Toward Universal Dependencies for Ainu

Hajime Senuma
University of Tokyo
National Institute of Informatics
senuma@nii.ac.jp

Akiko Aizawa
National Institute of Informatics
University of Tokyo
aizawa@nii.ac.jp

Abstract

Ainu is a language spoken by the Ainu people, a minority ethnic group in Japan. The development of syntactic corpora for Ainu is crucial both for the advancement of linguistic theories and for language revitalization. To solve the situation, we experimentally annotated a short poem of Ainu with Universal Dependencies version 2 scheme. In this paper, we report several language-specific suggestions for the UD annotation of Ainu, obtained through the experience of the initial experiments.

1 Introduction

"In the beginning was the Word"—assuming we have the definition of *word* in the first place. For many languages, the boundary between words and sentences is clear in many cases, not to say in most cases. For some languages, however, it is inherently ambiguous. More than one hundred years ago, linguists heatedly discussed a construction called *noun incorporation*, in which nouns and verbs are combined to form *one* word which has information comparable to a sentence, as a challenging phenomenon to identify such a boundary (Kroeber, 1910; Sapir, 1911; Kroeber, 1911). Even in this age the discussion is not resolved at all but further complicated (Mithun, 1984; Baker, 1988; Baker, 1996; Massam, 2001). Massam (2009, p. 1091) claimed that "[noun incorporation] studies all by themselves provide us with a microcosm of linguistic theory, demonstrating the struggle by linguists to answer major questions such as what constitutes a construction, what are the differences between words and sentences, and what is the relation between meaning and form."

Ainu is one of the representative languages equipped with noun incorporation and *polysynthesis* (words are synthesized by many morphemes), and it has been used as an important tool to uncover the universal nature of human languages (Baker, 1996). It has been spoken by the Ainu, an indigenous people in Japan who originally inhabited islands around the border of what is now Japan and Russia. It is a *language isolate* in the sense that no languages are confirmed to be genetically related to Ainu. The language is in need of linguistic resources, not only for its typological importance, but also for language preservation and revival; Ethnologue classified Ainu as *nearly extinct* (Lewis et al., 2016) and constant effort has been taken to revitalize it (Sato, 2012).

Unfortunately, to the best of our knowledge, there is no syntactically annotated corpus on Ainu thus far. Momouchi et al. (2008) annotated some Ainu poems of Ainu Shinyoshu only with part of speeches and parallel Japanese translation. Bugaeva (2011b) released a freely-accessible dictionary with in-depth glossing for sentence examples but the glossing system is mainly for traditional typological purposes. If we have syntactic annotations, they will enable us to develop advanced natural language processing (NLP) tools and may also serve as an educational tool to explain unique phenomena in Ainu clearly.

Considering the recent revival of dependency grammar (Tesnière, 1959; Tesnière, 2015; Hays, 1960) by theoretical refinement (Järvinen and Tapanainen, 1998; Kuhlmann, 2013; Nivre, 2015) and success in statistical/neural dependency parsing (McDonald et al., 2005; Nivre et al., 2007; Chen and Manning, 2014; Kiperwasser and Goldberg, 2016), Universal Dependencies (UD) is a natural choice to syntactically annotate Ainu texts. One notable feature of modern dependency theory such as UD is its simplicity in the annotation scheme (Nivre, 2015). Because Ainu is a complex language the theoretical aspect of which is not yet fully understood, it is crucial to make its annotation as easy as possible.

In addition, Tesnière (1959, Chapter 276) claimed that dependency diagrams are a useful tool for pedagogical purposes, e.g., deeper analysis of

133

sentences and learning new languages. On top of that, it should be accessible to Japanese students, because Japanese junior high school curriculum includes the *kakari-uke* theory, a variant of the dependency theory independently developed by Hashimoto (1932).

