

Bernard Comrie

Uilta (Orok) attributive adjective order: Grammar, text, typology

Abstract: Northern Eurasia, including the Transeurasian (Altaic) languages, is known for its near-exclusive dominant Adjective-Noun constituent order. WALS claims that Uilta, alone among the Transeurasian languages in its sample, has dominant Noun-Adjective order. This article aims to show that Uilta in fact has Adjective-Noun order, like its genealogical and areal neighbors. The article aims not only to correct a questionable claim in the literature, but also to explore a methodology for investigating such claims, by examining the behavior of attributive adjectives in a corpus of Uilta texts, concentrating on adjectives that express descriptive properties.

Keywords: Uilta (Orok) language, languages of Northern Eurasia, attributive adjective order, areal typology, corpus linguistics

1 Introduction

Northern Eurasia is well-known as the world's largest solid area of languages with dominant constituent order Adjective-Noun (AN) in the noun phrase; see, for instance, Dryer (2005, 2013b), on whose map the presentation in this paragraph is based. The core of this area can be considered to be the languages of the Transeurasian (Altaic) families: Turkic, Mongolic, Tungusic, Koreanic, and Japonic. However, it also includes the small families and language isolates of northern Asia: Yeniseian, Yukaghir, Nivkh, and Ainu; only Chukotko-Kamchatkan on the eastern periphery of the area departs from this pattern by including languages with no dominant order (NDO), a point to which I return in Section 6. To the west, it also includes Uralic, and the Balto-Slavic and Germanic branches of Indo-European. In the west, the boundary between Balto-Slavic and Germanic on the one hand with AN, and Celtic and Romance on the other with Noun-Adjective (NA), is sharp. Elsewhere, one moves from the solid AN area of northern Eurasia into zones with more mixed representation of AN, NA, and NDO before reaching solidly NA areas in Southwest Asia and North Africa and in Southeast Asia, and another solidly AN area in South

Bernard Comrie, University of California, Santa Barbara, e-mail: comrie@ucsb.edu

Asia south of the Himalayas. It is surprising, then, to find a single NA language, Uilta (Orok), in the middle of Sakhalin island, with AN languages to the north, west, and south (and the Pacific Ocean to the east). All the more surprising when one realizes that Uilta is a Tungusic language, i.e. belongs to one of the Transeurasian families usually considered to form the core of the northern Eurasian AN area. This areal-typological outlier certainly merits further investigation.

Dryer gives as the source of his information on Uilta attributive adjective order Piłsudski (2011).¹ Piłsudski actually prepared two versions of his Uilta grammar sketch, one in Russian and one in Polish. I will start with the Russian version (1) (even though this is not the one to which Dryer 2005, 2013b refers), because it is quite unequivocal.

(1) *Прилагательные*

Всегда ставится впереди существительного. (Piłsudski 2011: 688)

Adjectives

Always is placed in front of the noun. [My translation and minor orthographic correction of the original]

Dryer actually refers to the Polish version, or rather to the volume editors' English translation of Piłsudski's Polish text. The original Polish is given in (2).

(2) *Przymiotniki stoją zawsze (?) przed rodzajnikami.* (Piłsudski 2011: 656)

My translation of (2) is given in (3).

(3) Adjectives always (?) stand in front of "rodzajniki".

Two comments are in order. First, Piłsudski hedges the absolute statement of (1) with a question mark, suggesting that there might also be some instances of adjectives following the noun. But more importantly, the question arises of what Piłsudski means by a "rodzajnik" (plural: *rodzajniki*). This is not the usual Polish word for 'noun', which is *rzeczownik*, a word that is used elsewhere by Piłsudski in this sense. In modern Polish linguistic terminology *rodzajnik* means 'article' (as in *(in)definite article*), a category that is, however, quite alien to Uilta (as it is to Polish and Russian). The editors of the 2011 volume translate *rodzajniki* as "the words (nouns) they [i.e. the adjectives

¹ Dryer actually refers to a 1987 preprint that circulated before the 2011 publication. I have verified that there are no significant relevant differences between the two versions. Note that Piłsudski's original manuscripts are from the early 20th century; see Section 3.

– BCJ determine”, and I accept this interpretation. It means that *rodzajnik*, while not actually denoting ‘noun’, is to be interpreted as ‘noun’ in this context. Now, the full translation of (2) given by the editors is as in (4).

- (4) Adjectives always (?) follow the words (nouns) they determine. (Piłsudski 2011: 656)

It will be noted that while the Polish original (2), cf. my translation in (3), unequivocally says that adjectives precede their “*rodzajnik*”, i.e. their head noun, the published translation says that they follow. I can only interpret this as an unfortunate translation error; in a work that is undeniably a monumental contribution to Uilta studies. Piłsudski (2011: 656, 688), incidentally, provides an illustrative example ((5) below, in the transcription used in this article), which clearly shows AN order, with the attributive adjective *bərəmi* and the head noun *ulaa*.

- (5) *bərəmi* *ulaa*
 well-behaved reindeer
 ‘well-behaved reindeer’

In other words, Piłsudski actually claimed that Uilta’s dominant order is AN.

The published version Piłsudski (2011) thus contains an unfortunate mistranslation of the Polish original, and one might simply take note of this, correct the mistranslation, and correct the datapoint in Dryer (2013b). All the more so given that a number of other grammars and grammar sketches of Uilta also identify AN as the dominant order (with none known to me that claim NA), including: Nakanome (1928: 34), Tsumagari (2009: 10), Ozoliņa (2013: 246), Tsumagari and Yamada (2024: 451). Only Ozoliņa (2018: 22), while repeating this general statement, tempers it somewhat by saying that proprietive adjectives usually follow their head; see Section 5.5. However, in correspondence on this Dryer (p.c.) emphasized to me, quite correctly, that the final arbiter of what constitutes the order of attributive adjective and head noun in Uilta should not be what a grammar or grammars say, but rather what documentation of the language shows. Whence the present article.

2 The language and its speakers

The language name “Uilta” is based on the indigenous name (*uilta*, variant *ulta*) of the language that in earlier literature is usually referred to as “Orok”. In this, I follow Tsumagari and Yamada (2024), the most recent survey article on the language

and on which this section is based. Uilta is a Tungusic language, most closely related to Ulcha, and more distantly to other members of the Nanaic subgroup; for further information on subgrouping of Tungusic languages, see Janhunen (2024: 7–10).

The traditional home of the Uilta is the central horizontal belt of Sakhalin, with the unrelated (also to each other) languages Nivkh to the north and Ainu to the south. There is a shallow but clear division into Northern and Southern dialects of Uilta. From 1905 to 1945 Sakhalin was divided politically between Russia/the USSR (northern half) and Japan (southern half), with this boundary running between the areas where Northern and Southern Uilta are spoken. The southern half of the island reverted to the USSR/Russian Federation after World War II, and in the following period many speakers of Southern Uilta relocated to Hokkaido (Japan). In the historical period, ethnic Uilta have numbered in the hundreds, although the language is now heavily endangered; Tsumagari and Yamada (2024: 437) suggest that “[t]oday, only a couple of fully fluent speakers survive”.

Ozoliņa (2013) is the most extensive descriptive grammar of Uilta, supplemented by the article Ozoliņa (2018) specifically on attributive constructions, including attributive adjectives. The grammar sketches of Tsumagari (2009) and Tsumagari and Yamada (2024) are especially useful in providing more analysis, while Petrova (1967) remains an important resource. As dictionaries, I have consulted Ozoliņa (2001) and Ozoliņa and Fedjaeva (2003) (for the Northern dialect), and Ikegami (1997) and Magata (1981) (for the Southern dialect); the vocabulary in Ikegami et al. (2008: 89–107) also proved useful, especially in identifying Northern and Southern dialect variants.

Uilta examples are given in the transcription used by Ikegami (1997), which has two practical advantages. First, it corresponds to the transcription used in most of the source examples. Second, it corresponds closely to a broad IPA transcription, with the following exceptions: Palatal affricates are indicated as voiceless *č* and voiced *ǰ*; long segments (vowels and consonants) are indicated by doubling the letter; other combinations of a vowel followed by *i* or *u* indicate falling diphthongs. Examples from other sources are adapted to this system. Uilta has vowel harmony, and upper-case letters are used to indicate archiphonemes.

3 Sources

In assembling a corpus of Uilta textual material for the purposes of this investigation, I was guided by the following rather stringent selection principles:

- a) Only published texts were used. My material and analysis can thus readily be checked against the original.

- b) Only prose texts were considered. This was to avoid any possible bias from the influence of metrical or other constraints leading to atypical constituent orders.
- c) Only original texts were included. For the purposes of this investigation, involving as it does the study of a single language, this seems a reasonable constraint.² This involved excluding not only translations, but also one text, Pl_19, that is presented as a kind of phrasebook without clarifying how the Uilta material was collected, i.e. it is possible that the Uilta versions were elicited as translations from some other language. Text Ym_08 consists of a series of extensive responses by a native speaker to brief prompts in Uilta from the researcher; only the responses were included.
- d) Only transcriptions of oral texts were included. The recent development of Uilta written by native speakers is a major event in the social life of the language, but brings with it possible influences from the new media or from writing practices acquired in other languages. For consistency, therefore, only oral texts were considered.
- e) Duplication of texts was usually avoided. In particular, some of the texts in Ikegami (2002) part 3 are different editions of the same texts that appear in part 1; for details, see (iii) below. I used the texts in part 3, since they are provided with English translations in the publication and are thus more widely accessible. However, while Pl_02 and Ik_01_20, and likewise Pt_08 and Ik_01_17, are different versions of the same story, I included both since there are substantial differences between the two versions of each.

