Acta Linguistica Academica 72 (2025) 4, 455–515
DOI: 10.1556/2062.2025.00810

1828
AKADÉMIAI KIADÓ

The syntax of numeral modification in Ainu

ÉVA DÉKÁNY*

ELTE Research Centre for Linguistics and ELTE (Eötvös Loránd University), Budapest, Hungary

Received: March 31, 2024 • Revised manuscript received: September 29, 2025 • Accepted: October 1, 2025

ABSTRACT

This paper explores the morpho-syntax of numerals in Ainu, a critically endangered Northeast Asian language isolate. I argue that Ainu simplex numerals exhibit the direct vs. indirect modification split discussed for adjectives in Cinque (2010). My findings thus corroborate the direct/indirect bifurcation among adnominal modifiers, and show that this split applies to more types of modifiers than previously thought. In the domain of complex numerals, I argue that in addition to multiplicative, additive and subtractive numerals, Ainu also employs a rare type called overcounting numerals. Complementing and extending Ionin & Matushansky's (2018) approach to additives and subtractives, I put forth the first analysis of the internal structure of overcounting numerals.

KEYWORDS

Ainu, numerals, overcounting, subtraction, object counting vs. serial counting, relative clause, classifiers, PPs

1. INTRODUCTION

This paper explores the structure of simplex and complex numerals in Ainu, a critically endangered Northeast Asian language isolate. While most Ainologists have commented on the fact that the language has an interestingly complex numeral system, a morpho-syntactic investigation of the patterns is still lacking. This paper seeks to fill this gap and examines the implications of the Ainu data for current theories of numerals.

I will start the investigation with Ainu simplex numerals. I will argue argue that these numerals (as well as possessors) exhibit the direct vs. indirect modification split discussed for adjectives in Cinque (2010). My findings thus corroborate the direct/indirect bifurcation among

* Corresponding author. E-mail: dekany.eva@nytud.hu

adnominal modifiers, and show that this split applies to more types of modifiers than previously recognized. I will then turn to complex numerals. I shall argue that in addition to multiplicative, additive and subtractive numerals, Ainu also employs a rare type called overcounting numerals. I will complement and extend Ionin & Matushansky's (2018) approach to additives and subtractives, and put forth the first analysis of the internal structure of overcounting numerals.

The paper is structured as follows. Section 2 familiarizes the English speaking reader with features of Ainu morpho-syntax that will be important in the subsequent analysis, and also explains where the data come from. Section 3 examines simplex numerals, including the relationship between abstract counting and object counting numerals. Section 4 looks at the constituency of additive numerals. Section 5 treats the structure of odd multiples of 10, arguing that they instantiate overcounting rather than subtraction. Section 6 is dedicated to true subtractive numerals. Section 7 rounds off the paper with the conclusions.

2. BACKGROUND TO THE AINU AND THEIR LANGUAGE

The Ainu, traditionally hunter-gatherers and one of the first inhabitants of the Japanese archipelago, used to live on the Island of Hokkaido, the north of Honshu, the south of Sakhalin as well as the Kurile islands and likely also the southern part of Kamchatka. Around 900 BCE they started to be pushed north by agricultural settlers from the Korean Peninsula. Japanese settlers arrived in Hokkaido in the 16th century (de Graaf 2016). They established rule over the Ainu, who faced discrimination even at the end of the 20th century (Nakagawa 1999).