Note that within the UD framework we do not annotate noun incorporation and treat complex verbs formed by noun incorporation as ordinary one-word verbs, as the framework adopted the *lexical integrity principle* and separated morphology as being fundamentally different from syntax (de Marneffe et al., 2014). We will try to discover the nature of noun incorporation only when we resolve the problems of syntactic side. Practically speaking, this treatment will not be so problematic for Ainu, because it is in poetry where noun incorporation is highly productive and for day-to-day speech, noun incorporation is usually used only for idiomatic expressions. Still, annotating Ainu is by no means trivial. As we will see below, there are several unique constructions other than noun incorporation and we will discuss how to annotate them in UD.

To endeavor to develop the UD corpus of Ainu, as an initial effort, we examined a short poem in a collection of Ainu poetry and experimentally annotated texts with UD. This tiny experimental "corpus" contains only 36 sentences and 516 words in total but through this experiment we devised a prototype of the annotation scheme for Ainu. In this paper, we report several findings which may lead toward the complete corpus in the future.

2 Notation

For glossing, we adopted the Leipzig Glossing Rules (Comrie et al., 2008) with the following additions: n-VAL for valency (e.g., 1VAL indicates monovalent verbs) and INT for intensifiers.

We omitted the `root` dependency from each dependency diagram if visually redundant.

3 Resources

3.1 Source text

We used the poetry *Ainu Shin'yōshū* "Ainu Godly Tales" written by a female poet Yukie Chiri (Chiri, 1923) as our raw corpus. It contains 13 short traditional Ainu poems (*kamuyyukar* "godly tales") in romanized form as well as Japanese translations by herself. In Japan, it has been recognized as a great masterpiece and now published from Iwanami Bunko, a series of paperbacks for literary

classics, similar to Penguin Classics and Reclam. The English translation of the poetry by Peterson (2013) is openly accessible.[1]

We chose the work as our text because of the following reasons:

1. It is now under public domain and freely available from Aozora Bunko.[2]

2. It uses Classical Ainu, a register of Ainu which exhibits noun incorporation more extensively than colloquial one.

3. It was compiled by a native speaker.

In this experimental annotation, we used the sixth poem *Pon Horkewkamuy Yayeyukar: "Hotenao"* "A little wolf-god recites a song about himself: 'Hotenao' " ("6: The Song the Wolf Cub Sang" in Peterson (2013)'s translation) as reference.

The result was uploaded as the first author's GitHub project under CC BY 4.0.[3]

3.2 Orthography

Traditionally the Ainu language has no writing systems. From the 19th century, however, two writing systems have been co-developed: one with Latin characters and the other with Japanese katakana. While katakana system has an advantage that it is easy for Japanese people to learn, it has a disadvantage that it generates ambiguity for some cases. For instance, modern katakana orthography adopted by the Foundation of Research and Promotion of Ainu Culture (FRPAC) uses the same character ッ for germination and a consonantal syllable *t*, resulting that both *makkosanpa* "to brighten suddenly" (Tamura, 1996, p. 376) and *matkosanpa* "to wake up suddenly" (Tamura, 1996, p. 380) are written as マッコサンパ. On the other hand, the Latin-based writing system has less ambiguity and thus is suitable for research purposes, while elderly people and children in Japan may feel difficulty to learn it.

The original text of our corpus was written in Latin characters following Kyōsuke Kindaichi's romanization style. We, however, manually converted it into modern (Latin-based) romanization style because doing so will be useful for further research analysis as it is adopted by modern scholars (Tamura, 2000; Sato, 2008) and also for being read

---

[1] http://www.okikirmui.com
[2] http://www.aozora.gr.jp/cards/001529/files/44909_29558.html
[3] https://github.com/hajimes/ud-ainu

by beginners of Ainu as it is adopted in textbooks for non-experts by FRPAC (2011). Although there are several minor differences between these work, we basically followed FRPAC (2011) where there are conflicts. To name a few, 1. we use "=" as a delimiter between pronominal morphemes and other words (like *ku=kor* "I=have"), 2. the first letter of each sentence is *not* capitalized, and 3. we prefer to retain morphological values (e.g., *sanpa* because it consists of *san* and *pa*, in spite that its sound is *sampa*). We also corrected errors in the original text by using critical studies on the work (Kirikae, 2003; Sato, 2004).