This left a substantial body of texts, as detailed below, sufficient, in my opinion, to investigate reliably the position of attributive adjective relative to head noun in Uilta.

One caveat should, however, be mentioned. Petrova (1967: 55) says:

A characteristic of adjectives in Uilta is their more frequent use in the role of predicate than in the role of attribute. As far as can be judged from the textual material, the Uilta, having once delineated an object in terms of a feature, in what follows do not consider it necessary to repeat these features as an attribute of the given object. An exception is made only for a few words, such as *daaji* 'big', *nuuči* 'small' and some others. [My translation]

I have not investigated whether Uilta does indeed have a low density of attributive adjectives in comparison with other languages, and even if this should turn out to

² Those interested in the use of parallel translations for purposes of cross-linguistic comparisons will be pleased to note that there is now an Uilta translation of *Le petit Prince* (Saint-Exupéry 2016).

be the case, the number of attributive adjectives identified in the textual material seems sufficient on which to base conclusions.

The following are the texts that have been used as a source of material:

i) Piłsudski (2011): These texts were collected in Sakhalin at the beginning of the 20th century. The published version reflects the efforts of the volume editors in interpreting Piłsudski's materials. The texts are numbered in the volume and I have used this numbering to identify texts using the formula Pl_{xy} . In fact, nearly all of the (not particularly numerous) examples are from Pl_{02} , by far the longest of Piłsudski's prose texts. In the 2011 volume, lines are numbered, and this numbering has been adopted here following the formula pq (i.e. a sequence of two digits); note that a numbered "line" may in fact contain more than one line of printed text. The general pattern of presentation of examples is: Piłsudski's original Uilta version – he used a Latin-script spelling that is at least in part impressionistic rather than phonemic; then Piłsudski's Polish word-by-word translation; then the editors' English word-by-word translation. This is followed by a Cyrillic version of the text corresponding to the editors' understanding of the original, a free Russian translation of the editors' version, and a free English translation of the same. There are sometimes notes by Piłsudski or by the editors. It should be borne in mind that this invaluable early collection of Uilta material does require philological interpretation. The 2011 volume editors in general follow a conservative policy, adhering as closely as possible to what Piłsudski wrote. For the long text Pl_{02} , Tsumagari (2014a) offers an alternative analysis, as Piłsudski's text might be edited by a linguist working with native speakers a century later; this includes Tsumagari's amended version of Piłsudski's original in Latin script, interlinear morphological analysis, glossing using English as the metalanguage, and a free Japanese translation. There are occasional stark discrepancies between the two versions, including some relevant to attributive adjective order. The texts from Piłsudski (2011) included in the corpus are Pl_{01} , Pl_{02} , Pl_{07} , and Pl_{09} ; some of the other short texts might meet criteria (a)–(e) but do not contain any relevant material.

ii) Petrova (1967): These texts were collected in Leningrad (now: St. Petersburg) in 1936 and 1949 from Uilta students from Sakhalin, speakers of the Northern dialect. The texts are numbered 1–10 in the original. Most of the texts are traditional stories, though some are descriptive of traditional life, and Pt_{10} is a collection of riddles. The texts are identified using the formula Pt_{xy} , where "xy" indicates Petrova's text number. The starting line of an example is identified as mn_{pq} , where "mn" indicates the number of the paragraph within that text (and "00" indicates the title), "pq" the line number within that paragraph. Each text is presented in the original in Uilta in Cyrillic script without interruption, followed by a free Russian trans-

lation without interruption or (with rare exceptions) annotations; for text Pt_10, however, each individual riddle is presented separately and immediately followed by its Russian translation.

iii) Ikegami (2002): These texts were audio-recorded from speakers of the Southern dialect in Hokkaido in 1955–1958 and 1966 (parts 1 and 3, with the texts in part 3 supplemented from dictated versions from 1949–1951), and from speakers of both Southern and Northern dialects in Sakhalin in 1990–1992 (part 2). They provide by far the most extensive published original text material in Uilta, primarily traditional stories. The 2002 publication gathers together, with revisions, texts that had previously been published in various outlets. The texts are divided into three parts, with the texts numbered separately within each part. Texts are identified according to the formula Ik_vw_xy, where “vw” indicates the part and “xy” the number of the text within that part. For texts in parts 1 and 2, the line at which an example starts is identified by page number and line on that page, using the formula (1)mn_pq, where “(1)mn” indicates the page and “pq” the line. A “line” consists of the original Uilta text in Latin script followed by a word-by-word translation into Japanese. The texts in part 3 are presented first as continuous text in Uilta in Latin script, with each sentence numbered, then as a free English translation with sentences numbered in the same way. For texts in part 3, the sentence in which an example occurs is identified using the formula mn for the sentence as numbered by Ikegami. Text Ik_03_04, a set of riddles, departs somewhat from this convention, and here the formula mn corresponds to the Roman numeral in Ikegami which identifies an individual riddle – all examples are short, so I did not feel it necessary to number individual sentences separately. In all parts, texts are often accompanied by notes. Incidentally, the four texts in part 3 correspond closely to four texts in part 1 (Ik_01_12, Ik_01_04, Ik_01_11, Ik_01_19, following the order of texts in part 3), so these texts from part 1 are not included in the corpus. The texts that are included in the corpus are Ik_01_01 through Ik_01_17, Ik_01_20, and Ik_01_28; Ik_02_01 through Ik_02_04 and Ik_02_16 (all from speakers of the Northern dialect); and Ik_03_01 through Ik_03_04. There is also a Russian-language edition of this collection (Ikegami 2007), with the Uilta text in Cyrillic script and a word-by-word (occasionally phrase-by-phrase) translation into Russian. Tsumagari (2009: 18–19) presents an analysis of Ik_01_01 with morpheme divisions, interlinear glosses using English as the metalanguage, and a free English translation.

iv) Pevnov (2014): This single text was audio-recorded in 2012 in Sakhalin from a speaker of the Northern dialect. It is a renarration of a dream originally told to the speaker by her mother. The text is identified as Pv. The starting line of an example is identified using the formula mn, where for this purpose a “line” is a block of

Uilta text in Cyrillic script (with indication of morpheme boundaries) accompanied by interlinear glosses using Russian as the metalanguage, and a single continuous Russian translation; it thus occupies several physical lines of actual text. Note that line numbers are not indicated in the original.

v) Yamada (2011a, b, c, 2012, 2014, 2015, 2016, 2023): These texts were audio-recorded in Sakhalin in the early 21st century, from speakers of the Northern dialect. They include traditional stories and descriptions of traditional life, but also discussions of the present day, and some conversational material. Each text was published separately; Table 1 shows correspondences between my text-identifying formula Ym_xy and the publication. Some of Yamada’s text publications (Ym_01, Ym_02, Ym_04, Ym_07) contain more than one text, in which case the formula Ym_xy_vw is used to identify individual texts. Within each text, the formula mn indicates the number of the line using the numbering of the original, where “line” is a block consisting of original Uilta text in Latin script, interlinear morphological analysis, glossing using English as the metalanguage, and free Japanese and Russian translations.

Table 1: Correspondences between text codes and Yamada publications.

Text code	Publication	Text code	Publication
Ym_01	Yamada 2011a	Ym_05	Yamada 2014
Ym_02	Yamada 2011b	Ym_06	Yamada 2015
Ym_03	Yamada 2011c	Ym_07	Yamada 2016
Ym_04	Yamada 2012	Ym_08	Yamada 2023

4 Methodology

The present article adopts the same general typological approach to the notion “adjective” as does Dryer (2013b), so it will be useful to quote his characterization of the notion; see (6).

- (6) For the purposes of this map, the term *adjective* should be interpreted in a semantic sense, as a word denoting a descriptive property, with meanings such as ‘big’, ‘good’, or ‘red’. It does not include nondescriptive words that commonly modify nouns, such as demonstratives [. . .], numerals [. . .], or words meaning ‘other’ [. . .]. In some languages, like English, adjectives form a distinct word class. In other languages, however, adjectives do not form a distinct word class and are verbs or nouns [. . .].

It should be emphasized that this characterization serves to define the task at hand. It should not be interpreted as a “definition” of the word class “adjective”, indeed the last sentence of (6) explicitly excludes this as a definition of a language-specific word class.

I go, however, somewhat beyond (6) in the direction of using Uilta-specific word classes as follows. If an item is identified as an “adjective” on the basis of (6) and shares properties with a wide range of other such “adjectives”, then I consider this to define a language-specific class of adjectives. If an item does not share these properties, then I exclude it; relevant cases are discussed in Section 5.5. Conversely, if an item does not satisfy (6), or the issue is questionable, then I nonetheless assign it to the language-specific class of adjectives if it shares properties with “adjectives” as defined by (6) and the sentence preceding this one. I thus use a process of bootstrapping from (6) to a language-specific word class characterization.

