards', 'life'

that up to three lexical preverbs occur (e.g., string *kâ-kî-pê-isi-kanâci-pimâtisit* 's/he has been living such a clean life'), excluding grammatical preverbs and the hesitation preverbal element *aya-*,⁶ and that lexemes with up to seven morphemes occur (e.g., *kâ-kî-wîci-kiskinohamâkosîmakik* ← *wîci- + /kiskinw- /+/-h /+/-amaw /+/-ikw- /+/-isi /+/-m /* 'the ones with whom I went to school').⁷ Moreover, the maximum number of preverbs plus stem morphemes that we have observed to occur is nine, as in the form *ê-kî-pê-isi-postayiwînisahisocik* 'y clothe themselves such', (*pê-*, *isi-*, and stem */post- /+ /ay- /+ /-iwi /+ /-n /+ /-is /+ /-ah /+ /-iso /*). Furthermore, we observed a weak but clearly significant inverse correlation $r_{pearson} = 0.040$, $p(t = -3.9498, df = 9981) = 7.874e^{-5}$, including both the number of lexical preverbs (including reduplicative elements) and the number of stem morphemes. The more lexical preverbs there are in a form, the slightly fewer stem morphemes one might expect to observe, and vice versa. This suggests, as predicted, an upper limit to the derivational complexity found in a verb form, and that the theoretical maximally complex verb in Plains Cree may be a relatively rare occurrence.

Discussion

Overall, the patterns seen in preverb and stem morpheme combinations align well with semantic descriptions of individual morphemes and with preverbal templates. For preverbs, the most prevalent pattern is the frequent co-occurrence of manner and direction preverbs with each other: to come toward or away from, to leave or go home. For derivational morphology, the most frequent co-occurrences create familiar verb classes and subtypes, as well as frequent stems themselves. Perhaps unsurprisingly, the only striking pattern among the preverb-stem morphology co-occurrences are between manner/direction and motion morphemes. Overall, preverbs demonstrate considerably more combinatory freedom than stem derivational morphemes, with fewer discernible patterns, highlighting their more syntactic nature.

Finally, looking beyond combinations, we investigated the total numbers for preverbs and stem morphology, both separately and together, with respect to derivational complexity. We found that lexical preverbs reach upper limits of three in a string, and that lexemes contain as many as seven morphemes, but when lexical preverbs appear, we see as many as nine derivational morphemes in total (divided into two preverbs and seven stem morphemes). Furthermore, the weak though significant inverse correlation between the number of preverbs and the number of morphemes in the stem serves to illustrate that Plains Cree verbs in everyday language use rarely, if ever, display the maximum complexity made possible by the rich morphological system of the language. Further investigation into this relationship is needed, and a larger corpus, especially including different registers, or field experimentation with native speakers, may also offer more data that will reveal stronger effects or different patterns entirely.

Conclusion

The complexity of Cree morphosyntax, in both inflection and derivation, has been extensively studied in the literature. Now, with the use of an analyzed corpus, quantitative investigations of previously described phenomena are possible. We have presented a preliminary quantitative investigation into derivational complexity in Cree verbs, including both lexical preverbs and stem derivation under the umbrella of derivation, in opposition to inflectional categories such as person,

number, and tense, among others. We have found some patterns, generally in line with previous preverbal templates, of direction and motion preverbs, and semantic links with motion medials and finals. However, the counts predominantly only reflect the overall frequencies of certain preverbs and stems containing certain morphemes; while perhaps not of great theoretical interest, such counts may be useful in classrooms, where more frequent elements may be taught before those less frequent, and more frequent sequences or combinations before those less frequent.

Our findings do, however, suggest an upper limit to derivational complexity in Cree verbs not reflected in typical theoretical templates; in the corpus, the maximum total found was nine morphemes, though theoretical descriptions allow for upward of fifteen morphemes including preverbs and the roots, medials, and finals of stem derivation (e.g., Bakker 2006; Wolvengrey 2012). Further data collection and analysis may shed light on the actual complexity of verbal derivation in spoken Plains Cree. Finally, though lexical preverbs and stem morphology both carry semantic information and contribute to lexeme formation, and these similarities have motivated the present study, lexical preverbs occur more freely. Thus, lexical preverbs and stem morphology are functionally distinct; these categories found in descriptive literature have been confirmed using quantitative methods.

NOTES

1. Abbreviations: II = inanimate intransitive verb, AI = animate intransitive verb, TI = transitive inanimate verb, TA = transitive animate verb, BENEF = benefactive, REFL = reflexive, TRZ = transitivizer, 3SG = third person singular, 3 = third person.
2. Though this degree of ambiguity is considerable, it has little bearing on the results presented here, and more sophisticated disambiguation methods are under development.
3. This is the initial morpheme in *itêw* 's/he says thus to someone' and *itwêw* 's/he says so'. This form is known as a relative root, in that it requires an antecedent (e.g., to go to a specified place, say a certain thing). This form is also in the particle *isi* and its preverb counterpart *isi-*, and as such is very common in our corpus.
4. The actual token frequency is often much higher, but we are using here form types as the basis of our co-occurrence stats (to focus rather on possible forms and to rule out the undue influence of a single form that has a high frequency). For example, there are 67 word tokens with the combination *pê-...-isi*, nearly twice the type frequency.
5. There is overlap here: *iT-oht-ê* (in, e.g., *itohtêw* '(s)he goes') is a common sequence of

- three derivational morphemes.
6. We have included *aya* 'um' as a preverb in our computational model to increase recognition rates, as it commonly occurs among preverbs. However, it is not included among our counts.
 7. This is counted as a single lexeme, *wíci-kiskinohamâkosím-*, which is a lexicalized compound. The preverb *wíci-* and the suffix *-m* form the comitative VTA derivation 'to do something with someone'.