Tokenization is relatively simple with the exception of treatment for pronominal clitics. See Section 4 for the topic.

3.3 Dictionary and referential textbook

For dictionary and reference, we mainly referred to Tamura (1996) and Tamura (2000) (English translation of Tamura (1988)), respectively. Strictly speaking, Tamura (1996) deals with the Saru dialect while our text is in the Horobetsu dialect, but two dialects are so close that they have few differences in vocabulary and no grammatical differences.

We also used other dictionaries, e.g., Nakagawa (1995), Kayano (2002), and Bugaeva (2011b). For reference, we also consulted Refsing (1986), Shibatani (1990), and Sato (2008).

4 Pronominal Clitics and Polypersonal Agreement

Polysynthesis often comes with *polypersonal agreement* and Ainu is no exception. Ainu verbs are obligatorily inflected with subject and object. For example, a prefix *e=un=* denotes 2SG>1PL.EXCL "you (but not you all) [hit, took care of, etc.] us (but not including you)".

It is highly controversial whether these pronominal morphemes are words, clitics (that is, relatively dependent elements which are still separable from other words), or affixes (that is, partial elements of other words). Bugaeva (2011a, Section 2.3) argued that *=an* and *=as* are words, *a=* and *eci=* are clitics, and *ku=, ci=, e=, en=, un=, i=* are "fully-fledged prefixes". We applied the classical test of cliticness (Zwicky, 1985) and decided to treat all of these morphemes as clitics. We annotated them with PART (particles) for their part-of-speech category and used the aux relation when they depend on verbs and nmod:poss when nouns.

For convenience, we treated delimiters "=" as if they are a part of pronominal clitics. For example, *sap=as* "go=I" is tokenized as *sap* and *=as* rather than *sap*, *=*, and *as*. This style is consistent with Tamura (1996) and the glossary of FRPAC (2011) in which pronominal clitics are affixed with the delimiters in their entries.

5 Relative Clauses as Adjectival Expressions

In light of syntactic categories (or parts-of-speech), the most noticeable feature of Ainu is a lack of adjectives. Instead, the language employs relative clause constructions to obtain the same effect. This phenomenon is similar to Arapaho (Wagner et al., 2016).

For example, since Ainu is SOV in most cases,

horkew   *pon*
wolf   be.small.1VAL   'the wolf is small'
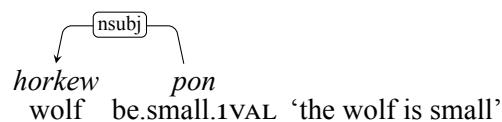
On the other hand, if we place the verb at the preceding position of the noun, it acts as relative clause modification.

pon   *horkew*
be.small.1VAL   wolf   'a wolf that is small'
   or 'a small wolf'
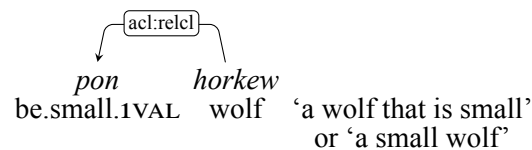
For *alienable* nouns, pronominal possession is also expressed by relative clause-like construction. For example (Tamura, 2000, p. 87),

ku=   kor   tennep
1SG=   have.2VAL   baby   'a baby that I have'
                     or 'my baby'
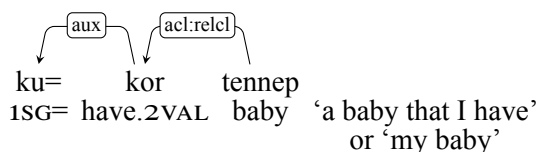
A condition to form relative clauses is that the remaining number of slots of a verb expression must be one, so it actively interacts with the valency property of verbs.

