A Computational Architecture for the Morphology of Upper Tanana

Olga Lovick[1], Christopher Cox[2], Miikka Silfverberg[3], Antti Arppe[4], Mans Hulden[3]
First Nations University of Canada[1] Carleton University[2] University of Colorado[3]
University of Alberta[4]
Olga@lithophile.com, christopher.cox@carleton.ca, Miikka.Silfverberg@colorado.edu, arppe@ualberta.ca, mans.hulden@colorado.edu

Abstract

In this paper, we describe a computational model of Upper Tanana, a highly endangered Dene (Athabaskan) language spoken in eastern interior Alaska (USA) and in the Yukon Territory (Canada). This model not only parses and generates inflected Upper Tanana verb forms, but uses the language's verb theme category system, a system of lexical-inflectional verb classes, to additionally predict possible derivations and their morphological behavior. This allows us to model a large portion of the Upper Tanana verb lexicon, making it more accessible to learners and scholars alike. Generated derivations will be compared against the narrative corpus of the language as well to the (much more comprehensive) lexical documentation of closely related languages.

Keywords: Dene languages, inflection, derivation

1. Upper Tanana

Upper Tanana (ISO 639-3: tau) belongs to the Alaskan subgroup of the Northern Dene language family. Documentation of the language began in 1959. There are two published text collections (Tyone, 1996; David, 2017) as well as two corpora of unpublished texts collected by James Kari and Olga Lovick. The language is severely endangered with fewer than 50 speakers, most of them elderly.

The verbal morphology of Upper Tanana is typical of a Dene language (Rice, 2000), and often involves a complex interweaving of non-continuous lexical, derivational, and inflectional prefix sequences. Verb stems vary with aspect and mode in a synchronically opaque fashion. The verb can be represented as a template:

12	11	10	9	8	7	#	6	5	4	3	2	1	0	-1
Postpositional object	Postposition	Adverbial-derivational	Iterative	Incorporate	Distributive	– Disjunct boundary –	Pronominal	Qualifier	Conjugation	Mode	Subject	Voice/valence	Stem	Suffix

Table 1: The Upper Tanana verb, represented as a template

This templatic approach, albeit popular in Dene linguistics from its inception (cf. Goddard (1911; Hoijer (1945; Li (1946; Sapir and Hoijer (1967)) is not without challenges. It is common for the prefixes in Positions 1–6 in the above template to coalesce into one or two syllables, often through relatively opaque morphophonological processes. Templatic analyses also typically require large numbers of Ø-morphemes in analysis, even for comparatively simple verb forms, and rely heavily on complex systems of morpheme co-occurrence restrictions to prevent overgeneration. All of these factors contribute to the perception that Dene languages are inherently "difficult", which undermines both revitalization and computational modelling efforts.

A possible solution to this problem is to represent inflection in Dene languages paradigmatically as sets of phonologically precomposed portmanteau morphemes (cf. Arppe et al. (2017) and Cox et al. (2016)). This allows learners to learn sets of paradigmatic 'chunks' like the ones in Table 2, required by all verbs belonging to same lexical-inflectional class of "operative" verb themes.[1]

	Ø-Ipfv	aa-Pfv	Fut	Opt
1s	ag	ag	t-ag	og
2s	į̨l	į̨l	t-į̨l	ų̨l
3s	el	al	t-al	ul
1p	ts'-el	ts'-al	ts'&t-al	ts'-ul
2p	al	at	t-al	al
3p	h-el	h-et	h&t-al	h-ul

Table 2: Inflection pattern for operative verb themes with L-voice/valence marker

When learning a new verb that belongs to the operative class and thus requires the inflection pattern in Table 2, a learner only needs to know where in the lexical entry the inflectional chunk is inserted. This is illustrated in Table 3 for the verb theme ch'+L+dzüh 'dance' in the Perfective mode; the precomposed chunks from Table 2 are inserted following the *ch'*- prefix and before the stem. The remaining morphophonemic adjustments (in this example, in the first and third person plural) are relatively minor and quite regular, which facilitates both language learning and the computational modelling discussed below.