The first task was thus to establish a subcorpus of examples of combinations of attributive adjective and head noun following (6). This was done by reading through the texts in the corpus, ensuring that I understood them, i.e. not only the meaning but also how the grammatical structures and lexical items combine to convey this meaning, at least with respect to data relevant to the enterprise. I was then able to extract clear instances of combinations of attributive adjective and head noun, plus less clear cases that might require further investigation. Each individual combination of attributive adjective and head noun identified in the texts constitutes one token. In cases where two (or more) adjectives are attributes to a single head noun, this constitutes two (or more) tokens, i.e. $[[A_1 \text{ and } A_2] N]$ is considered to be two tokens, $[A_1 N]$ and $[A_2 N]$ – note that the formulae given here do not presuppose any particular constituent order. Likewise, if two (or more) nouns are qualified by the same adjective, this constitutes two (or more) tokens, i.e. $[[A [N_1 \text{ and } N_2]]]$ is considered to be two tokens, $[A N_1]$ and $[A N_2]$. This establishes a list of tokens. All tokens consisting of the same combination of attributive adjective and head noun, and having the same syntactic construction, including constituent order (see below), are considered a single type. Needless to say, the actual process of working through the data involved numerous attempts at shortcuts (some successful, some less so), much backtracking, and useful hints from the existing literature on Uilta, both grammatical and lexical. Since the extraction of examples was done manually, some omissions surely occurred inadvertently, but this should not affect the overall picture.

On the basis of this methodology, I established four candidate constructions for combinations of attributive adjective and head noun, as follows. The formula for each construction is closed with an asterisk (*), to distinguish these four specific constructions from the more general distinction between whether the adject-

tive precedes or follows the noun. In these formulae, and more generally in what follows, the order of constituents is important.

(i) AN*: This is simply an attributive adjective followed by a head noun, as in (7).

(7) (Ik_01_10 42_06)

daaji bəjə

big bear

‘big bear’

While most such examples of this construction consist of just an adjective and a noun, it is also possible for the overall noun phrase to contain other constituents, e.g. demonstratives, numerals, or possessors. The adjective is usually a single word, although there are numerous examples of adjectives accompanied by intensifiers, the most frequent being *jiŋ* ‘very’. Other instances of branching attributive adjective phrases are rare, except in the case of proprietive adjectives – see Section 5.5. Complex heads are also occasionally found, as in (17), where the attribute *sagji* ‘old’ qualifies the complex head *uilta əkkə-səl-ni* ‘Uilta women’, lit. ‘Uilta’s women’. Construction (i) is by far the most frequent type. Constructions (ii)–(iv) are not found in the subcorpus with as much variation in terms of optional additions as is construction (i), but this may of course be a feature of the limited corpus. Indeed, several questions that can be answered from the subcorpus or the literature for construction (i) remain unanswered for the other constructions.

The noun phrase, and thus the head noun, can be marked for plurality – this is more frequent for nouns higher in animacy – and will be marked for case according to its syntactic-semantic role in the clause. Where appropriate, the head noun will also be marked for possessor, as Uilta has a head-marking possessive construction as illustrated in (8); the possessive suffix follows the case suffix (if any).³ Note that the possessive suffix can be 3SG even if the possessor is plural (Tsumagari 2009: 6).

(8) (Ik_01_12 49_05)

namu kira-tai-ni

sea shore-ALL-3SG

‘to the shore of the sea’

³ This is actually the inalienable possessive construction. Uilta also has an alienable possessive construction, with the suffix *-ŋu* before any case suffix and the possessor suffix. For further details on Uilta possessive constructions, see Ozoliņa (2013: 118–133).

Agreement of attributive adjectives with their head noun is complex in Uilta. Ozoliņa (2013: 103, 111, 247–248), and earlier Petrova (1967: 55), discuss the situation in the Northern dialect, whereby an attributive adjective may be marked for plural if its head noun is marked for plural, and an attributive adjective may be marked for accusative case if its head noun stands in the accusative. There is no agreement for other cases or for possessor. Tsumagari (2009: 10) says that in the Southern dialect there is no agreement of attributive adjectives.

(ii) NA*: This is simply a head noun followed by an attributive adjective, as in (9).

(9) (Pt_05 01_08)

nooni nəktə nari, moomi, masi
 3SG low person thick strong
 ‘He was a short person, thickset and strong’

The noun phrase resolves to three tokens, *nəktə nari*, *nari moomi*, and *nari masi*, of which the first shows the construction AN*, the other two NA*. Clear instances of NA* are rare, and a relevant factor in (9) seems to be that the adjectives form a list, apparently leading to greater likelihood of postposing some of them.

(iii) NA-ni*: In this construction, the attributive adjective follows the head noun and takes the 3SG possessor suffix *-ni*, as in (10).

(10) (Pt_09 00_01)

ulaa masi-ni
 reindeer strong-3SG
 ‘strong reindeer’

This construction is mentioned in the literature, e.g. Petrova (1967: 151), Tsumagari (2009: 10), Tsumagari and Yamada (2024: 451), but is usually excluded from consideration on the grounds that it is a nominalization. Petrova suggests a literal translation of the type ‘the reindeer’s strength’, but the resulting noun phrase does not denote an abstract quality, so it seems better to go with Tsumagari and a literal translation of the type ‘the reindeer’s strong one’. In Uilta adjectives can be used in headless noun phrases, so that an adjective like *masi* ‘strong’ can participate in a noun phrase *masi* meaning ‘strong one’. Here, there is no explicit nominalization, and no need to assume covert nominalization once one recognizes the existence of headless noun phrases. The precise function of construction (iii) is unclear from the limited subcorpus examples, but it seems to have a delimiting function of the type ‘the/a strong one from among the reindeer’.

If we follow (6), then there seems to be no reason not to consider this construction a combination of attributive adjective and head noun. The A constituent comes from the same set of lexical items as can be A in construction (i). The meaning is to attribute the property A to the referent N, and while literal translations of the kind given in the previous paragraph can be constructed, the most natural translation is usually simply ‘strong reindeer’. It is even possible that the analysis as a nominalization is only historically valid. I therefore include this construction.

I exclude, however, examples where the second component is overtly marked by derivational morphology as a noun, as in (11), since while *ulinga* ‘good’ is an adjective, *ulinga-hta* ‘good thing’ is not – see the entries *ulinga* and *ulingakta* in Ikegami (1997: 219).

- (11) (Pt_06_07_06)

pəttə-l ulinga-hta-či
 seal-PL good-NMLZ-3PL
 ‘best seals’

- (iv) NA-niN*: This construction is like construction (iii) but with repetition of the head noun at the end, as in (12).

- (12) (Ik_01_17_66_07)

gasa daaji-ni gasa
 bird big-3SG bird
 ‘big bird’

The structure may be an apposition involving type (iii) followed by a copy of the head noun, i.e. ‘the big bird, the bird’; the adjective thus goes with the preceding N, not with the following one. This construction is only found in Ikegami’s material, and the small number of attestations makes it difficult to discern a specific function.

In Section 5, then, I will work with these four types, one of which is straightforwardly AN, the other three all instances of NA, whether just this bare combination, or together with a possessive suffix, or together with a possessive suffix and repetition of the head noun.

Once the subcorpus of examples of attributive adjective plus head noun has been constructed, the next step is to count how many instances there are of each of the four constructions listed above, although for the purposes of establishing dominant order only the binary distinction between construction (i) on the one hand (AN) and (ii)–(iv) on the other (NA) is relevant. In then establishing whether Uilta is AN, NA, or NDO, I follow the heuristic rule in (13) proposed by Dryer (2013a).

- (13) The rule of thumb employed is that if text counts reveal one order of a pair of elements to be more than twice as common as the other order, then that order is considered dominant, while if the frequency of the two orders is such that the more frequent order is less than twice as common as the other, the language is treated as lacking a dominant order for that pair of elements. For sets of three elements, one order is considered dominant if text counts reveal it to be more than twice as common as the next most frequent order; if no order has this property, then the language is treated as lacking a dominant order for that set of elements.

Since we are dealing with a pair of elements (attributive adjective and head noun), only the first sentence of (13) is relevant and will be applied in Section 5.

5 Data and analysis

This section presents and, where necessary, comments on the data in the subcorpus of combinations of attributive adjective and head noun; the material is presented in table form. The tables are ordered following the order of the four constructions in Section 4, with some further subdivision in the case of construction (i) (AN*). In each table, the first column gives the adjective, the second column the nouns with which this adjective combines. Uilta lexical items are given in the citation form taken from Ikegami (1997), or using the same transcription system for items not included in Ikegami. Where there is a clear difference between Southern and Northern dialectal forms, this is indicated using the formula “Southern/Northern”; however, variation between the vowel qualities *ø* and *o*, *u* is not indicated, as this represents an ongoing loss of the distinct phoneme *ø* during the documentation of Uilta (Tsumagari and Yamada 2024: 439). There is often variation in indicated vowel length, even within the same source. English translations attempt to capture the general or basic meaning of the adjective, though sometimes a contextually appropriate form is given. Sometimes more than one English translation is given, though this is probably as much due to idiosyncrasies of the English lexicon as to those of Uilta. To avoid undue repetition, *ulaa* is translated in the tables as ‘reindeer’, although more accurately it means ‘domestic reindeer’; *nari* sometimes translates more naturally as ‘man’ than as ‘person’, but in the table I have given only the translation ‘person’. It should be borne in mind that all translations will have passed through Japanese or Russian, although Ikegami (1997) does give English translations for some lexical items. The next column lists the textual occurrences (i.e. of tokens) of each type (i.e. combination of attributive adjective and head noun); different texts are separated

by semicolons, while different occurrences within a text are separated by a comma without repetition of the text identifier. The last two columns give a total of tokens (penultimate column) or types (last column) for the given part of the table. Adjectives are ordered in terms of semantic similarity.

