

Developing a Universal Dependencies Treebank for Alaskan Gwich'in

Matthew Andrews
University of Tübingen
m.kirkandrews@gmail.com

Çağrı Çöltekin
University of Tübingen
cagri.coeltekin@uni-tuebingen.de

Abstract

This paper presents a Universal Dependencies (UD) treebank of Gwich'in, a severely endangered Athabascan language. The treebank, developed using instructional materials and dictionaries, includes 313 annotated sentences. This paper discusses the methodology used to construct the treebank, the linguistic challenges faced, and the implications of annotating a polysynthetic, morphologically complex language within the Universal Dependencies framework. The treebank was released with UD version 2.15 and available at https://github.com/UniversalDependencies/UD_Gwichin-TueCL/.

1 Introduction

We present a treebank of Gwich'in annotated following the Universal Dependencies framework (de Marneffe et al., 2021). To our knowledge, this is the first treebank of the Athabascan language family, a language family spread over a large part of North America. Recent efforts in the Universal Dependencies community have increased coverage of the world's languages substantially. As of the most recent data release, there are 319 treebanks of 179 different languages from 62 different language groups (Zeman et al., 2025). However, excluding the present effort, only one language indigenous to the United States or Canada is included in the database. Despite the growing interest in Universal Dependencies and in the development of tools and resources for low-resource languages among the NLP-community, languages from this linguistically diverse region of the world remain underrepresented among UD treebanks.

Since the early 1990s, linguists have been concerned with the issue of language endangerment. Athabaskanist Michael E. Krauss was one of the first to address this problem in his classic article "The World's Languages in Crisis", in which he compared the loss of linguistic diversity to the loss

of biodiversity (Krauss, 1992). In order to draw global attention to the critical situation of language endangerment, the United Nations General Assembly has declared the period between 2022 and 2032 the International Decade of Indigenous Languages (United Nations General Assembly, 2019). In line with this spirit and in the spirit of the Universal Dependencies project's commitment to linguistic diversity we present a small manually annotated treebank for Gwich'in, a critically endangered language with interesting linguistic properties, including polysynthesis and classificatory verbs.

Following an introduction of the Gwich'in language, this paper will present and discuss the methodology used and some of the linguistic challenges faced in the development of this treebank.

2 Background

Gwich'in is an indigenous language of North America spoken by the Gwich'in people, also known as *Dinjii Zhuh*, who live in what is now the US state of Alaska, Yukon and the Northwest Territory in Canada. Historically, all *Dinjii Zhuh najji* could speak their language, *Dinjii Zhuh k'yàa*, but today, no more than 300 of the 4000 *Dinjii Zhuh najji* can, most of whom are elders (Mishler and Frank, 2019). Due to colonial pressure on the *Dinjii Zhuh najji* and other indigenous peoples to give up their native tongues, parents are no longer able to speak *Dinjii Zhuh ginjik* to their *tr'iinin najji* ('children') at home.

Despite these challenges, there is a vibrant Gwich'in community interested in learning and revitalizing their language. Currently, Gwich'in is taught as a course at the University of Alaska-Fairbanks and in some K-12 schools in Canada. In recent years, some language learning products have also been developed, including books (Gwich'in Social and Cultural Institute, n.d.a), online dictionaries (Gwich'in Social and Cultural In-

stitute, n.d.b), apps, a language learning platform (Doyon Foundation, n.d.), and Gwich'in language episodes of the Emmy award winning PBS Kids show "Molly of Denali" whose main character Molly is part Gwich'in Athabascan.

Besides these few resources, there is very little documentation and description of the Gwich'in language. The vast majority of data and information on the language consists of scans of handwritten field notes, classroom worksheets, and unannotated audio recordings housed in the Alaska Native Language Archive (Alaska Native Language Archive, n.d.). From the perspective of a computational linguist, these resources in their current form are difficult to use for the development of language technology or a so-called "NLP-toolkit" (Zariquiey et al., 2022).

In addition, few academic works have been published on the linguistic aspects of the Gwich'in language. Publications on neighboring Athabascan languages such as Koyukon and Upper Tanana are much easier to find (Axelrod, 1990; Lovick, 2020, 2023). Moreover, the majority of research on these languages and on Athabascan in general has focused more on the rich morphology of the language, particularly the verbal morphology, than on any other area of linguistic study. To date, the syntactic structure of Athabascan has received far less attention. The present study hopes to address these gaps.