The basic lexical entry for verbs in Dene languages is known as a verb theme (Kari, 1979), and consists of a voice/valence marker and a stem, as well as zero or more re-

[1] The hyphen in all 1p and 3p forms indicates that morphological material may intervene here; the ampersand in the Future forms signals that additional morphophonemic adjustments have to be made.

	Surface form	Breakdown
1s	ch'agndzia'	ch'-ag-dzia'
2s	ch'įldzia'	ch'-įl-dzia'
3s	ch'aldzia'	ch'-al-dzia'
1p	nts'aldzia'	n-ch'-ts'al-dzia'
2p	ch'aldzia'	ch'-al-dzia'
3p	ch'ihaldzia'	ch'-h-al-dzia'

Table 3: *aa*-perfective of ch'+L+dzüh (*op.*) 'dance'

quired lexical prefixes. Verb themes fall into larger lexical-inflectional classes referred to in Dene linguistics as verb theme categories. These are characterized by shared conjugation markers in the Imperfective and Perfective, a shared primary aspect indicated by (historic) stem suffixation pattern, derivational potential, and semantic properties (durativity, telicity...). There are 10 verb theme categories, as shown in Table 4.

VTC	Example	Cnj	Aspect
Motion	Ø+haayh 's. arrive'	n, n	Momentaneous
Successive	O+Ø+got 'punch O'	Ø, aa	Durative
Operative	ch'+L+dzüh 'dance'	Ø, aa	Durative
Conversive	O+H+tsįį 'make s. O'	Ø, dh	Conclusive
Extension	Ø+'ah 'it extends'	n, aa	Neuter
Classific.	Ø+'ąą 'classify CO'	dh, aa	Neuter
Positional	Ø+dah 's. sit'	dh, aa	Neuter
Stative	H+ts'iik 'be sick'	dh, aa	Neuter
Dimens.	Ø+chaa 'be big'	Ø, aa	Neuter
Descriptive	Ø+łįį 'be'	Ø, aa	Neuter

Table 4: Verb theme categories in Upper Tanana

The verb theme category system is quite powerful at predicting a verb theme's behavior (Kari, 1979). By knowing that Ø+haayh 's. arrive' and O+Ø+'ąą 'handle compact O' belong to the motion theme category, we also know that they (1) require (n, n) conjugation markers, (2) have punctual, telic meaning, (3) require the Momentaneous Aspect in their primary derivation, (4) allow aspectual derivations in the Momentaneous, Perambulative, Reversative Aspects, (5) have enormous derivational potential (55 directional aspectual prefix strings plus many non-aspectual derivations), (6) allow the Progressive superaspect, (7) allow the Inceptive derivation, (8) restrict the Imperfective to second person subjects in most syntactic environments, and so on. As we discuss in 3., modelling verb theme categories allows much of the derivational morphology of Upper Tanana to be captured in a straightforward way, building on the representation of inflectional processes described in 2..

2. Computational model

Following prior work for related Tsuut'ina initially presented in Cox et al. (2016) and detailed in Arppe et al. (2017), we make use of finite state machines (FSMs) (see e.g. Beesley and Karttunen (2003)) as the computational formalism for implementing our model for Upper Tanana. FSMs have become one standard way for computationally modeling the morphological structure of words, and there are currently several open source implementations of FSM compilers, e.g. *xfst* (Beesley and Karttunen, 2003), *foma* (Hulden, 2009) and HFST (Lindén et al., 2011). The key advantages of FSMs are most crucially that they provide a calculus for powerful manipulations and are designed for rule-based definition of paradigms, which does not require large corpora from which to learn such rules, usually lacking for endangered languages. Furthermore, as well-established computational data structures, FSMs allow for easy integration with other software applications, for instance as spell-checking modules within word-processors, morphologically "intelligent" electronic dictionaries, and "intelligent" computer-aided language-learning applications. Here, we make use of the Giella infrastructure, developed by the Giellatekno and Divvun research teams at the University of Tromsø (Trosterud, 2006), which provide ready-made solutions for the integration of an FSM-based computational model as part of end-user applications that all types of "language workers" benefit from in practice.