5.1 Construction (i) (AN*)

The adjectives found in this construction are divided into two, with Table 2 showing adjectives with core adjectival semantics, Table 3 those with peripheral adjectival semantics. The dividing line proved to be somewhat subjective, and in case of doubt I followed the devil's advocate principle and assigned an item or group of items to the peripheral category. Adjectives in the core category are usually monomorphemic, at least synchronically, although some may contain the productive suffix *-uli* deriving adjectives related to perception (Ozoliņa 2013: 53). The adjective *irgala* 'patterned' (though English prefers 'spotted' in reference to a seal or its hide, identifying the pattern in question) is derived from the noun *irga* 'pattern', but this is not the usual proprietive suffix *-lu* discussed in Section 5.5.

Table 2: Construction (i) (AN*) with core-semantic adjectives.

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>daaji</i> 'big, great'	<i>nari</i> 'person'	Ik_01_08 28_17; Pt_04 05_03	2	
	<i>jangee</i> 'official'	Ik_01_17 69_05	1	
	<i>beje</i> 'bear, beast'	Ik_01_10 42_06	1	
	<i>moo</i> 'tree, wood'	Ik_01_06 16_11	1	
	<i>gasa</i> 'village'	Ym_08 51, 52	2	
	<i>duxu/duku</i> 'house'	Ik_01_16 62_06, 63_11; Ik_02_01 108_03	3	
	<i>møeluutu</i> 'well'	Ik_01_16 63_12	1	
	<i>tugdala</i> 'bridge'	Ym_05 20	1	
	<i>bajausa</i> PN	Ym_07_01 11	1	
	<i>torisa</i> PN	Ym_07_01 21	1	
<i>kadara</i> 'big'	<i>beje</i> 'bear, beast'	Ik_01_10 42_06	1	
	<i>duxu/duku</i> 'house'	Ik_02_01 108_02	1	
<i>nučiika</i> 'small'	<i>anduma</i> 'box'	Ik_01_15 60_06, 60_10	2	
	<i>xuldaa</i> 'box'	Ik_01_17 79_11	1	
	<i>abdu</i> 'goods'	Ym_03 03	1	

Table 2 (continued)

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>jiktu</i> 'thick'	<i>gara</i> 'branch'	Ik_01_06 16_12	1	
<i>moomi</i> 'thick, thickset'	<i>moo</i> 'tree, wood'	Ym_06_21	1	
<i>nəmdə</i> 'thin'	<i>moo</i> 'tree, wood'	Pt_02 03_05	1	
<i>gugda</i> 'high, tall'	<i>nari</i> 'person'	Ik_01_07 20_09	1	
<i>nəktə</i> 'low, short'	<i>nari</i> 'person'	Pt_05 01_08	1	
<i>sagji</i> 'old (aged)'	<i>ulaa</i> 'reindeer'	Ik_01_05 11_08	1	
	<i>sama</i> 'shaman'	Ik_01_09 37_05	1	
	<i>mama</i> 'old woman'	Ik_02_04 116_06, 116_06; Pt_05 01_06	3	
	<i>əakta</i> 'woman'	Ym_04_01 09	1	
<i>puriga/purəə</i> 'young'	<i>nari</i> 'person'	Ik_01_06 16_01; Ik_01_10 40_06, 41_11, 42_11	4	
	<i>əakta</i> 'woman'	Ym_07_02 25	1	
<i>aja</i> 'good, beautiful'	<i>nari</i> 'person'	Ym_04_02 25; Ym_07_03 13	2	
	<i>əakta</i> 'woman'	Ym_07_01 07	1	
	<i>inəŋi</i> 'day'	Ik_01_06 16_05, 16_05, 16_06	3	
	<i>baawui</i> 'k.o. seal'	Ik_01_08 30_04	1	
	<i>namba</i> 'load'	Ym_07_03 08	1	
<i>ulinga</i> 'good'	<i>nari</i> 'person'	Ym_04_01 19	1	
	<i>jakka</i> 'treasure'	Ik_01_16 63_07	1	
	<i>tətuwə</i> 'clothes'	Ik_01_16 63_07, 63_08	2	
	<i>billaatu</i> 'kerchief'	Ik_01_16 63_09	1	
	<i>xuisa</i> 'travel food'	Ym_02_02 13	1	
<i>orki</i> 'bad, weak'	<i>gewxaatu</i> PN	Ik_01_17 79_07	1	
	<i>nari</i> 'person'	Pt_03 03_01	1	
	<i>dəppi</i> 'food'	Pt_03 02_02	1	
<i>aptauli</i> 'delicious'	<i>dəppi</i> 'food'	Ym_02_01 12	1	
<i>saari/sagari</i> 'black'	<i>tətuwə</i> 'clothes'	Ik_01_03 07_04; Ym_01_01 03	2	
	<i>suli</i> 'fox'	Ik_01_03 09_04, 09_08	2	
	<i>xileepu</i> 'bread'	Ym_05 39	1	
<i>pakal</i> 'dark'	<i>ojo</i> 'skin'	Pt_09 09_02	1	
<i>taagda</i> 'white'	<i>ulaa</i> 'reindeer'	Ik_02_03 113_14	1	
	<i>soondoo</i> 'fawn'	Ik_03_03 31	1	
	<i>məə</i> 'water'	Ik_01_17 73_11, 73_12, 73_13, 74_06	4	

Table 2 (continued)

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>sægda</i> 'red'	<i>sinakta</i> 'hair, fur'	Ik_01_17 71_02, 72_06, 72_14	3	
	<i>nirukta</i> 'hair'	Pt_08 06_11, 06_13	2	
	<i>urakta</i> 'willow branch'	Pl_01 01	1	
	<i>mæə</i> 'water'	Ik_01_17 73_10, 73_11	2	
<i>irgala</i> 'patterned'	<i>pæta</i> 'seal'	Pt_06 04_02, 04_05	2	
	<i>natta</i> 'hide'	Pt_06 02_04	1	
<i>kækku</i> 'empty'	<i>ugda</i> 'boat'	Ik_01_06 16_15	1	
<i>bæðə</i> 'fat'	<i>ulaa</i> 'reindeer'	Pv 06	1	
<i>piłji</i> 'sturdy'	<i>nari</i> 'person'	Ik_01_09 32_03	1	
<i>manga</i> 'strong'	<i>nari</i> 'person'	Pt_05 01_05	1	
	<i>sama</i> 'shaman'	Pl_02 67, 75	2	
	<i>æakta</i> 'woman'	Ym_08 60	1	
<i>tanadala</i> 'skilful'	<i>nari</i> 'person'	Pt_05 13_09	1	
<i>iimau</i> 'fresh'	<i>bəjə</i> 'bear, beast'	Ik_01_17 72_08	1	
	<i>sundatta</i> 'fish'	Ym_02_01 02	1	
<i>gægda</i> 'clean'	<i>busu</i> 'cloth'	Ym_02_01 07	1	
<i>nunjuuli</i> 'cold'	<i>xədu</i> 'wind'	Ym_05 08	1	
<i>nama</i> 'warm'	<i>inaŋi</i> 'day'	Ik_01_06 16_05	1	
<i>namauli</i> 'warm'	<i>inaŋi</i> 'day'	Ik_01_06 16_06	1	
<i>ajakta</i> 'fierce'	<i>køərbə</i> 'male deer'	Pt_05 09_04	1	
	<i>bəjə</i> 'bear, beast'	Pt_09 10_03	1	
<i>bərami</i> 'modest'	<i>nari</i> 'person'	Ym_04_02 10	1	
<i>kəalə</i> 'speedy'	<i>ulaa</i> 'reindeer'	Ik_03_03 48	1	
<i>nilau</i> 'naked'	<i>nari</i> 'person'	Ik_01_02 05_14, 06_04, 06_05, 06_07	4	
<i>baja</i> 'rich'	<i>nari</i> 'person'	Ik_01_03 07_07	1	
<i>jobbee</i> 'poor'	<i>mapa</i> 'old man'	Ik_03_01 55, 56, 65	3	
<i>paskapsuuli</i> 'unusual'	<i>ulaa</i> 'reindeer'	Ik_01_05 11_07	1	
<i>əsarai</i> 'pregnant'	<i>æakta</i> 'woman'	Ik_01_08 27_01	1	
<i>tadə</i> 'true'	<i>təaluŋu</i> 'legend'	Ik_01_08 27_10, 27_11	2	
Total			107	76

The peripheral adjectives in Table 3 are all derived. The first block contains the suffix *-pčĩ*, which derives adjectives from temporal adverbs (Ozoliņa 2013: 53). The second block contains the suffix *-mA*, which derives adjectives from nouns indicating the material from which something is made (as with 'silk ribbon'), or at least one

of its major constituents (as with ‘soapy water’) (Ozoliņa 2013: 51); for simplicity, I have used the translation ‘of X’. The last example illustrates the similitive suffix *-ŋAčī* ‘similar to, -like’ (Ozoliņa 2013: 58) – while there are numerous instances of this formation in the corpus, most are adverbial or predicative.