3 Methodology

Text Source Materials referenced for the development of the Gwich'in treebank come from the Alaska Native Language Archive of the Alaska Native Language Center at the University of Alaska-Fairbanks. These materials consisted of scans of previous classroom handouts developed for university language courses (Marlow and Garnett, 1996) as well as English to Gwich'in dictionaries (Peter, 1979; Mueller and Garnett, March 1994). The classroom handouts included verb paradigms and short sentences which were selected for annotation in the Universal Dependencies framework. These materials also provided detailed grammatical descriptions of many linguistic aspects of Gwich'in. Select sentences with interlinear glosses were also chosen from two Master's theses on Gwich'in directionals (Busch, 2000) and classificatory verbs (Bushey, 2021). The treebank focuses exclusively on the Gwich'in variety spoken in Alaska.

Approach At a minimum, sentences in Gwich'in consist of a single morphologically complex verb form. A verb form consists of a verb stem, a subject prefix belonging to one of four classifiers, and zero or more inflectional and derivational prefixes. Verb stems are encoded for tense/aspect, and subject prefixes must agree in person and number as well as with the verb stem's classifier and tense/aspect. Information on a word's part of speech and lemma or dictionary form were annotated as were basic dependency relations between words. English translations, morpheme segmentation, and detailed interlinear glosses were included as well. The annotation tool UD Annotatrix (Tyers et al., 2017) was used to streamline annotations as was a Gwich'in dictionary web application developed by the first author to help determine word lemmas.

Corpus The treebank consists of a total of 313 annotated sentences. Approximately two-thirds of the corpus consists of one- to two-word sentences highlighting several different verb paradigms with a verb in the imperfective, perfective, and future tenses and zero or one noun objects. The remainder of the corpus consists of longer, more complex multi-word sentences covering a majority of part-of speech tags and dependency relations of the Universal Dependencies framework.

Language-specific features¹ include the feature Classifier and its values \emptyset , L, Ł, and D (See Appendix A for the descriptions). The name for this feature and its values are consistent with the Athabascan literature. A Gwich'in verb stem takes one of these four classifiers, each of which has its own set of subject prefixes. Compare the verb *yahtsii* 'he or she is making it' which has the verb stem *-tsii* and the subject prefix *-ah-* with the verb *yitsii* 'he or she is caching it' which has the same stem but takes the subject prefix *-i-*. The verb stems are the same, but the third-person subject prefixes are different as these verbs have different classifiers.

The language-specific feature StemClass and its values are shown in Appendix A. This feature is used to account for the noun class to which a Gwich'in classificatory verb stem belongs. This and other language-specific challenges are covered in the next section.

A Note on Tone Gwich'in is a tonal language. Low-tone, rising-falling, and falling-rising are

¹Morphological features are currently omitted from the treebank but will be provided in future releases.

Disjunct					Conjunct						Stem
0	1a	1b	2	3	4	5	6	7	8	9	10
postposition object	"null post- position"	adverbial- thematic	iterative	plural	direct object	deictic	adverbial- thematic	mode- aspect	subject	classifier	stem

Table 1: Verb template for Navajo, a Southern Athabascan language.

marked. There is much inconsistency of the marking of tone among the sources referenced in the creation of this treebank. Some sources mark tone, while others do not. Sentences in this treebank are transcribed exactly as they appear in the source text. As such, the same word may appear different elsewhere in the treebank although it is not, one with tone marking and one without.

4 Language-specific Features and Challenges

4.1 Verbs

Gwich'in like all Athabascan/Na-Dene languages is well known for its rich verbal morphology. Gwich'in is a polysynthetic, primarily prefixing, head-final language with the verb stem occupying the last position in a verb and the verb typically occurring as the last element in a clause. An Athabascan verb is composed of a verb stem and a set of prefixes. This is often represented in the Athabascan literature with a verbal template as in the one for Navajo shown in Table 1. There are 12 positions depicted in this template; however, not all positions are necessarily filled, and those that are are filled in a discontinuous manner. The maximum number of occupied slots found in the present treebank is seven.

The exact makeup of a (verb) word cannot be determined from its surface form alone, as the individual morphemes that make up the word are obscured by a number of morphophonological processes (Hale, 2003). This is illustrated for the verb *Giyahaqah'yaa* 'They will see it' in example 1 below.