6 Pronominal Arguments

In traditional understanding, Ainu has polypersonal agreement and we have the following analysis (the phrase is a shortened form of a sentence ky-6-33 *eani anak pon horkew sani e=ne ruwe tasi an ne* in our corpus):

eani anak horkew e= ne
you INT wolf 2= COP
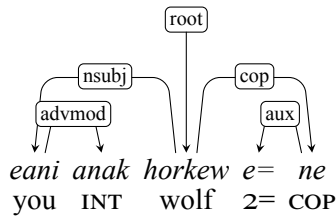
'You are a wolf'

Here, *eani* is used as a common pronoun (PRON) and a clitic *e=* is used as an auxiliary item (PART) to the verb. The problem is that in almost all cases explicit pronominal arguments like *eani* is omitted. In 516 words of our corpus, these "pronouns" occurred only once. To make the matter worse, these "pronouns" often behave like adverbials rather than nominals.

Because of this, several linguists (cf. Baker (1996)) claimed that the clitics are "true" arguments to verbs in these polysynthetic languages and the pronoun-like words such as *eani* in Ainu are mere adjuncts. According to Baker, this analysis was already pointed out by Wilhelm von Humboldt for Aztec in the 1830s but it was not until Eloise Jelinek discussed the matter in the 1980s that the analysis is widely recognized. Although Baker's analysis is based on Chomskyian constituency framework, if we somehow apply their discussion to our text, we may have the following analysis:

eani anak horkew e= ne
you.ADV INT wolf 2= COP
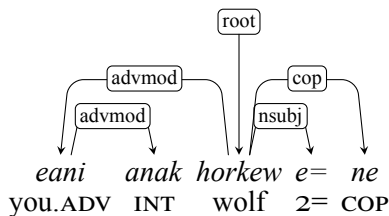
'Speaking of you, you are a wolf.'

In our annotation scheme, however, we maintain the first conservative approach. This approach has an advantage that it can consistently annotate a subtype of Ainu which adopted several constructions from Japanese and was used in some dialects. Izutsu (2006) reported that in some dialects the construction of Ainu sentence became close to the Japanese language and they lack pronominal clitics even though in original Ainu they are mandatory. Instead, they used pronoun-like words such as *eani* as true core arguments to verbs. If we adopted the first annotation scheme, there is no difficulty to handle this exceptional situation as:

eani anak horkew ne
you INT wolf are
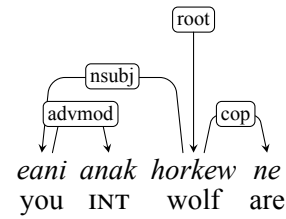
7 Language-Specific Features

7.1 Person: Fourth person

The Ainu language has a peculiar set of pronominal clitics: subject postfix *=an* for monovalent (intransitive) verbs, subject prefix *a=* for polyvalent (transitive) verbs, and object prefix *i=*. It fulfills various rolls depending on the context (Tamura, 2000, pp. 63–80): 1. most commonly first person plural (1PL) inclusive (whereas an ordinary prefix *ci=* for 1PL exclusive); 2. first person singular (1SG) in quotational sentences (such as "I" in "He said 'I saw the man.' "); 3. 1SG in oral literature, because the Ainu literature often employs the style with which someone talks about himself/herself in quotational sentences, 4. general laws ("In general, people do ..."); 5. indefinite person; 6. used for passive construction; and 7. at least in the Saru dialect, traditionally used as second person polite form from females to males of higher position such as their own husbands and village-leaders.

There is no unanimous agreement what we should call these pronouns. Nakagawa (1995) used the term "fourth person", while Tamura (2000) used "indefinite person" and Shibatani (1990), Sato (2008), and Bugaeva (2011a) used "first person plural inclusive".

In our scheme, we used Person=4 for expedience, as available in UD version 2. Whereas UD defines Person=4 as "a third person argument morphologically distinguished from another third person argument" that is suitable for Ainu to some extent, its usage in Navajo, cited as a typical example, is much different from that in Ainu. One possible solution (owing to Francis Tyers) is to annotate it as first person inclusive with Clusive=Inc. We hope to gain insights from research into other languages in the future.

7.2 Number: pluractional

In Ainu, some verbs are inflected by pluractionality, that is, the plurality of actions.

For example, *tuypa*, the pluractional form of *tuye* "to cut" indicates (Tamura, 2000, Section 4.3.2.4): 1. two or more people do cutting, 2. one person cuts two or more objects, or 3. one person cuts one object more than once.