Table 3: Construction (i) (AN*) with peripheral-semantic adjectives.

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>goro-pči</i> ‘old (traditional)’	<i>saxuri</i> ‘tale’	Ik_01_13 54_11; Ik_01_16 65_14; Ik_03_01 68	3	
	<i>dooro</i> ‘path’	Ik_01_06 16_15	1	
	<i>sama</i> ‘shaman’	Ik_01_09 37_05	1	
<i>ooro-pči</i> ‘aforementioned’	<i>manga</i> ‘hero’	Ik_01_10 43_02	1	
	<i>mөө</i> ‘water’	Pt_08 10_04	1	
<i>čīisə-pči</i> ‘former’	<i>saŋa</i> ‘hole’	Pt_08 07_01, 12_01	2	
	<i>duxu/duku</i> ‘house’	Pt_08 08_02	1	
<i>xuja-ma</i> ‘of horn’	<i>nari</i> ‘person’	Pt_10 02_01	1	
<i>talū-ma</i> ‘of birch bark’	<i>ugda</i> ‘boat’	Ik_01_06 16_07, 16_07	2	
<i>orokto-mo</i> ‘of grass’	<i>duxu/duku</i> ‘house’	Ik_01_17 66_02, 66_05, 77_15, 79_05	4	
<i>čaawakta-ma</i> ‘of claw’	<i>pokto</i> ‘track’	Ym_01_02 14	1	
<i>saura-ma</i> ‘of silk’	<i>leenta</i> ‘ribbon’	Ym_03 22	1	
<i>sala-ma</i> ‘of iron’	<i>sabuu</i> ‘chopstick’	Ik_01_07 24_11	1	
	<i>masaari</i> ‘ax’	Ik_01_07 24_14	1	
	<i>nari</i> ‘person’	Ik_01_09 32_03	1	
	<i>koori</i> ‘mythical bird’	Ik_01_17 66_08, 68_11	2	
	<i>urakta</i> ‘willow branch’	Ik_01_20 89_17, 90_03	2	
	<i>olfjiga</i> ‘hook’	Pl_02 54, 87	2	
	<i>poroxootu</i> ‘steamer’	Ik_01_17 69_08	1	
<i>aisi-ma</i> ‘of gold’	<i>xunaaptu</i> ‘ring’	Ik_01_17 73_05, 78_04, 78_15	3	
	<i>xuldaa</i> ‘box’	Ik_01_17 79_11	1	
<i>məŋu-ma</i> ‘of silver’	<i>xunaaptu</i> ‘ring’	Pt_08 01_11	1	
<i>xulmäu-ma</i> ‘of container’	<i>namba</i> ‘load’	Ym_07_03 08	1	
<i>putta-ŋəči</i> ‘childlike’	<i>nari</i> ‘person’	Ym_06 25	1	
Total			36	24

5.2 Construction (ii) (NA*)

None of the examples of this construction is straightforward, as indicated in the discussion following Table 4, and even so, only core-semantic adjectives are attested, though this could be an artefact of the small number of examples.

Table 4: Construction (ii) (NA*).

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>daaji</i> ‘big, great’	<i>bəjə</i> ‘bear, beast’	Pl_02 34	1	
<i>nučiikə</i> ‘small’	<i>poo</i> ‘relative’	Ym_07_01 10	1	
<i>moomi</i> ‘thick, thickset’	<i>nari</i> ‘person’	Pt_05 01_08	1	
<i>masi</i> ‘strong’	<i>nari</i> ‘person’	Pt_05 01_08	1	
Total			4	4

The first example, *bəjə daaji*, is the only one consisting solely of a noun followed by an attributive adjective. Tsumagari (2014a: 221) amends this to *bəjə daaji-ni*, but while this may well be correct, accepting it for present purposes would be circular, so we retain the version given by Piłsudski. The second one is more fully as in (14).

(14) (Ym_07_01 10)

poo-l-bi *nučiikə-səl-bə*
 relative-PL-REFL small-PL-ACC
 ‘her own small relatives (i.e. junior kin)’

The plural suffix *-səl* normally goes on nouns, and (Ozoliņa 2013: 32) notes that this plural suffix added to some adjectives also has a nominalizing function. So it is possible that (14) actually consists of two noun phrases in apposition, i.e. ‘relatives, small ones’. The two examples at Pt_05 01_08 are those already discussed as example (9). So extracting even 4 examples from the corpus involves a fair amount of application of the devil’s advocate principle.

The text Pl_02 contains a number of instances of an item, in Piłsudski’s spelling, *nuci* or *núci*, sometimes written hyphenated to the preceding noun and sometimes written separately, and which he glosses with the Polish lexical item *mały* ‘small’, i.e. identifying the Uilta item as *nuučī* ‘small’. The editorial team of Piłsudski (2011), following their general principle of adhering closely to Piłsudski’s presentation and

analysis, translate this item into English as ‘small’ or ‘little’. However, in his Uilta dictionary, Piłsudski (2011: 386) has an entry *núci*, with a variant *múci*, distinct from the entry *núci* (i.e. *nuučĩ*) ‘small’ (Piłsudski 2011: 385), and with the definition “is added to nouns about which is narrated in tales” [my literal translation], i.e. with no reference to smallness or other descriptive properties. Tsumagari (2014b: 86–90) argues cogently against the analysis of this item as an adjective meaning ‘small’, even leaving aside general doubts concerning the viability of the construction NA* in Uilta. The following is a summary of his argument (except for my addition of the parallel to *mərgə* ‘wise’). First, the interpretation ‘small’ often does not make sense. Second, the adjective could have been conventionalized as a postposed bound diminutive morpheme, though once again this is often semantically problematic, and this conventionalization as a suffix is not attested in other sources or by contemporary speakers. Note, however, that Ikegami (2002) does show one instance of conventionalization of an adjective when postposed, namely of *mərgə* ‘wise’ as an honorific marker in (15).

- (15) (Ik_01_16 64_03)
andaxa mərgə (or perhaps: *andaxa-mərgə*)
 guest wise (or perhaps: guest-HON)
 ‘honored guest’

Third, it is conceivable that this is the third person plural pronoun *noočĩ* ‘they’ in apposition to the preceding noun, though this hypothesis runs into two problems: The difference in vowel quality is unexpected even given Piłsudski’s impressionistic spelling; the plural pronoun would be inconsistent with instances of a preceding noun singular in both form and interpretation. Finally, Tsumagari proposes that *nuci*, as well as its variant *muci* (found after *o* by assimilation), might reflect a basic representation *-ŋu-čĩ* – Piłsudski sometimes fails to distinguish *n* and *ŋ* elsewhere too – ‘AL-3PL’, i.e. ‘their’, but reinterpreted as a clitic topic marker; Tsumagari cites close parallels from work by Ikegami and Magata. Given this, I have not included examples of postnominal *nuci* as instances of the NA* construction. Even if they were included, this would reflect idiosyncratic ordering of a single adjectival lexical item.

Pevnov’s text provides another example that might be parallel to (9) in terms of providing a list of properties expressed by adjectives, namely (16).

(16) (Pv 05)⁴

j̃iŋ bara ulaa-l: peemura=ddaa, gəlč̣i=ddəə, karau=ddaa,
 very many reindeer-PL dappled=CONC gray=CONC brown=CONC
taagda=ddaa
 white=CONC
 ‘very many reindeer: both dappled and gray and brown and white’

However, there are also reasons for thinking that this is a different structure, as is suggested by Pevnov’s translation, with a colon between the noun *ulaa-l* and the string of adjectives, implying at least a different prosody from simple postposition of a string of adjectives.

5.3 Construction (iii) (NA-ni*)

With this construction again, only core-semantic adjectives are attested (Table 5), though again this could reflect the small number of instances.

Table 5: Construction (ii) (NA-ni*).

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>gugda</i> ‘high, tall’	<i>nari</i> ‘person’	Ik_01_07 20_09, 20_09	2	
<i>aja</i> ‘good, beautiful’	<i>nari</i> ‘person’	Ik_01_03 07_05	1	
<i>manga</i> ‘strong’	<i>nari</i> ‘person’	Ik_01_09 32_02; Pt_07 06_04, 06_07	3	
<i>masi</i> ‘strong’	<i>ulaa</i> ‘reindeer’	Pt_09 00_01, 01_02, 04_01, 05_01, 06_01, 07_01, 09_02, 09_03, 11_02, 11_03, 11_07	11	
<i>ajakta</i> ‘fierce’	<i>bəjə</i> ‘bear, beast’	Pt_09 00_01, 01_03, 02_01, 02_08, 04_01, 05_01, 06_01, 07_01, 07_07, 09_03	10	
Total			27	5

It is noticeable that the vast majority of tokens – 21/27 or 77.8% – are provided by two types found exclusively in one text, where they refer to the two main participants: *ulaa masi-ni* ‘the strong reindeer’ and *bəjə ajakta-ni* ‘the fierce bear’. This is the clearest instance of a substantial discrepancy between token and type frequency.

⁴ The concessive particle is frequently used as a conjunctive coordinator ‘and’ (Ozoliņa 2013: 362–363).