Ainu was first documented in the 17th century, in the form of glossaries. In the early 20th century the language was still spoken on Hokkaido, Sakhalin and the Kuriles (Tamura 1999). With the 1945 Russian takeover of Sakhalin and the Kuriles, most Ainu relocated to Hokkaido. Sakhalin and the Kuriles now have no speakers left. Revival efforts on Hokkaido started in the postwar era, and in 1997 the Japanese government passed the 'Ainu Shinpo' (New Ainu Law), which recognizes the Ainu as a separate ethnic group and calls for respect of their culture (Nakagawa 1999). While the Ainu population today is estimated to be between 24,000 and 50,000 (de Graaf 2016), there are very few speakers. Tamura (1999) estimated that as of 1999, there were only twenty to thirty older speakers on Hokkaido who used the language when they were young and who could become fluent with some training. This is echoed in Bugaeva (2004, 7): "there are only a few persons aged 80–90 on Hokkaido who remember their native language. The Ainu language is practically extinct now. All the Ainus speak Japanese". While Southwestern Hokkaido Ainu has a small population of speakers/rememberers to date, fieldwork on Ainu is no longer possible (Bugaeva 2018).

Historically Ainu had three main dialects: Hokkaido Ainu (Japan), Sakhalin Ainu (Russia) and Kurile Ainu (Russia). Tamura (1988/2000) characterizes the little-documented Kurile Ainu variety as significantly different from the other two dialects, and according to Refsing (2014), Hokkaido and Sakhalin Ainu are mutually unintelligible. Hokkaido Ainu's two main dialect groups are Southwestern (further divided into southern and western) and Northeastern (further divided into northern, eastern and central); these have major differences (Bugaeva 2016, 2024).

This paper focuses on Southwestern Hokkaido Ainu (especially its Chitose and Saru varieties in the southern group), which is the best documented variety of the language (though as pointed

out in Refsing 1986, 111, "the numerals do not differ much from one Hokkaidō dialect to another"). The data come from previous descriptive-typological publications on this dialect (which are, in turn, based on oral tradition, as Ainu is traditionally an unwritten language) as well as Nakagawa et al. (2016–2024), a glossed audio corpus of Chitose and Saru Ainu folklore. The corpus data fall into two genres: 'prosaic folktales' (*uepeker*, abbreviated as UP) and 'divine epics' (*kamuy yukar*, abbreviated as KY). At the appropriate points, some examples were included from the Ishikari variety of the Central Hokkaido dialect and the Shizunai variety of the Eastern Hokkaido dialect; these are marked as such.

Not all sources on Ainu use the same transliteration system. Examples are presented as spelled in the original, but the glosses have been adapted to match the abbreviations used in Nakagawa et al. (2016–2024). When not explicitly cited from a different source, numerals appear in the transliteration of Tamura (1988/2000).

Previous work on Ainu was conducted in the descriptive-typological framework. Ainu is polysynthetic, incorporating, head-marking and has SOV basic word order. It has no case marking on arguments. The indexing of arguments on verbs, on the other hand, is obligatory, with mixed alignment: nominative/accusative in the first person singular, neutral in second and third person, and tripartite elsewhere. Pronouns are usually dropped; third person agreement is zero. Adjuncts are marked with postpositions. There are no tense markers; aspectual, modal and evidential markers appear in the post-verbal position.

There are three aspects of Ainu grammar that will recur in the glosses and the analysis, and which are therefore worth pointing out in advance. Firstly, the first person plural inclusive, the second person honorific, the indefinite proper (i.e., generic or unknown speaker or addressee) and logophorics share the same verbal indexing. As arguments, the NPs of these categories are all cross-referenced on the verb with *a*= for transitive subjects (A), *an*= for intransitive subjects (S) and *i*= for objects (O). Ainu studies use the cover-term 'indefinite person' or 'fourth person' to refer to the above-mentioned NP types. Following Nakagawa et al. (2016–2024), I will use the term 'fourth person' and will gloss *a*=, *an*= and *i*= as 4.A, 4.S and 4.O, respectively.

Secondly, the Ainu copula *ne* 'be, become' is a bivalent verb, which is primarily used in equative sentences with a nominal predicate and always relates two terms. It has unique agreement properties. (In persons with tripartite alignment) Ainu intransitive verbs take S agreement, while transitive verbs take A and O agreement. Unlike intransitive verbs, the copula takes A agreement, but unlike transitive verbs, it does not exhibit O agreement. Existential sentences and locative sentences with PP predicates employ the existential verb *an* 'to exist.SG' rather than *ne*. (Plural subjects trigger the suppletive plural allomorph *oka* 'to exist.PL'.) Sentences with 'adjectival' predicates generally do not feature either *ne* or *an*/*oka*, as 'adjectives' in Ainu are grammatically intransitive verbs (and show S agreement).