- (1) *gī- yī- hī- nī- ah-*
 PL 3O FUT THEM 3SG+L+FUT
 'yaa
 see+FUT
 'They will see it'

The verb stem *-yaa* 'will see' is part of the *Ł* classifier and so takes the subject prefix *-ah-* for third person from the set of subject prefixes for the *Ł* classifier. Each of the four verb classifiers in

Gwich'in has its own distinct set of subject prefixes. The subject prefix is always closest to the verb stem. The next prefix *-nī-* is a thematic prefix and together with the verb stem and classifier forms what is called the verb theme. Then we have *-hī-*, a prefix that indicates future tense or aspect, preceded by *-yī-* to indicate the third person direct object of a third person subject. Third person marking in Gwich'in and Athabascan languages is quite complex and is the subject of a major theoretical question which we will return to below. Finally, the prefix *-gī-* occupies the position for plurality and is used for third person subjects.

Notice that many of these prefixes undergo a vowel change in their surface form. The use of *i* in these prefixes is to indicate the underlying vowel phoneme /i/ that is subject to vowel harmony with the verb stem and that surfaces as [a] in the presence of [a], [o], and [w]. However, *gī-* does not undergo this change, as plural morphemes belong to the disjunct zone of the verb template as seen in Table 1. All other prefixes belong to the conjunct zone. Only prefixes of the conjunct zone are subject to phonological changes.

The deletion of /n/ of the thematic prefix and the nasalization of adjacent vowels can also be seen in the surface form. This phonological process occurs when the thematic prefix is preceded by another prefix. A second phonological rule is applied to the vowel of the future prefix *-hī-*, where the vowel is deleted when the following vowel is long.

4.1.1 Classificatory Verbs

The noun classification system of Gwich'in and other Athabascan languages is quite elaborate. Bushey (2021) identifies nine classes for noun class membership, so-called 'stem-classes' in the Athabascan tradition. However, in Athabascan languages, noun class membership is not indicated on the noun but rather on the verb.

So-called 'classificatory verbs' are a hallmark of every Athabascan language including Gwich'in. Classificatory verb stems are unique in the sense that they classify a particular noun by its physical characteristics, e.g., shape, as well as describe the movement, handling, or position of the noun.

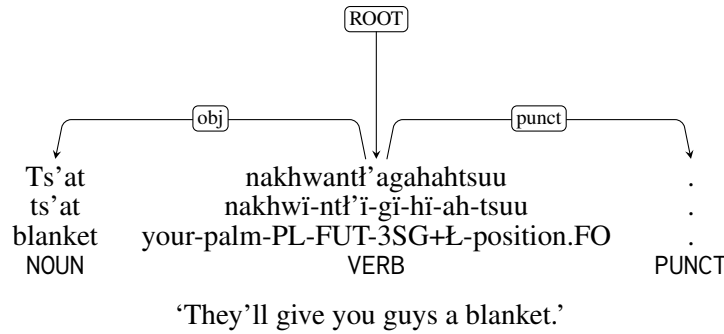


Figure 1: Verb stem *-tsuu* meaning ‘give soft, fabric object’.

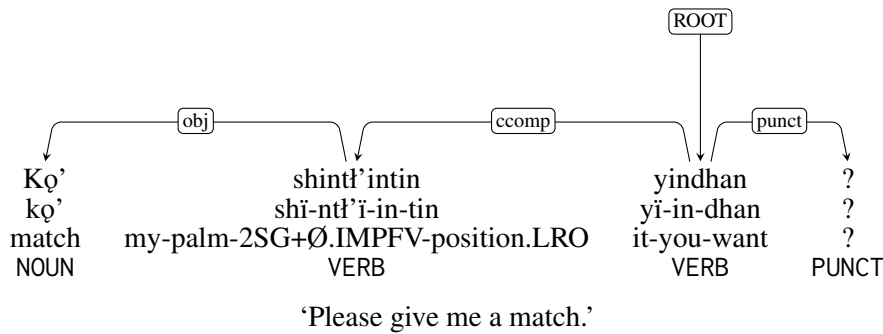


Figure 2: Verb stem *-tin* meaning ‘give long, rigid object’.

This phenomenon is somewhat similar to the English verb *roll* which is typically used only with round objects (Lovick, 2023). However, classificatory verbs are much more common and complex in Gwich'in than in English. Take for instance the verb stem *-tsuu* seen in Figure 1 which together with its classifier \mathbb{L} can be glossed as ‘position cloth-like object.’ This particular verb may be used with fabric objects (FO) like shirts and blankets. Compare this with the verb stem *-tin* found in Figure 2, which is used with stick-like or long, rigid objects (LRO) such as matches or guns. Notice how both of these verbs are translated with the English word ‘give.’ Although the complete details of classificatory verbs are beyond the scope of this paper, it is worth noting that nine stem classes have been identified for Gwich'in. Put another way, there are potentially nine different Gwich'in words for the English word ‘give.’² This is also true for other verbs having to do with the movement, handling, and positioning of objects, e.g., ‘fall’, ‘carry’, ‘put’, etc.