While Upper Tanana verb structure is clearly complex and multimorphemic, it is possible to divide it into a more basic, three-zone structure, as has been common practice in many preceding descriptive studies of other Dene languages (cf. Kari (1975), Kari (1989)). The three primary zones—the disjunct domain at the left edge of the verb (template positions 12–7), the inner conjunct domain (positions 6–1), and the rightmost stem domain (position 0), cf. Table 5—are distinguished not only by their linear order within the verb, but also by a number of phonological and morphosyntactic criteria, including the set of phonological processes that are observed in each zone (Kari, 1975; Kari, 1989). These differences between zones have immediate consequences for computational modelling: using only general-purpose morphophonological adaptation rules that apply indiscriminately to all parts of the verb-word risks modifying prefix combinations in zones where those rules should not apply, and makes it difficult to model processes that operate only in one of these domains.

12	11	10	9	8	7	=	6	5	4	3	2	1	.	0	-1
Postpositional object	Postposition	Adverbial-derivational	Iterative	Incorporate	Distributive	=	Pronominal	Qualifier	Conjugation	Mode	Subject	Voice/valence	.	Stem	Suffix
Disjunct						=	Conjunct						.	Stem	

Table 5: The Upper Tanana verb, represented as a template

We thus follow Arppe et al. (2017) in including boundary symbols in our lexical representations to mark off the disjunct (=), conjunct (_), and stem (.) zones in our lexical entries, which can thus together be seen as defining a discontinuous lexical 'tier'. Where boundary markers are not explicitly indicated in a lexical entry, it is possible to conclude that no morphological material appears within the corresponding zone, and thus insert the missing boundary markers automatically. Thus, a lexical entry for a verb like na#D+kuyh 'vomit' (imperfective), which contains only the stem *kuyh* and a single lexical disjunct prefix *na-*, can be specified as na=kuyh, and subsequently expanded by the model automatically into na=_.kuyh. This automatic insertion of unspecified boundary markers serves to align the form of our lexical entries more closely with common lexicographic practice in Dene linguistics (e.g., Kari (1990), Jetté and Jones (2000), among others), while avoiding one source of possible human error when such boundary markers are added manually.

Automatically inserting disjunct, conjunct, and stem zone marker symbols into lexical entries also provides an opportunity to flag the presence or absence of morphological material in each of these zones for use in other parts of the model. As an example, when this model encounters the lexical entry na=kuyh 'vomit', it not only inserts the appropriate additional boundary markers, but also sets a corresponding flag diacritic (@P.PREFIX.OUTER@) to indicate that a disjunct prefix is present in this verb. This information is linguistically important, as the forms that inflectional morphology takes in Dene languages vary considerably based on the kind of prefixes that precede it. Using flag diacritics allows the model to insert the correct set of inflectional allomorphs for the observed prefixation condition (e.g., selecting the allomorphs of the Ø-imperfective paradigm that appear with a preceding conjunct prefix when the conjunct prefix flag diacritic is set, or the allomorphs of that same inflectional paradigm that appear with a preceding disjunct prefix when the disjunct prefix flag diacritic is set, etc.).