Thirdly, relative clauses are prenominal and employ the gap strategy. Since Ainu has no mophological marking of subordination on verbs, and consequently also lacks non-finite verbs, relative as well as complement clauses are inflected in the same way as the main verbs of matrix clauses.

An additional factor which will be relevant in the context of numerals, albeit primarily via its absence, is morphological number marking. Number marking in Ainu typically appears on verbs rather than nouns. While Tamura (1988/2000, 81) makes a categorical statement about the lack of pural marking on nouns, Bugaeva (2004, 19) and Bugaeva (2024, 580) point out that plural

marking is possible, though always optional, for nouns denoting humans and some nouns denoting animals. Number is typically indicated on verbs, in the form of pluractional marking. Only a subset of verbs can take this marking (some in the form of a suffix, some in the form of suppletion), however, and a semantically plural intransitive subject or object does not necessarily trigger pluractional marking on the verb (Tamura 1988/2000, 39–40). Therefore the morphologically singular form can be characterized as the unmarked default (Nakagawa 2022b, 585). NPs containing a numeral never trigger the plural form of verbs (Tamura 1970, 609).

Equipped with this background, we can now turn to the morphosyntax of numeral expressions in Ainu.

3. SIMPLEX NUMERALS

Ainu has simplex numerals for 1 through 5, for 10 and for 20.[1] As will become clear from the discussion below, 'simplex' is to be understood as not containing more than one numeral morpheme, but not necessarily as monomorphemic.

3.1. Direct vs. indirect modification among numerals

Truly monomorphemic numerals exist for *tu* 'two' and *re* 'three' only.

(1) tu pa, re pa
two year three year
'two years, three years' (Tamura 1988/2000, 255)

The other simplex numerals are construed with a verbal element: either the copula *ne*, or the existential verb *an*.

The expression for 1 is *sine*. Kindaichi (1931/1993, cited in Arapu 1977) suggests that it goes back to the reflexive prefix *si-* 'oneself', illustrated in (2), and the copula *ne*.

(2) si-kir-tektek hine isam
REFL-turn-momentary and not.exist
'[he] he turned around and left' (Nakagawa et al. 2016–2024, K7908032UP.108)

Kindaichi & Chiri (1936/1974, cited in Refsing 1986), on the other hand, identify the *si-* part of 1 as the word meaning 'real, original, genuine, indigenous', cf. (3) (see also Kindaichi 1960, cited in Ochiai 2021).

(3) si poro kamuy
true be.big/old bear
'a really big bear' (Nakagawa et al. 2016–2024, K8303243UP.241)

[1] 6 through 9 are subtractive from 10 and will be discussed later in the paper.

The word for 4 is *ine*, which both Kindaichi (1931/1993) and Kindaichi & Chiri (1936/1974) relate to *inne* 'many, be numerous' (see also Tamura 1988/2000, 260). Tamura (1988/2000, 254) suggests that the numeral root is *i-*, and the *ne* part is the copula.[2] This is supported by the fact that in the subtractive numeral 6 (*i-wan* lit. 'four-ten') only the first vowel of *ine* is kept, while *ne* is dropped (see also Kindaichi 1960, cited in Ochiai 2021).

The number 5 is *asikne*. Laufer (1917), Kindaichi (1931/1993, cited in Arapu 1977), Kindaichi (1960, cited in Ochiai 2021), as well as Kindaichi & Chiri (1936/1974, cited in Refsing 1986) relate this to *aske*, which means 'hand' in compounds (e.g., *askepet* 'finger' Refsing 1986, 111). The remaining part is the by-now familiar *ne*.