A list of suggested values for the feature StemClass to account for the noun classification system of Gwich'in can be found in Appendix A

²In actuality, only eight of these are attested in the data collected by Bushey (2021). A verb for the stem class of deteriorated objects is unattested in the data.

4.1.2 (Pro)nominal Incorporates

As seen in example 1 above, the verb word functions as a complete sentence, with both subject and object pronominals marked on the verb. The status of these markers has been the subject of considerable theoretical debate. When compared with pro-drop languages like Italian and Spanish, it is often assumed that Gwich'in and other Athabaskan languages also feature null arguments.

The Pronominal Argument Hypothesis (PAH) offers an alternative view, arguing that pronominal markers are not simply agreement markers but core arguments of the sentence (Jelinek, 2014; Hale, 2003). Free noun phrases, when they do occur, are treated as adjuncts. This suggests that, unlike in pro-drop languages, there are no null arguments; instead, arguments are marked directly on the verb (Gelderen and Willie, 2012).

In Southern Athabaskan languages such as Navajo, both subject and object pronominals must be marked on the verb. This contrasts with Northern Athabaskan languages like Gwich'in, where nominal objects are in complementary distribution. Specifically, a nominal object can either be a free noun phrase or expressed as a pronominal marker, but not both at the same time (Gelderen and Willie, 2012). For instance, when *dinjik* ‘moose’ is present

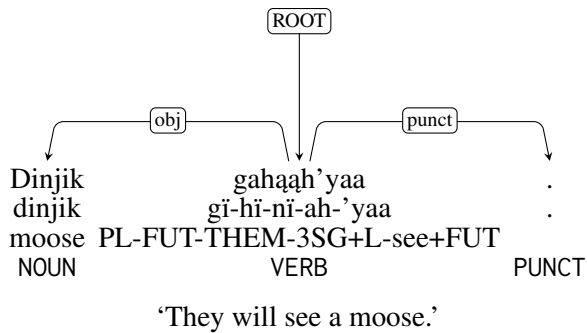


Figure 3: Verb with free noun phrase and no pronominal object marker.

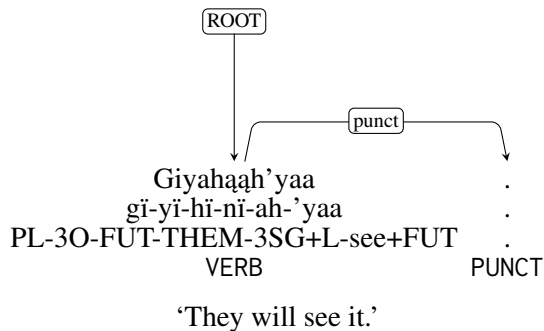


Figure 4: Verb with pronominal object marker and no free noun phrase.

in Figure 3, the pronominal object marker *-yi-* is absent, whereas in Figure 4, when the free noun phrase is not present, the pronominal marker appears.

This presents challenges for annotation schemes that require core arguments to be lexical (i.e., separate words), as stipulated by the Universal Dependencies (UD) framework. Although UD allows subwords, or ‘syntactic words’ as units of analysis, word segmentation for Gwich’in is non-trivial. Similar concerns have been raised by others working on Athabascan and other polysynthetic languages, and multiple proposals for segmenting verbs in polysynthetic languages exist (Spence et al., 2017; Park et al., 2021; Tyers and Mishchenkova, 2020). The present treebank does not introduce word segmentation, as illustrated in Figures 3 and 4. However, we include detailed glosses to help future revisions that would segment the complex words following a consensus between the alternatives.