Pronominal clitics may enforce pluractionality. Our corpus contains a phrase *pis ta sap=as* "I go to the beach" where *sap* "to go" is in a pluractional form though the action is semantically singular. This is possibly because, as we stated in previous section, in the oral literature 1PL inclusive (*=as*) is used as 1SG.

To annotate this phenomenon, following the tradition of Ainu studies in which it is treated as number, we tentatively used a language-specific feature `Number=Pluract`. However, this phomonenon clearly deviates from number, which basically serves as an agreement system between nouns, often purely syntactically, rather than semantically. In the literature of typology there seems to be a difference of opinion as to whether pluractionality is related to aspect or it is an independent feature. We will make further research on this point, and we may adopt either `Aspect=Pluract` or `Pluract=Yes` in the future.

7.3 Valency

Valency (number of core arguments attachable to verbs) enjoys the central role in the Ainu morphology and syntax. Morphologically it governs the system of noun incorporation. It also affects syntactic phenomena such as selection of pronominal clitics and construction of relative clauses.

Commonly used valency are monovalent (intransitive), divalent (monotransitive), and trivalent (ditransitive). In addition, avalent verbs are not so rare, e.g., *sirkunne* (it expresses a sentential phrase "it is dark" by one word). The maximum of valency in Ainu is not known but Bugaeva (2015, p. 828) reported the existence of a tetravalent verb (`Valency=4`) *korere* "to make somebody give something to somebody".

In the above example, *kor-e-re* is formed by doubly suffixing causatives: "have-CAUS-CAUS". Here *e* / *re* (allophones of *te*) add valency and causativity. The degree of modification depends on morpheme; for example, while *te* obligatorily increases valency, an indefinite causative morpheme *yar* does not change it. Turkish has a similar system, which allows multiple causative voices. In Ainu, however, causative is just a part of verb formation and has no special status.

The morphological system of Ainu (notably Classical Ainu) productively affects the valency of verbs: e.g., valency-increasing operators such as causative suffixes and applicative prefixes (such as instrumental *e-*, adding an argument slot meaning 'by using') and valency-decreasing ones such as incorporated nouns and reflexive prefixes.

To annotate this phenomenon, we used a language-specific feature `Valency=0`, `Valency=1`, etc. Possibly we may borrow the terminology of UniMorph (Sylak-Glassman et al., 2015) features, that is, `Valency: DITR, IMPRS, INTR, TR` in the future.

7.4 Inalienable possession

The only inflectional system of Ainu nouns is by inalienable possession, with the form similar to ezāfe in Persian (Bugaeva, 2011a, p. 520).

Only a handful of nouns are classified as being inalienable, e.g., body parts (including names) and family members. They are inflected if being possessed by someone. For instance, *re* "name" is inflected to *réhe*; and with 1SG pronominal clitic *ku=*, we obtain *ku=réhe* "my name".

Likewise, a class of nouns called *locative nouns* has possessed forms. For example, a locative noun *or* is used as a concept of "place" as in *atuy or un* "sea place to", that is, "into the sea", while its possessed form *oro* is used to indicate "its place" as in *oro wano* "its.place from", that is, "from there".

Several nouns act as *nominalizer* in their possessed forms. For instance, *hawe*, the inflected form of *haw* "voice", is used as a nominalizer related to auditory sense, as in *itak=as hawe* "speak=I fact" or "the fact that I speak'. Unlike relative clauses (see Section 5), these nominals are prefixed only by complete sentences, and thus we used `acl` relationship to mark the dependency rather than `acl:relcl`.

To annotate this phenomenon, we used a language-specific feature `Possessed=Yes`.

8 Conclusion

This paper proposed several UD annotation schemes for Ainu as a prototype, obtained through experimental annotation for a short poem. We plan to annotate the whole collection of poetry by using this scheme and publish it as an open-source corpus. We hope this endeavor serves as a basis for building huge corpora of Ainu in the future.

References

Mark C. Baker. 1988. *Incorporation*. Chicago University Press.