The expression for 10 is *wan*. Kindaichi (1931/1993, cited in Arapu 1977) and Kindaichi (1960, cited in Ochiai 2021) suggest that this form, in fact, combines the reciprocal suffix (normally *u-*) and the existential verb *an*. 'RECP-exist' designates that (the fingers of) both hands are to be counted. Finally, 20 is variably given in grammars as *hot* or *hot ne*. (We will come back to the two versions in Sect. 3.3.) According to Refsing (1986), *hot* means 'a set' and designates the set of the fingers and the toes together.

The composition of simplex numerals is summarized in (4). Although I will refer to the roots of the bimorphemic expressions in (4) as 'numerals', I take at least *asik* 'hand' and *hot* 'score/set' to fall into the morpho-syntactic class of nouns. In this, they are similar to English *dozen*, which has the distribution of nouns but its lexical-conceptual content is a quantity.

(4)

		Decomposition	Construction
one	*si-ne*	truly/oneself-COP	copular
two	*tu*		monomorphemic
three	*re*		monomorphemic
four	*i-ne*	$\sqrt{\text{four}}$(/many?)-COP	copular
five	*asik-ne*	$\sqrt{\text{aske}}$(hand)-COP	copular
ten	*w-an*	RECP-exist (i.e., both hands)	existential
twenty	*hot (ne)*	score/set (COP)	copular

In the previous literature (4) represents the etymologies of simplex numeral expressions but synchronically 1, 4, 5, 10 and 20 are not necessarily taken to be bimorphemic, and it is not discussed what a decomposition analysis would entail for the syntax of these numerals. Below I will suggest that a synchronic decomposition of 1, 4, 5, 10 and 20 into a stem and verbal element is fully justified, and compares favourably with a monomorphemic analysis.

I suggest that the monomorphemic numerals *tu* 'two' and *re* 'three' merge to the extended projection of the head noun directly.

(5) $[_{\text{NumP}}$ tu $[_{\text{NP}}$ pa]]
two year
'two years' (Tamura 1988/2000, 255)

[2] Laufer (1917) also proposes a bimorphemic analysis, but he simply calls *-ne* a suffix.

The bimorphemic 1, 4, 5, 10 and 20, on the other hand, correspond to N-modifying clauses (gapped relative clauses or other noun-modifying clauses, to be made more precise below). The clause is projected from the copular or existential verb, and the numeral stem is an argument of this verb.

(6) a. [$_{RC}$ __$_i$ si-ne] [$_{NP}$ pa$_i$]
truly/oneself-COP year
'one year'
(Tamura 1988/2000, 255, glosses added)

b. [$_{RC}$ __$_i$ hot-ne] [$_{NP}$ pa$_i$]
score-COP year
'twenty years'
(Tamura 1988/2000, 255, glosses added)

(7) [$_{clause}$ w-an] [$_{NP}$ pa]
RECP-exist.SG year
'ten years', lit. 'year such that both-there [in number]'
(Tamura 1988/2000, 255, glosses and paraphrases added)

In other words, the (lexeme with the meaning of a) numeral does not merge directly with a projection of the head noun, as in the case of *tu* 'two' and *re* 'three'. Instead, it is embedded in a larger structure, and the semantic relationship between it and the head noun is established indirectly, via the mediation of the verb.

N-modifying clauses headed by *ne* 'be, become' and *an/oka* 'exist' are independently attested, as in the b) examples below, thus the proposal in (6) and (7) falls into the range of available syntactic structures in the language.