Another distinction between Southern and Northern Athabascan languages is the phenomenon of

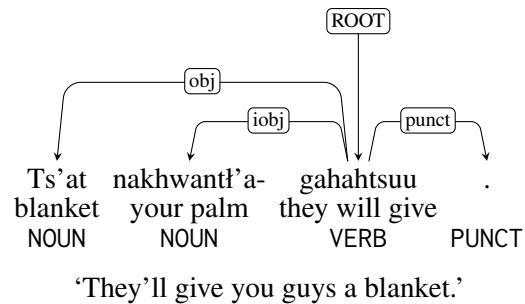


Figure 5: An alternative annotation splitting the complex verb (cf. Figure 1).

noun incorporation. It has been suggested that northern Athabascan languages allow a noun to be incorporated as part of the verb word, while Southern Athabascan languages do not. This may be because of the obligatory pronominal object marker on the verb in Southern Athabascan languages (Gelderen and Willie, 2012). The linguistic data used for this treebank does not appear to have incorporated direct objects on the verb; however, given the limited data and Gwich’in proficiency, we do not make a claim to the status of incorporated direct objects in Gwich’in. However, the corpus does contain example sentences with incorporated noun recipients in ditransitive verb themes. This is well attested and documented in neighboring languages such as Upper Tanana (Lovick, 2020, 2023). For instance, the verb in Figures 1 and 2 has the incorporated possessed noun ‘palm’ to indicate the recipient of the fabric object. A literal translation given for this sentence would be “They will position a blanket at rest in all your hands” (Bushey, 2021). Two separate annotations for this sentence can be seen in Figure 1 and Figure 5, one with and one without splitting the verb.

4.2 Postpositions

Adpositions follow their complement in Gwich’in. There is a three-way distinction in the use of these postpositions depending on whether the object of the postposition is a non-human, human, or areal noun (Marlow and Garnett, 1996). In other words, postpositions must agree with their object. If the object is human, the postposition is inflected for person and number. The pronominal prefixes used for this are identical to the direct object and possessor prefixes. If the object is a noun referring to an area, the postposition takes the areal prefix

gwī-. Postpositions that modify all other nouns are unmarked. An example of postpositions with non-human, human, and areal objects is shown in Figures 6, 7, and 8 respectively, where the postposition *zhit* ‘inside’ is compared in all three instances.

Objects that are personified in myth and folklore may be treated as human as demonstrated in Figure 9.

When a non-areal object is unspecified in the sentence but understood as in the sentence in Figure 10, the postposition takes the third person object prefix. This can result in an unconventional dependency relation between the predicate of the sentence and the postposition of the unspecified object as is shown in Figure 10. The dependency relation oblique is normally used for a nominal argument or adjunct of a verb, adjective, or other adverb, but here it is used for the adposition of a dropped pronoun object. The syntactic analysis of Figure 7 and the relation between the root verb and its dependent is therefore preserved when the nominal argument is omitted as in Figure 10. Notice the similarity and difference between Figure 7 and Figure 10 when the object of the postposition is dropped.

4.3 Lemmas

All Athabascan verb stems are themselves morphologically complex, which consist of a verb root and aspectual suffixes. The root provides information about the meaning of a verb, while the aspectual suffixes provide information about its tense. In many Athabascan languages, the root and the aspectual suffixes have fused so tightly together that the same verb has different stems for each aspect (Lovick, 2020).

Take for instance the verb sentence *Ch’qahtsii* ‘She is picking berries’ and *Ch’iitstaii* ‘She picked berries’ which can be seen in example 2 and example 3 below. This verb takes the classifier Ł subject prefixes and has the stem *-tsii* in the imperfective. In the perfective, the verb has the stem *-tsaii*. Some verbs, including this one, also take the prefix *dhi-* in the perfective. Because of this, the third-person singular subject prefix surfaces as *-it* and not *-ah* in the perfective.

- (2) *ch’i-* *nī-* *ah-*
 3O.INDF THEM 3SG+Ł
tsii
 pick-berry.IMPFV
 ‘She is picking berries’

- (3) *ch’i-* *nī-* *dhi-* *it-*
 3O.INDF THEM CNJ 3SG+Ł
tsaii
 pick-berry.PFV
 ‘She picked berries’

However, not all Gwich’in verbs that take the *dhi-* prefix do so in the perfective; some take it in the imperfective. No verbs take it in both.

The future *Ch’ahqahtsyàa* ‘She will pick berries’ has the stem *-tsyàa*, as seen in example 4.