This approach to capturing recurring patterns of inflectional allomorphy associated with different prefixation conditions allows for considerable simplification of the final computational model. In the case of Upper Tanana, each inflectional paradigm (e.g., the *Ø*-imperfective mentioned above) can thus be represented as a finite number of sets of allomorphs that appear under particular prefixation conditions (e.g., when preceded by no other prefixes, or preceded by one or more conjunct prefixes, or by one more disjunct prefixes). Within each such prefixation condition, different phonological forms for each subject person-number combination are hard-coded into the model (e.g., the first-person singular form of the *Ø*-imperfective paradigm when preceded by a disjunct prefix). In practice, these inflectional 'chunks' are treated as portmanteau morphemes that combine not only subject person and number, but also the voice-valence markers found in position 1 of the verbal template given above. Thus, a lexical entry for a verb tagged as taking the *Ø*-imperfective paradigm, which has no preceding prefixes, and has an **l*- voice-valence marker (cf. 2) would follow a series of continuation lexica to finally arrive at the following set of inflectional chunks, which give the forms of the core inflectional markers that appear in this context:

```
LEXICON 0-IPFV-L-AFF-NoPrefix
@U.SBJPERSON.1@@U.SBJNUMBER.SG@ag      #;
@U.SBJPERSON.2@@U.SBJNUMBER.SG@įl      #;
@U.SBJPERSON.3@@U.SBJNUMBER.SG@el      #;
@U.SBJPERSON.1@@U.SBJNUMBER.PL@el      #;
@U.SBJPERSON.2@@U.SBJNUMBER.PL@al      #;
@U.SBJPERSON.3@@U.SBJNUMBER.PL@el      #;
@U.SBJPERSON.INDEF@@U.SBJNUMBER.SG@el #;
```

On this approach, lexical entries in the model can be represented parsimoniously as lexical combinations of prefixes and stems that are associated with a particular aspect (e.g., imperfective, perfective) and a lemma (in most cases, the third-person singular imperfective form), and subsequently 'tagged' through continuation lexica for their transitivity (e.g., transitive, intransitive, etc.), conjugation class and aspect (e.g., *n*-perfective, *aa*-perfective, etc.), and voice-valence marker (i.e., Ø-, **d*-, **l*-, or **ł*-). Whole verb themes, such as for the verbs 'vomit' and 'cry' below, are captured by listing the distinctive stem forms and inflectional patterns associated with each aspect in which this verb appears in the corresponding *lexc* definitions:

```
LEXICON VerbThemes
etsüh[cry]:tsüh             INTR-0-IPFV-0;
etsüh[cry]:tsia'            INTR-aa-PFV-0;
etsüh[cry]:tsüh             INTR-aa-FUT-0;
etsüh[cry]:tsüü             INTR-u-OPT-0;

na'etkuyh[vomit]:na=kuyh    INTR-0-IPFV-D;
na'etkuyh[vomit]:na=kuyh    INTR-dh-PFV-D;
na'etkuyh[vomit]:na=kuyh    INTR-aa-FUT-D;
na'etkuyh[vomit]:na=kuyh    INTR-u-OPT-D;
```

Many core features of the Upper Tanana verbal lexicon are thus captured through tag-like continuation lexica: the imperfective form of 'cry' is marked as being intransitive (INTR-) and taking the Ø-voice/valence marker allomorphs of the Ø-imperfective paradigm (0-IPFV-0) by the lexical entry proceeding into the continuation lexicon associated with these features (INTR-0-IPFV-0). These continuation lexica lead through a series of unification flags that constrain the range of inflected forms that are ultimately recognized and produced by the model, as in the example below:

```
LEXICON INTR-0-IPFV-0
@U.VALENCE.INTRANSITIVE@
@U.ASPECT.IPFV@@U.TAMA.%0@@U.VV.%0@ Verbsuffixes;
```

The overall morphological model for Upper Tanana can be seen as a concatenation of several smaller models such as these, interleaving a discontinuous 'inflectional tier' (the 'chunks' mentioned above as well as morphemes for direct and postpositional objects) with a potentially discontinuous lexical tier for the verb theme. In terms of our computational implementation, this is achieved by first compiling the specifications of the verb theme as well as the applicable morphemes or morpheme chunks in the three slots representing the inflectional tier, as four separate FSMs. Then, the three inflectional tier FSMs are inserted using finite-state calculus within the lexical tier FSM, by replacing the special characters (=), conjunct (_), and stem (.) marking the inflectional boundaries with the entire corresponding inflectional FSMs. Careful use of flag diacritics shared by all of these components ensures that only forms licensed by the morphology are generated and accepted by the resulting composite model (cf. Arppe et al. (2017)).