Mark C. Baker. 1996. *The Polysynthesis Parameter*. Oxford University Press.

Anna Bugaeva. 2011a. A diachronic study of the impersonal passive in Ainu. In *Impersonal Constructions: A Cross-Linguistic Perspective*, Studies in Language Companion, pages 517–546. John Benjamins.

Anna Bugaeva. 2011b. Internet Applications for Endangered Languages: A Talking Dictionary of Ainu. *Waseda Institute for Advanced Study Research Bulletin*, 3:73–81.

Anna Bugaeva. 2015. Valency classes in Ainu. In Bernard Comrie and Andrey Malchukov, editors, *Valency Classes in the World's Languages*. Mouton de Gruyter.

Danqi Chen and Christopher D. Manning. 2014. A Fast and Accurate Dependency Parser using Neural Networks. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 740–750, Stroudsburg, PA, USA. Association for Computational Linguistics.

Yukie Chiri. 1923. *Ainu Shin'yōshū [Collection of Ainu Kamuyyukar]*. Kyodo Kenkyusha.

Bernard Comrie, Martin Haspelmath, and Balthasar Bickel. 2008. The Leipzig Glossing Rules. Technical report, Max Planck Institute for Evolutionary Anthropology and University of Leipzig.

Marie-Catherine de Marneffe, Timothy Dozat, Natalia Silveira, Katri Haverinen, Filip Ginter, Joakim Nivre, and Christopher D. Manning. 2014. Universal Stanford Dependencies: A cross-linguistic typology. In *In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 4585–4592.

FRPAC. 2011. *Chūkyū Ainugo: Horobetsu [Intermediate Ainu: Horobetsu Dialect]*. FRPAC (The Foundation of Research and Promotion of Ainu Culture).

Shinkichi Hashimoto. 1932. *Kokugogaku Gairon [Elements of Japanese Linguistics]*. Iwanami Shoten.

David G. Hays. 1960. Grouping and Dependency Theories. In *Proceedings of the National Symposium on Machine Translation*, pages 258–266.

Katsunobu Izutsu. 2006. Ainugo no Hinshi Bunrui Saikō: Iwayuru Ninshō Daimeishi o Megutte [Ainu Parts of Speech Revisited: with Special Reference to So-called Personal Pronouns]. *Journal of Hokkaido University of Education: Humanities and Social Sciences*, 56(2):13–27.

Timo Järvinen and Pasi Tapanainen. 1998. Towards an implementable dependency grammar. In *Processing of Dependency-Based Grammars: Proceedings of the Workshop, COLING-ACL'98*, pages 1–10.

Shigeru Kayano. 2002. *Kayano Shigeru no Ainugo Jiten [The Ainu Dictionary by Shigeru Kayano]*. Sanseido, 2nd edition.

Eliyahu Kiperwasser and Yoav Goldberg. 2016. Easy-First Dependency Parsing with Hierarchical Tree LSTMs. *Transactions of the Association for Computational Linguistics*, 4:445–461.

Hideo Kirikae. 2003. *Ainu Shin'yōshu Jiten [Lexicon to Ainu Shin'yōshū]*. Daigaku Shorin.

A. L. Kroeber. 1910. Noun incorporation in American languages. In *Verhandlungen der XVI Internationaler Amerikanisten-Kongress [Reprinted in The Collected Works of Edward Sapir, Vol. 5, 1990]*, pages 569–576. Hartleben.

A. L. Kroeber. 1911. Incorporation as a Linguistic Process. *American Anthropologist*, 13(4):577–584.

Marco Kuhlmann. 2013. Mildly Non-Projective Dependency Grammar. *Computational Linguistics*, 39(2):355–387, June.

M. Paul Lewis, Gary F. Simons, and Charles D. Fennig. 2016. *Ethnologue: Languages of the World, Nineteenth edition*. SIL International.

Diane Massam. 2001. Pseudo Noun Incorporation in Niuean. *Natural Language and Linguistic Theory*, 19(1):153–197.

Diane Massam. 2009. Noun Incorporation: Essentials and Extensions. *Linguistics and Language Compass*, 3(4):1076–1096.