- (4) *ch’i-* *hī-* *nī-* *ah-*
 3O.INDF FUT THEM 3SG+Ł
tsyàa
 pick-berry.FUT
 ‘She will pick berries’

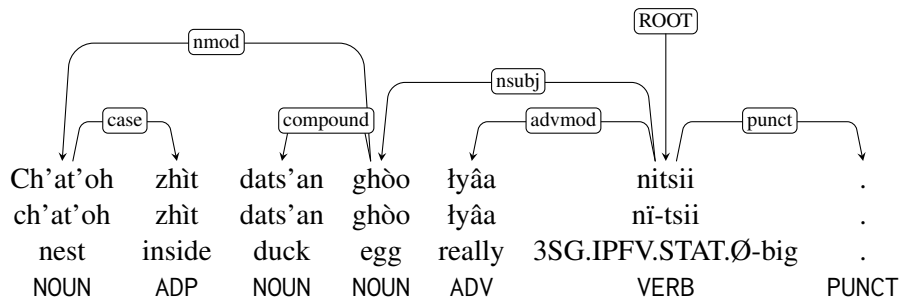
Given the morphological complexity of the verb in Gwich’in described throughout, the task of determining lemmas is not an easy one. Native speaker and teacher Lillian Garnett describes this challenge as follows: “While there are some basic tendencies, there are no rules to predict how the stem will change. All we can do is memorize the past stem (along with the present and future stems) as part of the word” (Marlow and Garnett, 1996).

Previous lexicographic work for Gwich’in uses third-person singular forms for all prefixes for all dictionary entries. Both imperfective and perfective forms are usually given. We follow this tradition of using third-person singular forms for lemmas in this treebank. This includes all parts of speech that require person inflection such as verbs and inalienable nouns. Following the Universal Dependencies framework, we also use one canonical form for each word. However, changes may need to be made to account for the verb stem variation characteristic of Gwich’in and other Athabascan languages described above.

5 Conclusion, Limitations and Future Work

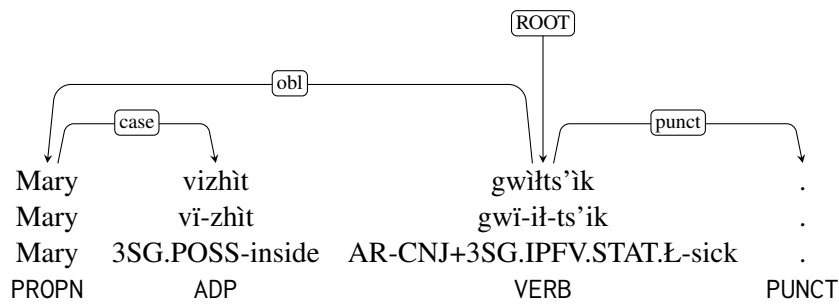
The development of the UD_Gwichin-TueCL treebank is an important step towards expanding the representation of diverse languages in the Universal Dependencies Framework. However, several limitations need to be acknowledged.

First, the corpus is relatively small, with only 313 annotated sentences. This may limit the scope of syntactic and morphological phenomena represented in the treebank. While we attempted to



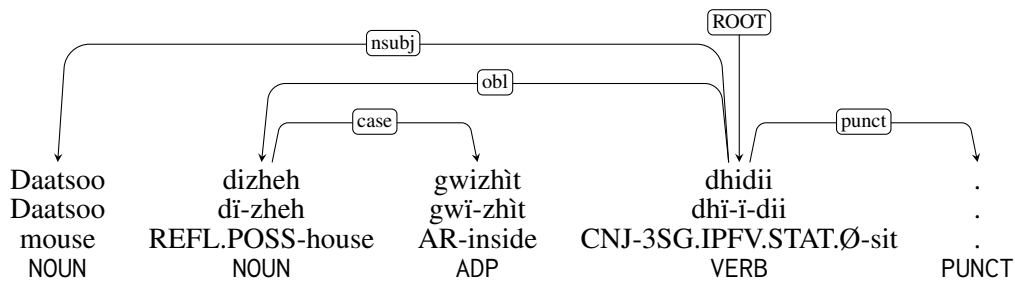
'The duck eggs in the nest are really big.'

Figure 6: Postposition *zhìt* 'inside' with inanimate object.



'Mary is hungry.'

Figure 7: Postposition *zhìt* 'inside' with human object.



'Mouse is inside her house.'

Figure 8: Postposition *zhìt* 'inside' with animate (non-human) object.