Despite the morphological complexity evidenced by the template presented in Table 1, this approach to representing the basic morphological 'scaffolding' of Upper Tanana verb forms appears to be adequate for the purposes of computational modelling. A small number of morphophonological rules apply to the model described above, handling regular processes of prefix vowel syllabification and epenthesis, as well as well-attested reductions in certain common prefix sequences. For example, the third-person plural subject marker *he-* may be realized (a) as *hih-* word-initially before /t, n, d/; (b) as *hi-* before /h/; or (c) as either *hi-* or *he-* (depending on dialect and/or speaker) when neither preceded nor followed by another vowel. These patterns can be modelled phonologically by treating the /e/ in this prefix as epenthetic, using the special character `E` to represent this vowel in the model (i.e., as `hE`). In the regular expression notation included in FST compilers (e.g. *foma*), these realizations can be captured through the following expressions:

```
def hEpenthesis E -> i h || .#. h _ [t|n|d];
def Epenthesis  E -> i || _ h         .o.
                E -> 0 || _ PrefixV   .o.
                E -> [i|e];
```

These few phonological rules, together with another set that deals with common reductions in prefix consonant sequences and instances of resyllabification, appear to be adequate in modelling much of the morphophonology associated with the 'core' verbal inflection implemented thus far.

3. Modelling verb themes

Utilizing the verb theme category system allows us to further increase the efficiency of the model. By tagging each verb theme for verb theme category, we predict much of its morphological behaviour, as described in section 1. This allows us to simplify the lexical entry for a verb theme. Instead of indicating that `na=kuyh` 'vomit' requires the Ø-Imperfective and *dh*-Perfective, we indicate that it is a member of the Conversive verb theme category, which is characterized i.a. by these two conjugation markers. This requires some adjustment to the structure of the lexicon, which would now be organized by verb theme category, transitivity, and voice/valence marker:

```
LEXICON Operative-Intransitive-0
etsüh[cry]:tsüh             IPFV ;
etsüh[cry]:tsia'            PFV ;
etsüh[cry]:tsüh             FUT ;
etsüh[cry]:tsüü             OPT ;

LEXICON Conversive-Intransitive-0
na'etkuyh[vomit]:na=kuyh    IPFV ;
na'etkuyh[vomit]:na=kuyh    PFV ;
na'etkuyh[vomit]:na=kuyh    FUT ;
na'etkuyh[vomit]:na=kuyh    OPT ;

LEXICON Motion-Intransitive-0
nihaayh[s.go]:haayh         IPFV ;
nihaayh[s.go]:shyah         PFV ;
nihaayh[s.go]:haał          FUT ;
nihaayh[s.go]:shya'         OPT ;
```

This approach has two advantages. First, it minimizes lexical data entry by allowing us to only list the information unique to each verb theme—its prefixes and aspectual stem variation. The remaining information—conjugation markers, transitivity, voice/valence-marker—are all given in the heading of each section of the lexicon (i.e., through these continuation lexica and associated flag diacritics that they set). Second, and more importantly, this approach allows us to model significant aspects of the derivational system of Upper Tanana. By marking each verb theme category and any associated derivational prefix strings with their own flag diacritic (e.g. `@U.VTC.OPERATIVE@` for the operative verb theme category), we ensure that these forms of derivation will occur with all and only the members of this verb theme. Since many derivational prefix strings select their own conjugation markers, this further allows the model to override the specification of a given verb theme. We illustrate this with an example from the Motion verb theme category.

If we flag a verb theme such as Ø+haayh 's. go' as a member of the Motion verb theme category, we indicate that it has certain semantic (+punctual, +telic situation type, Momentaneous aspect) and morphological (e.g. requiring *n*-Imperfective and *n*-Perfective) characteristics. Any Motion verb theme is however also compatible (theoretically) with all 55 derivational strings associated with Motion verbs, such as *ski-* 'across', *ki-* 'climbing', or *da+tl'a* 'up onto shore', as well as certain aspectual and superaspectual derivations, some of which are shown in (1). As the example shows, each derivational string is specified for conjugation marker and stem aspect, which overrides the primary theme's specifications. All forms in (1) are in the third person singular perfective; only (1f) is imperfective.