5.4 Construction (iv) (NA-niN*)

Only three adjectives are attested in this construction (Table 6), and all belong to the semantic core.

Table 6: Construction (ii) (NA-niN*).

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>aĵa</i> ‘good, beautiful’	<i>patala</i> ‘girl’	Ik_01_16 63_13	1	
	<i>ǎakta</i> ‘woman’	Ik_01_17 67_02, 68_12	2	
	<i>inǎŋi</i> ‘day’	Ik_01_17 66_06	1	
<i>daaĵi</i> ‘big, great’	<i>gasa</i> ‘bird’	Ik_01_17 66_07, 68_09	2	
<i>gugda</i> ‘high, tall’	<i>nari</i> ‘person’	Ik_01_06 16_10	1	
Total			7	5

5.5 Questionable and excluded examples

A number of items that might at first sight appear to be adjectives, at least following the characterization in (6), should probably not be considered within the remit of the present investigation. I discuss them in this subsection, starting with those that I would most surely exclude and working toward cases that are less clear. All of the excluded and questionable classes have the attributive element before the head noun, either as the only possibility or as the clearly more frequent possibility, so that including these items would only increase the incidence of AN relative to NA. Once again, the devil’s advocate principle is also at work.

Nationality attributive expressions of the type ‘Uilta woman’ are expressed in Uilta as ‘woman of the Uilta’, i.e. using a possessive construction, as in (17).

- (17) (Ym04_01 09)
sagĵi uilta ǎkkǎ-sǎl-ĵi-ni
 old Uilta woman-PL-INS-3SG
 ‘with old Uilta women’

I consider these to be possessive constructions, not attributive adjective constructions. This may seem inconsistent with the treatment of the construction NA-ni* as an attributive adjective construction, but there is a crucial difference. In an example like (10), *masi* is qua lexical item an adjective. By contrast, in (17) *uilta* is

qua lexical item a noun and cannot be used as an adjective. There is one possible counterexample to this in the corpus, alongside several examples of the construction in (17), namely (18).

- (18) (Ym_04_01 12)
 uilta nari-sal
 Uilta person-PL
 ‘Uilta people’

I assume that this is an instance of apposition between two nouns.

Sex-expressing attributes, comparable to English ‘male’ and ‘female’, precede the noun that they qualify. The most general are *xusə* ‘male’ and *əəktə* ‘female’, though there are also more specific terms used with some animals. However, all of these words also function as nouns, e.g. *xusə* ‘man’, *əəktə* ‘woman’, so collocations like (19) are best analyzed as nouns in apposition.

- (19) (Ik_01_10 42_02)
 əəktə bəjə
 woman bear
 ‘female bear’

Quantifiers do not seem to form a single formal category, or a set within a single formal category, in Uilta. Numerals precede their head noun, but are excluded from consideration by Dryer’s characterization (6). The quantifier *čipal* ‘all’ never appears with nominal morphology, and is probably an adverb, following Ikegami (1997: 34). The quantifier *bara* ‘much, many’ is more complex. It appears both before and after the noun to which it refers, and in the latter case when relating to a direct object it takes accusative case marking, as does the head noun, as in (20).

- (20) (Pt_06 07_08)
 pəttə-l-bə baram-ba
 seal-PL-ACC many-ACC
 ‘many seals’

I have not found this pattern with any candidate for adjective status in the corpus other than the quantifiers *bara* ‘much, many’ (numerous examples) and *ojuuka* ‘a little, a few’ (one example at Ym_05 01). I conclude that *bara* and *ojuuka* are either not adjectives, or constitute an irregular subclass of adjectives whose behavior cannot be taken to represent dominant order. Ikegami (1997: 16), however, does consider *bara* and *ojuuka* to be adjectives.

Uilta has a very productive suffix *-lu* deriving propriative adjectives, with the meaning ‘having X, with X’. Indeed, in predicative position this is the usual Uilta translation equivalent of English ‘to have’, and presumably any noun phrase, however complex its internal structure, can occupy the position X. Thus, although *-lu* is attached to the last word in the base noun phrase, it has scope over the whole of that noun phrase. Note that attributes dependent on the propriative adjective stand in the instrumental case, as in (22).

- (21) (Ik_01_17 67_17)
lauta-lu *ǰangee*
 sword-PROP official
 ‘an official with a sword’

- (22) (Pt_08 11_02)
du-ǰi *ǰili-lu* *səwəm-bə*
 two-INS head-PROP idol-ACC
 ‘a two-headed idol’

The order in clear instances of this construction in the subcorpus is exclusively with the attribute before the head; see Table 7.

Table 7: Construction (i) (AN*) with propriative adjectives.

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>bokko-lu</i> ‘having belly’	<i>mama</i> ‘old woman’	Ik_01_15 56_03, 56_06, 57_08, 57_14, 59_06, 61_03	6	
<i>lauta-lu</i> ‘having sword’	<i>ǰangee</i> ‘official’	Ik_01_17 67_17	1	
<i>ǰili-lu</i> ‘having head’	<i>kuukku</i> ‘swan’	Pt_07 04_06	1	
	<i>səwə</i> ‘idol’	Pt_08 11_02, 11_08	2	
<i>ananiǰu-lu</i> ‘having year’	<i>kəərbə</i> ‘male deer’	Pt_09 02_05	1	
	<i>xusə</i> ‘boy’	Ym_07_02 24	1	
<i>meela-lu</i> ‘having soap’	<i>məə</i> ‘water’	Ym_03 10	1	
<i>dausu-lu</i> ‘having salt’	<i>məə</i> ‘water’	Ym_03 11	1	
Total			14	8

This is perhaps surprising, since Ozoliņa (2018: 21) says that propriative adjectives usually follow their head noun. However, she qualifies this claim (Ozoliņa 2018: 22) by noting that when propriative adjectives themselves have attributes (i.e. are

branching), then the usual order is for them to precede, while otherwise they usually follow. Of the examples in our subcorpus, only 3 tokens representing 3 types are non-branching (*lauta-lu ʃangee*, *meela-lu mөө*, and *dausu-lu mөө*), while the other 11 tokens representing 5 types are branching (with the type *bokko-lu mama* accounting for 6 of these tokens). The prevalence of preposed nonbranching proprietive adjectives may therefore be an artefact of the small corpus. Overall, proprietive expressions present a mixed picture. On the one hand, they have a derivational suffix that affects the noun phrase status of their base, so treating them as derived adjectives would be plausible, even if they do not belong to the semantic core. On the other hand, their frequent complex internal structure with branching is atypical for Uilta adjectives. I have not included them in the statistics for constructions (i)–(iv).

Attributes of spatial orientation denote the location of the referent, as illustrated by the English translations of the examples from the Uilta subcorpus in Table 8. In Uilta, those without the suffix *-duma* can certainly be used as nouns, and are usually classified only as nouns by Ikegami (1997), while Ozoliņa (2001) classifies them as both nouns and adjectives. Those with the suffix *-duma* seem to be more clearly adjectives in Uilta. If the attribute of spatial orientation is an adjective, then all these Uilta examples illustrate construction (i) AN*. If it is a noun, then the exact nature of the construction is unclear; it does not appear to be simple apposition, as this does not make sense semantically, in contrast to examples like (19). Of the types discussed in this subsection, attributes of spatial adjectives are the ones I would be most inclined to classify as adjectives. However, I have not included them in the statistics for constructions (i)–(iv).

Table 8: Construction (i) (AN*) with adjectives denoting spatial orientation.⁵

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>nauramji</i> ‘front’	<i>ɲinda</i> ‘dog’	Ik_01_03 07_08	1	
<i>xamarree</i> ‘back’	<i>bagji</i> ‘leg’	Pt_01 02_09	1	
<i>uuwu</i> ‘top’	<i>kumulta</i> ‘fur-lined ski’	Ik_03_04 12	1	
	<i>kipere</i> ‘fish-frying rack’	Ik_03_04 12	1	

⁵ A couple of comments are in order concerning the English translations of some of the items in Table 8. Petrova consistently translates *aanʃee* as ‘left’, but this seems to be an error, as only the translation ‘right’ is given by Ozoliņa (2001: 25–26), Ikegami (1997: 1), and the translations in Ikegami (2002). The translation of *gajau* as ‘little finger’ does, however, seem to be correct for the Northern dialect, where it contrasts with *gajau aan-ni* ‘ring finger’, lit. ‘little finger’s older brother’ (Ozoliņa 2001: 55). In the Southern dialect, *gajau* means ‘ring finger’, and ‘little finger’ is a completely different word, *čimutəə* (Ikegami 1997: 33, 63; Magata 1981: 27, 62).