Ryan McDonald, Fernando Pereira, Kiril Ribarov, and Jan Hajič. 2005. Non-projective Dependency Parsing using Spanning Tree Algorithms. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing - HLT '05*, pages 523–530. Association for Computational Linguistics.

Marianne Mithun. 1984. The Evolution of Noun Incorporation. *Language*, 60(4):847–894.

Yoshio Momouchi, Yasunori Azumi, and Yukio Kadoya. 2008. Research Note: Construction and Utilization of Electronic Data for Ainu Shin-yosyu. *Bulletin of the Faculty of Engineering at Hokkai-Gakuen University*, 35:159–171.

Hiroshi Nakagawa. 1995. *Ainugo Chitose Hōgen Jiten [The Ainu–Japanese Dictionary: Chitose Dialect]*. Sōfūkan.

Joakim Nivre, Johan Hall, Jens Nilsson, Atanas Chanev, Gülşen Eryiğit, Sandra Kübler, Svetoslav Marinov, and Erwin Marsi. 2007. MaltParser:

A language-independent system for data-driven dependency parsing. *Natural Language Engineering*, 13(2):95–135, June.

Joakim Nivre. 2015. Towards a Universal Grammar for Natural Language Processing. In *Computational Linguistics and Intelligent Text Processing: 16th International Conference, CICLing 2015, Cairo, Egypt, April 14-20, 2015, Proceedings, Part I*, pages 3–16. Springer International Publishing.

Benjamin Peterson. 2013. *The Song The Owl God Sang: The collected Ainu legends of Chiri Yukie*. BJS Books.

Kirsten Refsing. 1986. *The Ainu Language: The Morphology and Syntax of the Shizunai Dialect*. Aarhus University Press.

Edward Sapir. 1911. The Problem of Noun Incorporation in American Languages. *American Anthropologist*, 13:250–282.

Tomomi Sato. 2004. Chiri Yukie Ainu Shinyōshu no Nandoku Kasho to Tokui na Gengo Jirei o Megutte [Notes on Some Linguistic Problems and the Interpretation of Lines Doubtful in Yukie Chiri's Ainu sin'yôshû]. *The Annual Reports on Cultural Science, Hokkaido University*, 10.

Tomomi Sato. 2008. *Ainugo Bunpō no Kiso [The Basics of the Ainu Grammar]*. Daigaku Shorin.

Tomomi Sato. 2012. Ainugo no Genjō to Fukkō [The Present Situation of the Ainu Language and Its Revitalization]. *Gengo Kenkyu*, 142:29–44.

Masayoshi Shibatani. 1990. *The Languages of Japan*. Cambridge University Press.

John Sylak-Glassman, Christo Kirov, David Yarowsky, and Roger Que. 2015. A Language-Independent Feature Schema for Inflectional Morphology. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 674–680.

Suzuko Tamura. 1988. Ainugo [The Ainu Language]. In *Gengogaku Daijiten*, volume 1, pages 6–94. Sanseidō.

Suzuko Tamura. 1996. *Ainugo Saru Hōgen Jiten [The Ainu–Japanese Dictionary: Saru Dialect]*. Sōfūkan.

Suzuko Tamura. 2000. *The Ainu Language*. ICHEL Linguistic Studies. Sanseidō.

Lucien Tesnière. 1959. *Éléments de syntaxe structurale [Elements of structural syntax]*. C. Klincksieck, Paris.

Lucien Tesnière. 2015. *Elements of Structural Syntax*. John Benjamins Publishing Company, Amsterdam. Translated by Timothy Osborne and Sylvain Kahane.

Irina Wagner, Andrew Cowell, and Jena D. Hwang. 2016. Applying Universal Dependency to the Arapaho Language. In *Proceedings of the 10th Linguistic Annotation Workshop held in conjunction with ACL 2016 (LAW-X 2016)*, pages 171–179, Stroudsburg, PA, USA. Association for Computational Linguistics.

Arnold M. Zwicky. 1985. Clitics and Particles. *Language*, 61(2):283–305.