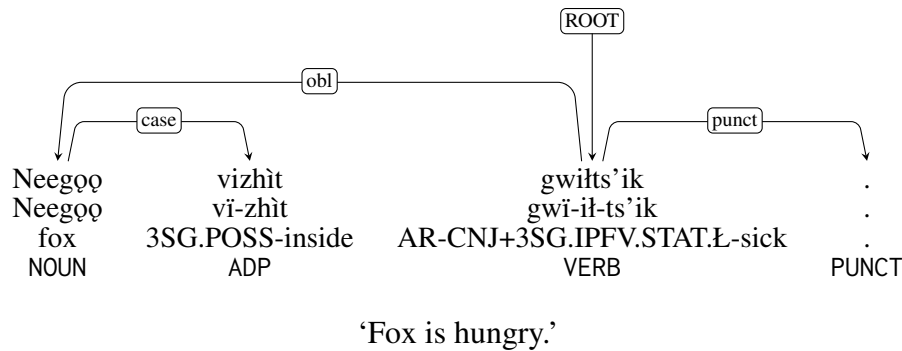


Figure 9: Human-form of postposition used with non-human object because of personification.

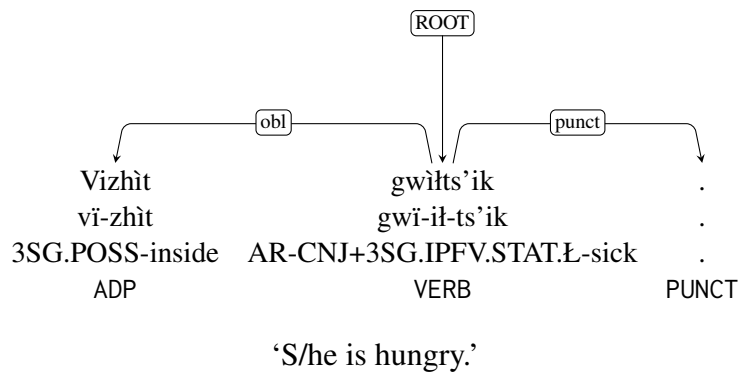


Figure 10: Postposition is promoted due to 'elided' noun phrase object.

include a diverse range of phenomena in the treebank, the scarcity of linguistic data and information on the language made this a challenging task. Future expansions of the corpus will benefit from additional data sources including native speakers.

Second, the polysynthetic and morphologically complex nature of the language poses challenges for annotation. In particular, the widely held theoretical position among Athabaskanists is to treat pronominal prefixes as core arguments of the verb (Jelinek, 2014; Hale, 2003; Lovick, 2023). Combined with challenges of determining morph(eme) boundaries, this poses some challenges to the UD framework, which requires core arguments to be (syntactic) words. As more North American languages are annotated in UD, it remains to be seen whether the annotation scheme would need major changes to adequately describe languages with verb-internal syntax such as Gwich'in.

Third, the present treebank was annotated by a non-native second language learner of Gwich'in. While every effort was made to ensure accurate annotations, the absence of native speakers in the construction of the treebank may result in inaccuracies and incomplete representations of the language. Furthermore, annotation decisions are those

of a single annotator. Therefore, inter-rater agreement scores cannot be obtained at this time. Future work on the language will benefit from collaboration with native speaker consultants as well as additional annotators.

This treebank was developed to contribute to the Universal Dependencies project as well as the study of Athabaskan morphology and syntax. We also hope that it will be instrumental for the Gwich'in community in their efforts to develop tools that support language revitalization. It is also hoped that this treebank will not only engender interest in Athabaskan languages and North American Indian languages more broadly in the field of Natural Language Processing, but also that Athabaskanists and others working with endangered languages will be encouraged to contribute more treebanks to the Universal Dependencies project in the future.