Table 8 (continued)

Adjective	Noun	Textual occurrences	Frequency	
			Token	Type
<i>uuwu-dumə</i> ‘upper’	<i>xan̄ki</i> ‘room’	Pt_08 07_02	1	
<i>pəduu</i> ‘bottom’	<i>kumultə</i> ‘fur-lined ski’	Ik_03_04 12	1	
	<i>kipərə</i> ‘fish-frying rack’	Ik_03_04 12	1	
<i>baj̄jee</i> ‘opposite (side of)’	<i>namu</i> ‘sea’	Ik_01_08 29_06, 09	2	
<i>dəun̄jee</i> ‘left’	<i>xoldo</i> ‘side’	Ik_01_10 43_08	1	
	<i>isal</i> ‘eye’	Ik_01_17 67_05, 68_16	2	
	<i>xawani</i> ‘armpit’	Ik_01_17 70_08, 71_10	2	
<i>aan̄jee</i> ‘right’	<i>xoldo</i> ‘side’	Ik_01_10 43_07	1	
	<i>gaj̄au</i> ‘little finger’	Pt_08 01_11	1	
	<i>ŋaala</i> ‘hand’	Pt_08 03_05	1	
	<i>xawani</i> ‘armpit’	Ik_01_17 70_07, 71_09	2	
	<i>isal</i> ‘eye’	Ik_01_17 67_06, 68_15	2	
<i>ojo-duma</i> ‘outer’	<i>duxu/duku</i> ‘house’	Pt_08 10_01	2	
Total			23	17

5.6 Frequency summary

Table 9 sets out the relative frequencies of the four different constructions, combining Tables 2–6, for both tokens and types, and both excluding and including peripheral-semantic adjectives (though this turns out to be relevant only for the AN* construction). For ease of comparison, both raw figures and percentages of the relevant total are given.

Table 9: Frequency of different constructions.

	Token				Type			
	Core		Core + Peripheral		Core		Core + Peripheral	
AN*	107	73.8%	143	79.0%	76	84.4%	100	87.7%
NA*	4	2.8%	4	2.2%	4	4.4%	4	3.5%
NA-ni*	27	18.6%	27	14.9%	5	5.6%	5	4.4%
NA-niN*	7	4.8%	7	3.9%	5	5.6%	5	4.4%
Total NA	38	26.2%	38	21.0%	14	15.6%	14	12.3%
Total	145	100.0%	181	100.0%	90	100.0%	114	100.0%

Dryer's rule of thumb given in (13) refers to tokens, so the ratio of AN: NA in Uilta, on the basis of my subcorpus of core-semantic attributive adjectives, is 73.8: 26.2, clearly above the minimum needed to identify AN as the dominant order (66.7: 33.3). Including peripheral-semantic adjectives, or shifting from tokens to types, serves only to increase the ratio in favor of AN. I conclude that the dominant order of attributive adjective and head noun in Uilta is AN.

6 Conclusions and prospect

I hope to have shown that Uilta should be classified as an AN language in the terminology of WALS. This has been quite a time-consuming task, and one might well wonder whether it was worth the effort. It may result in a change to the classification of Uilta for feature 87 in WALS, in which case the present article might serve as a footnote to history. But now that Grambank (Skirgård et al. 2023 a, b) is available, it is quite possible that in the future linguists will consult Grambank rather than WALS, and Grambank (feature GB193, with Jay Lataarche and Jeremy Collins as patrons) classifies Uilta as AN (using WALS terminology), citing Tsumagari (2009) as its source; in a world where Grambank gave this information and WALS gave either the same or no information, I would not have been impelled to write this article.

But there are nonetheless indications that not all is well, or at least clear, at the edges of the vast AN area in northern Eurasia, even if elsewhere within this area, with the exception of Uilta, WALS and Grambank are in agreement. At the eastern edge stands the Chukotko-Kamchatkan family. WALS gives information for three languages of the family, with Itelmen classified as NDO, Chukchi and the closely related Koryak as AN. Grambank includes only Chukchi, but classifies it as “Both orders possible” (BOP; see further below). Both cite Dunn (1999), more specifically p. 161ff (WALS) or p. 292 (Grambank). Reading these pages confirms that both AN and NA are possible orders in Chukchi, but does not include any clear statement that one or the other is dominant. On the basis of the sources cited, at this eastern edge Grambank seems to do better than WALS.

What about the Western edge? Here we find discrepancies within the Uralic family and within the Slavic branch of Indo-European. WALS includes 16 Uralic languages for this feature, all with the value AN. Grambank includes 30 Uralic languages, all AN with two exceptions, since both Eastern Khanty (more specifically: Surgut Khanty) and Livvi are classified as BOP. Eastern Khanty is included in both WALS and Grambank, and both refer to the same source, Filchenko (2007). Filchenko's account is unusually detailed in that it gives percentages for the relative frequency of AN and NA in his corpus: AN 93.5%, NA 6.5% (Filchenko 2007: 132). Faced

with these figures, I would have had no hesitation in calling AN as the dominant order in Eastern Khanty, and the relevant Grambank entry does include a note that AN is more frequent than NA. In order to see why Grambank might have classified Eastern Khanty as BOP, it is necessary to examine the coding definitions for this feature in Grambank, the relevant parts of which are given in (23).

(23) **Summary**

Adnominal property words are also known as “adjectives”, in particular in those languages where they make up a separate word class. We want to include elements that mark ‘adjectival’ function but that might not be described as ‘adjectives’ [. . .] The question concerns the pragmatically neutral order(s).

Procedure

Code 1 if most adnominal property words are placed before nouns.

Code 2 if most adnominal property words are placed after nouns.

Code 3 if adnominal property words can be placed before or after nouns, either because both orders are possible or because some adjectives precede and some follow the noun.

Grambank’s term “adnominal property word” is equivalent to WALs’s “adjective”, and for consistency I will continue to use the abbreviations AN (Grambank: 1) and NA (Grambank: 2) in relation to Grambank. It is less clear that Grambank’s 3 can be identified with WALs’s NDO, so I will use a distinct abbreviation BOP. The problem that arises in practice with (23) is determining the dividing line between AN and BOP or between NA and BOP. For each of these pairs, many languages are in principle consistent with either coding; in Eastern Khanty, for instance, it is true that most attributive adjectives are placed before nouns, but it is also true that adjectives can be placed before or after nouns, and this will hold for any language where the ratio of the two orders is neither 100: 0 nor 50: 50, which probably means most languages of the world. Unlike WALs, where Dryer suggests a cut-off point where one possibility occurs two-thirds of the time – see (13) – Grambank provides no guidance, and it remains unclear if different authors have used the same criteria. For Eastern Khanty, WALs seems to do better than Grambank. For the other Uralic language where Grambank has BOP, Livvi, I lack evidence comparable to the statistical evidence provided by Filchenko (2007) for Eastern Khanty, so the issue must for now remain open.

Turning now to Slavic, WALs includes 10 languages, all classified as AN, while Grambank includes 8 (a subset of the WALs sample), of which 5 are classified as AN, 1 as BOP (Ukrainian), 2 as NA (Czech and Polish). The classification of Czech as NA seems to be a coding error, since the source (Naughton 2005: 49) says that attributive adjectives are “normally” placed before their nouns, only “occasionally”

after, which I would interpret as AN, though with recurrence of the already noted problem in dividing AN from BOP in Grambank. For Ukrainian, the source (Pugh and Press 1999: 166) says: “the attributive adjective as a rule precedes, although it may follow when emphatic or for stylistic reasons”. Since emphasis is pragmatically marked, the order used for emphasis should be disregarded according to (23); it is unclear whether the same applies to “stylistic reasons”. My call here would again be AN, though again noting the problem with the dividing line between AN and BPO. Polish is a more complicated case, and certainly has a much higher incidence of NA than does the general run of Slavic languages; further discussion would exceed the bounds of this section and must for now remain open. But in general, for Slavic, and thus for the western edge in general, WALS seems to do better than Grambank.

The moral is clear: The price of typology is eternal vigilance.

Abbreviations

3	third person
A	Adjective
ACC	accusative
AL	alienable
ALL	allative
AN	Adjective–Noun
BOP	Both orders possible, or some adjectives precede and some follow the noun
CONC	concessive
HON	honorific
Ik	Ikegami (2002)
INS	instrumental
k.o.	kind of
N	Noun
NA	Noun–Adjective
NDO	No dominant order
NMLZ	nominalizer
PI	Piśsudski (2011)
PL	plural
PN	proper name
PROP	proprietary
Pt	Petrova (1967)
Pv	Pevnov (2014)
REFL	reflexive
SG	singular
Ym	Yamada [See Table 1]

References

Citing sources in different languages and scripts and from different publishing traditions (print, digital, and database) has been challenging. For authors whose last name is cited in the main text, I have adopted a single version of the last name, with the author's name as it appears in the original in square brackets. For titles of books, articles, and journals in languages other than English or German, an English translation is added in square brackets, using the translation given in the original publication if there is one, even if this is not a literal translation of the original; for Uilta titles in Cyrillic script, a version using Ikegami's transcription is added in parentheses. Names of publishers are given as in the original, with an English translation only when provided in the original publication and following the version in that publication.