REFERENCES

- Ahenakew, Alice. 2000. *âh-âyítaw isi ê-kí-kiskêyihthahkik maskihkiy / They Knew Both Sides of Medicine: Cree Tales of Curing and Cursing Told by Alice Ahenakew*, ed. by H. Christoph Wolfart. Winnipeg: University of Manitoba Press.
- Bakker, Peter. 2006. Algonquian Verb Structure: Plains Cree. *What's in a Verb? Studies in the Verbal Morphology of the Languages of the Americas*, ed. by E.B. Carlin and C.J. Rowicak, pp. 3–28. Utrecht: LOT.
- Bear, Glecia, Minnie Fraser, Irene Calliou, Mary Wells, Alpha Lafond, and Rita Longneck. 1992. *kôhkomínawak otâcimowinivâwa / Our Grandmothers' Lives as Told in Their Own Words*, ed. by Freda Ahenakew and H. Christoph Wolfart. Regina: Canadian Plains Research Center.
- Beesley, Kenneth R. and Lauri Karttunen. 2003. *Finite State Morphology*. Stanford, CA: CSLI Publications.
- Evert, Stefan. 2004. The Statistics of Word Co-occurrences: Word Pairs and Collocations. PhD thesis, University of Stuttgart.
- Harrigan, Atticus, Katherine Schmirler, Antti Arppe, Lene Antonsen, Sjur Moshagen, and Trond Trosterud. 2017. Learning from the computational modeling of the Plains Cree Verb. *Morphology* 27(4):565–598.
- Kâ-Nîpîtêhtêw, Jim. 1998. *ana kâ-pimwêwêhahk okakêskihkêmwina / The Counselling Speeches of Jim Kâ-Nîpîtêhtêw*, ed. by Freda Ahenakew and H. Christoph Wolfart. Winnipeg: University of Manitoba Press.
- Masuskapoe, Cecilia. 2010. *piko kikway ê-nakacihât: kékêk otâcimowina ê-nêhiawastêki*, ed. by H. Christoph Wolfart and Freda Ahenakew. Winnipeg: Algonquian and Iroquoian Linguistics.
- Minde, Emma. 1997. *kwayask ê-kî-pê-kiskinowâpatihicik / Their Example Showed Me the Way: A Cree Woman's Life Shaped by Two Cultures*, ed. by Freda Ahenakew and H. Christoph Wolfart. Edmonton: University of Alberta Press.

- Snoek, Connor, Dorothy Thunder, Kaidi Lõo, Antti Arppe, Jordan Lachler, Sjur Moshagen, and Trond Trosterud. 2014. Modeling the Noun Morphology of Plains Cree. Paper read at ComputEL: Workshop on the Use of Computational Methods in the Study of Endangered Languages, Fifty-Second Annual Meeting of the ACL, Baltimore, June 26.
- Vandall, Peter and Joe Douquette. 1987. *wâskahikanîwîyîniw-âcimowîna / Stories of the House People, Told by Peter Vandall and Joe Douquette*, ed. by Freda Ahenakew. Winnipeg: University of Manitoba Press.
- Whitecalf, Sarah. 1993. *kinêhiyawîwîniwaw nêhiyawêwin / The Cree Language Is Our Identity: The La Ronge Lectures of Sarah Whitecalf*, ed. by H. Christoph Wolfart and Freda Ahenakew. Winnipeg: University of Manitoba Press.
- Wolfart, H. Christoph. 1973. *Plains Cree: A Grammatical Study*. Transactions of the American Philosophical Society, n.s., vol. 63, part 5. Philadelphia.
- . 1996. Sketch of Cree, an Algonquian Language. *Handbook of North American Indians*, vol. 17: *Languages*, ed. by Ives Goddard, pp. 390–439. Washington, DC: Smithsonian Institution.
- Wolvengrey, Arok. 2001. *nêhiyawêwin: itwêwîna / Cree: Words*. Regina: University of Regina Press.
- . 2012. The Verbal Morphosyntax of Aspect-Tense-Modality in Dialects of Cree. Paper read at the 2012 International Conference on Functional Discourse Grammar, Ghent, June 8.
- . 2015. Preverb Combinations, Co-occurrences, and Sequences: Preliminary Findings from a Preliminary Plains Cree Corpus. Paper read at the Forty-Seventh Algonquian Conference, Winnipeg, October 22–25.