(8) a. ne tonomat a-unúhu ne
this lady 4.(A)-mother.POSS COP
'the lady is my mother' (Tamura 1988/2000, 188)

b. [$_{RC}$ a-unúhu ne] tonomat
4.(A)-mother.POSS COP lady
'the lady who is my mother' (Tamura 1988/2000, 188)

(9) a. ene oka tumunci kamuy to or ta oka
that exist.PL demon god pond place at exist.PL
'That kind of demon-gods are in that pond.' (Tamura 1988/2000, 189)

b. [$_{RC}$ ene oka tumunci kamuy oro ta oka] to
that exist.PL demon god place.POSS at exist.PL pond
'a pond in which that kind of demon-gods are' (Tamura 1988/2000, 189)

My proposal for simplex numerals means that numerals can entertain two different structural configurations with the noun they modify. This type of scenario is already known from the

realm of adjectives. Cinque (2010) persuasively argues that adjectives can merge either in the specifier of a dedicated functional head or they can be embedded in (reduced) relative clauses. He calls the former structure direct modification and the latter type indirect modification, a terminology that I will adopt here. Thus 2 and 3 in Ainu involve direct modification, while the other simplex numerals involve indirect modification.[3]

In addition to the cross-linguistic parallel with adjectives studied by Cinque, there are Ainu-internal considerations, too, which render a split between direct and indirect modification numerals feasible. Firstly, a similar distinction is found among Southern Hokkaido Ainu possessors (Bugaeva 2004, 2012; Dougherty 2017; Bugaeva, Nichols & Bickel 2022). The possessors of inalienables (body part nouns, part-whole relations, some kinship nouns and a few other nouns) combine with the possessed noun directly. The possessum is marked with the possessive suffix (allomorphs: *-V, -hV*, and *-VhV*, with the quality of V determined by vowel harmony). In addition, the possessor is cross-referenced on the possessum by a prefix which is identical to the transitive subject (A) marking prefixes of verbs.

(10) e=ona-ha
2SG.(A)=father-POSS
'your(sg) father' (Nakagawa et al. 2016–2024, K7708241UP.030)

(11) eci=ona-ha
2PL.(A)=father-POSS
'your(pl) father' (Nakagawa et al. 2016–2024, K7803231UP.082)

(12) ∅-ona-ha
3.(A)-father-POSS
'his father' (Nakagawa et al. 2016–2024, K7803231UP.167, explicit zero suffix for third person added)

(13) a=ona-ha
4.(A)=father-POSS
'Father' ('fourth person' possessor, here: logophoric, hence, 'my father') (Nakagawa et al. 2016–2024, K7708241UP.032)

Essentially, the possessive suffix *-V, -hV*, and *-VhV* is non-agreeing head-marking and the 'A' prefix is agreeing head marking in the possessive construction. Such double marking is also found on possessors in Hungarian (Uralic, Bartos 1999; Dékány 2021) and Hixkaryana (Carib,

[3]It should be noted that indirect modification numerals in Ainu involve full clauses. Similarly to garden variety relatives such as (8) and (9), they show no reduction in verbal morphology compared to matrix clauses. In this they slightly differ from the indirect modification adjectives studied by Cinque.

Derbyshire 1985).[4] Building on earlier work on the morphology of the Hungarian possessive construction (Bartos 2000; É. Kiss 2002; Dékány 2015), I will assume that the non-agreeing possessive suffix heads a dedicated functional projection (PossP) in the extended Noun Phrase, with the possessor occupying spec, PossP.

(14) $[_{\text{PossP}}$ possessor $[_{\text{Poss}'}$ $[_{\text{NP}}$ possessum] Poss=*-V/hV/VhV*]]

The possessors of alienables, on the other hand, are introduced in relative clauses headed by the verb *kor* 'to have'. *Kor* is a transitive verb; its subject is the possessor and its object is the possessed noun. In (15) the 1SG subject is pro-dropped but is indexed on the verb; the indexing for the third person object is zero.