- Dryer, Matthew S. 2005. Order of adjective and noun. In Matthew S. Dryer, Martin Haspelmath, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 354–357. Oxford: Oxford University Press.
- Dryer, Matthew S. 2013a. Determining dominant word order. In Matthew S. Dryer & Martin Haspelmath (eds.), WALS Online (v2020.4) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.13950591> (Available online at <http://wals.info/chapter/s6>, Accessed on 2024-11-15.)
- Dryer, Matthew S. 2013b. Order of adjective and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), WALS Online (v2020.4) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.13950591> (Available online at <http://wals.info/chapter/87>, Accessed on 2024-11-15.)
- Dunn, Michael John. 1999. *A grammar of Chukchi*. Australian National University PhD thesis.
- Filchenko, Andrey Yuri. 2007. *A grammar of Eastern Khanty*. Rice University PhD dissertation.
- Ikegami, Jirô [池上二良]. 1997. ウイルタ語辞典 [A dictionary of the Uilta language spoken on Sakhalin]. Sapporo: 北海道大学図書刊行会 [Hokkaido University Press].
- Ikegami, Jirô [池上二良]. 2002. 増訂 ウイルタ口頭文芸原文集 [Uilta oral literature: A collection of texts. Revised and enlarged edition.] Osaka: 環太平洋の言語 [ELPR Endangered Languages of the Pacific Rim].
- Ikegami, Jirô [Икегами, Дзиро]. 2007. *Сказания и легенды народа уилта [Uilta oral literature: A collection of texts]*. Sapporo: 北海道大学大学院文学研究科 [Graduate School of Letters, Hokkaido University].
- Ikegami, Jirô [Икэгами, Дзиро], E[lena] A. Bibikova [Е. А. Бибикова], L[jubov] R. Katazima [Л. Р. Катазима], S[irjuko] Minato [С. Минато], T[at'jana] P. Roon [Т. П. Роон] & I[rina] Ja. Fedjaeva [И. Я. Федяева]. 2008. *Уилта даирису (Uilta dairisu) / Говорим по-уилтински [Let's speak Uilta]*. Yuzhno-Sakhalinsk: Сахалинское книжное издательство [Sakhalin Book Publisher].
- Janhunien, Juha. 2024. Tungusic as a language family. In Alexander Vovin, José Andrés Alonso de la Fuente & Juha Janhunien (eds.), *The Tungusic languages*, 1–18. London: Routledge.
- Magata, Hisaharu [潤潟久治]. 1981. ウイルタ語辞典 [A dictionary of the Uilta language]. Abashiri: 北方民俗文化保存協会 [Society for the Preservation of Northern Region Culture and Folklore].
- Nakanome, Akira. 1928. *Grammatik der Orokko-Sprache*. Osaka: Osaka Asiatic Society. [Translated from the Japanese original of 1917.]
- Naughton, James. 2005. *Czech: An essential grammar*. London: Routledge.
- Ozoliņa, Larisa [Озолиня, Л. В.]. 2001. *Орокско-русский словарь [Orok-Russian dictionary]*. Novosibirsk: Издательство СО РАН [Publisher of the Siberian Branch of the Russian Academy of Sciences].
- Ozoliņa, Larisa [Озолиня, Л. В.]. 2013. *Грамматика орокского языка [A grammar of Orok (Uilta)]*. Novosibirsk: Издательство Гео [Geo Publisher].

- Ozoliņa, Larisa [Озолиня, Л. В.]. 2018. О непредикативных синтаксических структурах в тунгусо-маньчжурских языках: атрибутивная конструкция в орокском языке [On non-predicative syntactic structures in the Orok language]. *Языки и фольклор коренных народов Сибири [Languages and Folklore of Indigenous Peoples of Siberia]* 36. 20–32.
- Ozoliņa, Larisa [Озолиня, Л. В.] & I[rina] Ja. Fedjaeva [И. Я. Федяева]. 2003. *Орокско-русский и русско-орокский словарь [Orok-Russian and Russian-Orok dictionary]*. Yuzhno-Sakhalinsk: Сахалинское книжное издательство [Sakhalin Book Publisher].
- Petrova, T[aisija] I. [Петрова, Т. И.]. 1967. *Язык ороков (ульта) [The language of the Oroks (Uilta)]*. Leningrad: Наука [Nauka].
- Pevnov, A[leksandr] M. [Певнов, А. М.]. 2014. Текст на орокском (ульта) языке с комментариями [A text in the Orok (Uilta) language with notes]. In Valentin F. Vydrin [В. Ф. Выдрин] & Natalia V. Kuznetsova [Н. В. Кузнецова] (eds.), *От Бикина до Бамбалюмы, из варяг в греки. Экспедиционные этюды в честь Елены Всеволодовны Перехвальской [From Bikin to Banbaluma, from Varangians to the Greeks. Field-inspired essays in honour of Elena V. Perekhval'skaya]*, 55–66. St. Petersburg: Издательство Нестор-История [Nestor-Istoria].
- Piłsudski, Bronisław. 2011. Materials for the study of the Orok / Uilta language and folklore. In Alfred Majewicz (ed.) with the assistance of Larisa V. Ozoliņa, Mikhail D. Simonov, Tatyana Bulgakova, Elżbieta Majewicz, Tatyana P. Roon, Tomasz Wicherkiewicz & Werner Winter, *The collected works of Bronisław Piłsudski, Volume 4: Materials for the study of Tungusic languages and folklore*, 113–745. Berlin: Mouton de Gruyter.
- Pugh, Stefan M. & Ian Press. 1999. *Ukrainian: A comprehensive grammar*. London: Routledge.
- Saint-Exupéry, Antoine de [Антуану зе Сенту Экзупери (Antuaanu je Seentu Ekjuperi)]. 2016. *Нучйкэ принсу (Nučika prins) [The little prince]*. Sapporo: 北海道大学大学院文学研究科 [Graduate School of Letters, Hokkaido University]. [Translator: Elena A. Bibikova; Editor: Yoshiko Yamada; Supervisor: Toshiro Tsumagari.]
- Skirgård, Hedvig, Hannah J. Haynie, Harald Hammarström, Damián E. Blasi, Jeremy Collins, Jay Latache, Jakob Lesage, Tobias Weber, Alena Witzlack-Makarevich, Michael Dunn et al. 2023a. Grambank v1.0 (v1.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7740140> (Available online at <https://grambank.cld.org/>, Accessed on 2024-11-15.)
- Skirgård, Hedvig, Hannah J. Haynie, Damián E. Blasi, Harald Hammarström, Jeremy Collins, Jay Latache, Jakob Lesage, Tobias Weber, Alena Witzlack-Makarevich, Sam Passmore et al. 2023b. Grambank reveals global patterns in the structural diversity of the world's languages. *Science Advances* 9. DOI:10.1126/sciadv.adg6175
- Tsumagari, Toshiro. 2009. Grammatical outline of Uilta (revised). *Journal of the Graduate School of Letters, Hokkaido University* 4. 1–21.
- Tsumagari, Toshiro [津曲敏郎]. 2014a. Пiuсyтyкu 採集のウイльта語民話テキスト「ネズミの母親とカエルの母親」 [An Uilta folktale text by B. Piłsudski: “A rat mother and a frog mother” with grammatical analysis and Japanese translation]. *北方言語研究 [Northern Language Studies]* 4. 213–231.
- Tsumagari, Toshiro. 2014b. Remarks on the Uilta folktale text collected by B. Piłsudski. *北方人文研究 [Journal of the Center for Northern Humanities]* 7. 83–94. [Revised English version of Japanese original of 2013.]
- Tsumagari, Toshiro & Yoshiko Yamada. 2024. Uilta. In Alexander Vovin, José Andrés Alonso de la Fuente and Juha Janhunen (eds.), *The Tungusic languages*, 436–462. London: Routledge.
- Yamada, Yoshiko [山田祥子]. 2011a. Уйльта語北方言テキスト: 思い出話 2 編 [Uilta texts in Northern dialect: Two recollections]. *北海道民族学 [Hokkaido Journal of Ethnology]* 7. 48–59.

- Yamada, Yoshiko [山田祥子]. 2011b. ウイルタ語北方言テキスト: スルクタの作り方 [Uilta texts in Northern dialect: How to cook *sulukta*]. 北方言語研究 [*Northern Language Studies*] 1. 217–228.
- Yamada, Yoshiko [山田祥子]. 2011c. ウイルタの振楽器「ヨードプ」: 復元工程についてのウイルタ語北方言テキストを中心に [A type of shaken musical instruments *joodopu* of the Uilta: with the text in Uilta Northern dialect representing process of restoration]. 北方人文研究 [*Journal of the Center for Northern Humanities*] 4. 25–49.
- Yamada, Yoshiko [山田祥子]. 2012. ウイルタ語北方言テキスト: ありがとう、池上先生 [Uilta texts in Northern dialect: Thank you, Ikegami-sensei]. 北方人文研究 [*Journal of the Center for Northern Humanities*] 5. 159–172.
- Yamada, Yoshiko [山田祥子]. 2014. ウイルタ語北方言テキスト: アザラシ肉に関する体験談 [An Uilta text in Northern dialect: A story about seal meat]. 北海道民族学 [*Hokkaido Journal of Ethnology*] 10. 104–113.
- Yamada, Yoshiko [山田祥子]. 2015. ウイルタ語北方言テキスト: 人喰いお化けの話 [An Uilta text in Northern dialect: A story about ogre]. 北方言語研究 [*Northern Language Studies*] 5. 261–280.
- Yamada, Yoshiko [山田祥子]. 2016. ギシクタウダ (マリーヤ・ミヘエワ) の生涯: ウイルタ語北方言テキスト [The life of Gisiktauda (Marija Mikheeva): An Uilta text in Northern dialect]. 北方言語研究 [*Northern Language Studies*] 6. 179–201.
- Yamada, Yoshiko [山田祥子]. 2023. ウイルタ語北方言テキスト: お母さんのこと [A text in the Northern dialect of Uilta: Mother]. 北海道言語文化研究 [*Journal of Language and Culture of Hokkaido*] 21. 1–16.