(1) a. primary theme (n, n), Momentaneous

nįįshyah

's/he arrived'

b. ski# (n, n), Momentaneous

skinįįshyah

's/he went across'

c. ki# (Ø, aa), Momentaneous

ki'įįshyah

's/he climbed up [e.g. a tree]'

d. da+tl'a# (Ø, dh), Momentaneous

datl'a'eeshyah

's/he went up onto the bank'

e. na#t+D (aa, ee), Perambulative

natetshyah

's/he was walking around (pfv.)'

f. – (aa, –), Progressive (ipfv. only)

aahaał

's/he is/was walking along'

By flagging both prefix strings and verb themes for verb theme category, we can harness the enormous derivational power of this system for our model. Since all prefix strings in (1) are available in principle to all motion verb themes, our model can correctly generate derived forms from basic themes as schematized in Table 6. Some semantic restrictions apply: while all of the forms in Table 6 are morphologically and semantically well-formed and are attested in Lovick's corpus, ki# 'up' cannot be combined with the verb theme Ø+mbeeyh 'swim'. The form *ki'įįmįį* 'she climbed up while swimming' generated by our model is not accepted by speakers, presumably for semantic reasons. Such exceptions will need to be specified manually.

Prefix string	L+tthiit 's. run'	Ø+mbeeyh 'swim'
ski	ski'iltthät	skinįįmįį
'across'	's/he ran across'	's/he swam across'
ki	ki'altthät	–
'up'	's/he ran up [a tree]'	–
da+tl'a	datl'a'eltthät	datl'a'eemįį
'onto bank'	's/he ran onto the bank'	's/he swam onto the bank'
na#t+D	nateltthät	natetmįį
'Peramb.'	's/he was running around'	's/he was swimming around'
–	altthäł	almbeeł
'Prog.'	's/he was running along'	's/he was swimming along'

Table 6: Derivation of motion verb themes

In total, there are fewer than 100 primary motion themes and about 55 derivational strings that can apply to motion themes. By flagging each primary theme as compatible with each derivational string, we allow our model to generate 5,500 derived verb themes: a sizeable portion of the Upper Tanana verb lexicon. We follow the same process for the other verb theme categories and their associated derivational strings.

The accuracy of this model can be evaluated by comparing the resulting forms against the Upper Tanana documentation record. Attested forms might be listed in the model and weighted above predicted forms. Given the relatively small size of the corpus, however—Lovick's narrative corpus comprises about 9,000 utterances—it is likely that our model will produce many forms that are not attested in the documentation. We will not be able to check all generated forms with a speaker, due to the severe endangerment of the language. Our solution to this is to compare predicted forms against their equivalents in related languages with substantially larger corpora and, importantly, larger lexical databases, such as Ahtna (Kari, 1990) or Koyukon (Jetté and Jones, 2000). Forms with cognates in those languages may then also be listed in our model, albeit with a lower weighting than forms attested in Upper Tanana. In this fashion, the richness of the inflectional as well as the derivational system of Upper Tanana can be made accessible to language learners even when they do not have access to a speaker.

4. Conclusion

The computational model of Upper Tanana presented here builds on and expands that designed by Cox et al. (2016) and Arppe et al. (2017) for Tsuut'ina, as part of a more general research program to model Dene and other Indigenous languages (cf. `altlab.artsrn.ualberta.ca`). Verb lexica structured by verb theme category, transitivity, and voice/valence marker allow our model to predict much of a verb theme's inflectional and derivational behaviour, which drastically reduces the amount of information that needs to be entered manually. By adding a derivational component to our model, we make the rich derivational system of this conservative Dene language accessible to language learners.

5. Acknowledgements

This work has been funded by the Social Sciences and Humanities Research Council of Canada (SSHRC) through a Connections Outreach Grant (611-2016-0207) and a KIAS Research Cluster Grant during 2015-2018 (University of Alberta). Lovick's research on Upper Tanana grammar has been funded as part of her International Polar Year Postdoctoral Fellowship at the University of Alaska Fairbanks (2006-2008) as well as by a Documenting Endangered Languages grant from the National Science Foundation (2012-2015, held jointly with Dr. Siri Tuttle).