(15) poro seta ku-kor
big dog 1SG.A-have
'I had a big dog.' (Bugaeva, Nichols & Bickel 2022)

Relative clasues with *kor* have an object gap that is co-referential with the external head of the relative clause.[5]

(16) $[_{\text{RC}}$ (potentially dropped) subject $\__i$ Subj.clitic=*kor*] N_i

(17) $[_{\text{RC}}$ a=kor] pon seta
4.A=have be.small dog
'my small dog', lit. 'the small dog that I have ("fourth person" subject, here: logophoric)' (Nakagawa et al. 2016–2024, K8303243UP.195)

(18) $[_{\text{RC}}$ e=kor] pon seta
2SG.A=have be.small dog
'your(sg) dog', lit. 'the dog that you have' (Nakagawa et al. 2016–2024, K8303243UP.107)

[4]The Ainu agreeing head-marking is syntactically much more independent than its Hungarian counterpart, though, as it can be separated from the possessum by modifiers. (Note that *nispa* receives various different translations in Nakagawa et al. (2016–2024): 'rich man, wise man, master, elder', as well as 'old man'. In (i) it does not receive an explicit translation in the corpus.)

(i) **a**=hekote nispa ona-utar-**i** or-o ta arpa
4.A=be.married.to rich.man father-PL-POSS place-POSS at go.SG
'who went to our father's place' (Nakagawa et al. 2016–2024, K7908032UP.211)

The agreeing head-marking in Hungarian is a postnominal suffix; it cannot be separated from the possessum.

(ii) az én papír-**ja**-i-**m**
the 1SG paper-POSS-PL-1SG
'my papers' (own knowledge)

[5]Note that (17) and (18) are taken from the same folk tale, but the corpus translates *pon* 'small' only in the first case, while in the latter case it leaves *pon* out of the translation. This is not relevant for the point being made here.

Note that the external head is marked neither with the possessive suffix *-V/hV/VhV* nor with the agreeing 'A'-prefix. This is because there is no PossP in the structure. The head is modified by an RC only, and the understood possessor is introduced within this clause as the subject of *kor*. The possessive relationship is established between the subject of *kor* and the object gap; the relationship between the understood possessor and the external head is indirect. The syntax of possession thus confirms that in Southern Hokkaido Ainu the same modifier type can be construed in two different syntactic structures: direct modification and a relative-clause structure.

The second reason why assuming an underlying verbal modification structure for some numerals is plausible is that several concepts which Indo-European languages express with direct modification by NPs or APs are expressed by verbal constructions in Ainu. For instance, Ainu has no morphologically distinct class of adjectives (Bugaeva 2012; Nakagawa 2022a). The relevant concepts are expressed by intransitive verbs (they can take 'S' prefixes as well as causative and transitivizing suffixes, and they lack a comparative or superlative form). Recurring 'adjective endings' include *-ne* 'COP' and *-kor* 'have' (Refsing 1986), which again underscores that what is adjectival direct modification in Indo-European is expressed by a modifying clause in Ainu.

In addition, demonstratives are also built on the copula and the existential verb. So-called nominal or spatial demonstratives (used when the referent is visible to the speaker) comprise the adverbs *tap* 'this, like this', *ta* 'hey, here' and *to* 'hey, there' and the existential verb *an* 'to exist.SG'/*oka* 'to exist.PL' (data based on Tamura 1988/2000, 262).

(19)

	same place/present/new topic	nearby	distant
SG	*t-an*, *tap-an* (emphatic)	*ta-an*	*to-an*
PL	*tap-oká* (old fashioned)	*ta-oká*	*to-oká*

It is reasonable to assume that these, too, are relative clauses: 'the N that is here/there', and so the term 'nominal demonstratives' in Ainu designates a function rather than a word class.[6]

(20) a. ta-an kur, to-an kur
here-exist.SG person there-exist.SG person
'this person, that person', 'lit. the person that is here/there'

b. ta-oká utar, to-oká utar
this-exist.PL people that-exist.PL people
'these/those people', lit. 'the people who are here/there'
(data from Tamura 1988/2000, 262, glosses and paraphrases added)

The other group of demonstratives, i.e., the so-called anaphoric ~ cataphoric/conceptual demonstratives ('the referent we mentioned earlier'; essentially, discourse demonstratives), also feature either the copula or *iki*, the intransitive form of the verb *ki* 'do'. These demonstratives are listed below (see Tamura 1988/2000, 263–264). Discussion of the syntax of these demonstratives

[6]The morphological decomposition is made explicit in Tamura (1988/2000) but it is not discussed what sort of syntactic structure this corresponds to.

would take me too far afield. It suffices to say here that these 'demonstratives' also involve verb-mediated, indirect modification.