References

- Alaska Native Language Archive. n.d. [Gwich'in collection - Alaska Native Language Archive \(ANLA\)](#). Accessed: 15-Oct-2024.
- Melissa Axelrod. 1990. Incorporation in Koyukon Athabaskan. *International Journal of American Linguistics*, 56(2):179–195.
- John Busch. 2000. [Finding your way through a story: Direction terms in Gwich'in narrative](#). Research paper (M.A. in Anthropology), University of Alaska Fairbanks. Accessed: 15-Oct-2024. 42 pp. in English with some Gwich'in vocabulary.
- Scott T Bushey. 2021. *Western Gwich'in Classificatory Verbs*. University of Alaska Fairbanks.
- Marie-Catherine de Marneffe, Christopher D. Manning, Joakim Nivre, and Daniel Zeman. 2021. [Universal dependencies](#). *Computational Linguistics*, 47(2):255–308.
- Doyon Foundation. n.d. [Doyon languages online](#). Accessed: 15-Oct-2024.
- Elly van Gelderen and Mary Willie. 2012. Are there null arguments in Athabaskan? *Acta Linguistica Hafniensia*, 44(2):227–245.
- Gwich'in Social and Cultural Institute. n.d.a. [Gwich'in language store](#). Accessed: 15-Oct-2024.
- Gwich'in Social and Cultural Institute. n.d.b. [Gwich'in online dictionary](#). Accessed: 15-Oct-2024.
- Kenneth Hale. 2003. On the significance of Eloise Jelinek's pronominal argument hypothesis. *Formal approaches to function in grammar: In honor of Eloise Jelinek*, pages 11–44.
- Eloise Jelinek. 2014. Empty categories, case, and configurationality. In *Pronouns, Presuppositions, and Hierarchies*, pages 7–40. Routledge.
- Michael Krauss. 1992. The world's languages in crisis. *Language*, 68(1):4–10.
- Olga Lovick. 2020. *A grammar of Upper Tanana, volume 1: Phonology, lexical classes, morphology*. University of Nebraska Press.
- Olga Lovick. 2023. *A grammar of Upper Tanana, Volume 2: Semantics, Syntax, Discourse*, volume 2. U of Nebraska Press.
- Patrick Marlow and Lillian Garnett. 1996. Beginning Athabaskan Gwich'in ANL142. <https://www.uaf.edu/anla/record.php?identifier=KU985G1996b>. Accessed: October 22, 2023.
- C. Mishler and K. Frank. 2019. *Dinjii Vadzaih Dhidlit*. IPI.
- Dick Mueller and Lillian Garnett. March 1994. [Western Gwich'in topical dictionary](#). Accessed: 15-Oct-2024.
- Hyunji Hayley Park, Lane Schwartz, and Francis Tyers. 2021. [Expanding Universal Dependencies for polysynthetic languages: A case of St. Lawrence Island Yupik](#). In *Proceedings of the First Workshop on Natural Language Processing for Indigenous Languages of the Americas*, pages 131–142, Online. Association for Computational Linguistics.
- Katherine Peter. 1979. *Dinjii zhuh ginjik nagwan tr'iltasai: Gwich'in junior dictionary*. Alaska Native Language Center.
- Justin Spence, Zoey Liu, Kayla Palakurthy, and Tyler Lee-Wynant. 2017. Syntactic annotation of a Hupa text corpus. *Working papers in Athabaskan languages*, pages 37–53.
- Francis Tyers and Karina Mishchenkova. 2020. [Dependency annotation of noun incorporation in polysynthetic languages](#). In *Proceedings of the Fourth Workshop on Universal Dependencies (UDW 2020)*, pages 195–204, Barcelona, Spain (Online). Association for Computational Linguistics.
- Francis M. Tyers, Mariya Sheyanova, and Jonathan North Washington. 2017. [UD Annotatrix: An annotation tool for Universal Dependencies](#). In *Proceedings of the 16th International Workshop on Treebanks and Linguistic Theories*, pages 10–17, Prague, Czech Republic.
- United Nations General Assembly. 2019. [Resolution A/RES/74/135: International Decade of Indigenous Languages 2022-2032](#). Accessed: 15-Oct-2024.
- Roberto Zariquiey, Arturo Oncevay, and Javier Vera. 2022. [CLD² language documentation meets natural language processing for revitalising endangered languages](#). In *Proceedings of the Fifth Workshop on the Use of Computational Methods in the Study of Endangered Languages*, pages 20–30, Dublin, Ireland. Association for Computational Linguistics.
- Daniel Zeman, Joakim Nivre, Mitchell Abrams, Elia Ackermann, Jephthe Adolphe, Noëmi Aeppli, Hamid Aghaei, Željko Agić, Amir Ahmadi, Lars Ahrenberg, Chika Kennedy Ajede, Arofat Akhundjanova, Furkan Akkurt, Gabrielè Aleksandravičiūtė, Ika Alfina, Avner Algom, Khalid Alnajjar, Chiara Alzetta, Antonios Anastasopoulos, and 693 others. 2025. [Universal dependencies 2.16](#). LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.

A Gwich'in Specific Features and Values

Values:	∅	L	Ł	D
---------	---	---	---	---

Table 2: Values for feature Classifier.

Value	Description
1	stick-like
2	food
3	cloth-like
4	plural/rope-like
5	animate/dead
6	open container
7	sack of granules, enclosed/sheathed
8	compact
9	deteriorated

Table 3: StemClass values key.