6. Bibliographical References

Arppe, A., Cox, C., Hulden, M., Lachler, J., Moshagen, S. N., Silfverberg, M., and Trosterud, T. (2017). Computational modeling of the verb in Dene languages. the case of Tsuut'ina. In *Working Papers in Athabascan Linguistics ("Red Book" series)*, Fairbanks. Alaska Native Language Center.

Beesley, K. R. and Karttunen, L. (2003). *Finite State Morphology*. CSLI Publications, Stanford, CA.

Cox, C., Hulden, M., Silfverberg, M., Lachler, J., Rice, S., Moshagen, S. N., Trosterud, T., and Arppe, A. (2016). Computational modeling of the verb in Dene languages. the case of Tsuut'ina. In *Dene Languages Conference, Yellowknife, Northwest Territories, Canada, 6–7 June 2016*. http://altlab.artsrn.ualberta.ca/wp-content/uploads/2016/07/Tsuutina-FST-DLC-160607B.pdf.

David, C. (2017). *Teedlęy t'iin naholndak niign: Stories by the Tetlin people. Expanded edition. Edited by Olga Lovick*. Alaska Native Language Center, Fairbanks.

Goddard, P. E., (1911). *Athapascan (Hupa)*, pages 85–158. Number 40 in Smithsonian Institution Bureau of American Ethnology Bulletin. Government Printing Office, Washington, DC.

Hoijer, H. (1945). The Apachean verb, part I: Verb structure and pronominal prefixes. *International Journal of American Linguistics*, 11(4):193–203.

Hulden, M. (2009). Foma: a finite-state compiler and library. In *Proceedings of EACL*, pages 29—32, Athens. Association for Computational Linguistics.

Jetté, J. and Jones, E. (2000). *Koyukon Athabaskan Dictionary*. Alaska Native Language Center, Fairbanks.

Kari, J. (1975). The disjunct boundary in the Navajo and Tanaina verb prefix complexes. *International Journal of American Linguistics*, 41(4):330–345.

Kari, J. (1979). *Athabaskan verb theme categories: Ahtna*. Alaska Native Language Center Research Papers, Number 2. Alaska Native Language Center, Fairbanks.

Kari, J. (1989). Affix positions and zones in the Athapaskan verb complex: Ahtna and Navajo. *International Journal of American Linguistics*, 55(4):424–454.

Kari, J. (1990). *Ahtna Athabaskan Dictionary*. Alaska Native Language Center, Fairbanks.

Li, F.-K. (1946). Chipewyan. In Cornelius Osgood et al., editors, *Linguistic structures of native America*, pages 398–423. The Viking Fund publications in anthropology No. 6, New York.

Lindén, K., Axelson, E., Hardwick, S., Silfverberg, M., and Pirinen, T. (2011). HFST - framework for compiling and applying morphologies. In *Proceedings of the Second International Workshop on Systems and Frameworks for Computational Morphology (SFCM)*, pages 67–85. Springer, Berlin, Heidelberg.

Rice, K. (2000). *Morpheme order and semantic scope*. John Benjamins, Amsterdam, New York.

Sapir, E. and Hoijer, H. (1967). *The Phonology and Morphology of the Navaho Language*, volume 50 of *University of California Publications in Linguistics*. University of California Press, Berkeley, CA / Los Angeles, CA.

Trosterud, T. (2006). Grammatically based language technology for minority languages. In *Lesser known languages of South Asia*, pages 293—316. Mouton de Gruyter, The Hague.

Tyone, M. (1996). *Ttheek'ädn Ut'iin Yaaniidą' Oonign': Old Time Stories of the Scottie Creek People. Stories told in Upper Tanana Athabaskan by Mary Tyone, Ts'ą' Yahnik.* Alaska Native Language Center, Fairbanks.