(21)		
	this	*ne* < COP *ne* lit. 'the one that is'
	self-same, discussed now	*ne wa an/oka* < COP *ne* + *wa* 'and' + *an/oka* 'exist.SG/PL'
	that/those, discussed earlier	*né-a/rok* < COP *ne* + perfect marker *a/rok* (SG/PL) lit. 'the one that was'
	that/those (animate N that did sth previously)	*iki-a/rok* < *iki* 'do' + perfect marker *a/rok* (SG/PL) lit. 'done'

As Indo-European adjectives and demonstratives correspond to verbal structures in Ainu, the treatment of 1, 4, 5, 10 and 20 as verbal structures fits into the overall picture of nominal modification in the language.[7]

While I suggested that *tu* 'two' and *re* 'three' combine with the head noun in a direct modification structure as in (5), it is also possible to merge these in relative clauses, similarly to the other quantity-denoting expressions. In (22) we have a double relative clause: 'its hand exists' modifies 'tree', and 'it is three (things)' modifies 'hand'.

(22) [[re-p ne] ni-tek-ehe an] cikuni
three-NMN COP tree-hand-POSS exist.SG tree
'a place where the branches formed a three-pronged fork' lit. 'a tree that there are branches which are three' (Nakagawa et al. 2016–2024, K8303243UP.048, glossing adapted)

Based on the corpus of Nakagawa et al. (2016–2024), this is clearly the exception rather than the rule (this is the only example of this kind), and no grammars mention *tu ne* and *re ne* as versions of 'two' and 'three' either. This example clearly shows, however, that a quantity-denoting expression can be merged in a relative clause structure, and when this materializes, the RC is headed by the equative copula *ne*. The relevant portion of the structure is schematized in (23).

[7]It is reasonable to ask why 1, 4, 5, 10 and 20 involve a different underlying structure than 2 and 3. Here I can only speculate on the reasons. In her reconstruction of Proto-Ainu numerals, Ochiai (2021) suggests that at the initial stage expressions existed only for 2 and 3, and the other numerals emerged later. If this is so, then it is perhaps not so surprising that the newer numerals were created via periphrasis. Cinque (2010) suggests that direct modifiers are functional and form a closed class, while indirect modifiers are an open class. If Proto-Ainu only had 2 and 3, and these constituted a closed class of functional direct modifiers, it is less of a mystery why the newly added numerals (some of which clearly originated as members of open, lexical classes, cf. (4)) had to be accommodated in an indirect modification structure. On the other hand, a reviewer points out that they find it more likely that Ainu first had expressions for 'one' and 'many' (recall that 4 is related to 'many, be numerous'), and the two paucal numerals were added at a later step. From a typological viewpoint, this is also a plausible scenario, as there are languages which are reported to lack exact numerals above 'one' (e.g., several pre-contact languages in the Amazon, Hammarström 2010).

(23) [$_{RC}$ ___$_i$ re-p three-NMN ne] COP ni-tek-ehe$_i$ tree-hand-POSS
lit. 'its treebranch which is three'

(22) thus provides an important piece of additional evidence for the plausibility of the RC analysis of 1, 4, 5, 10 and 20.[8]

To summarize the discussion so far, Ainu has simplex numerals for 1 through 5 as well as 10 and 20. These show a split between a direct modification structure (for 2 and 3) and a modifying clause structure (for 1, 4, 5, 10 and 20). In the remainder of this section I will focus on numerals with an indirect modification structure and show that there is a finer distinction to be made among them: while 1, 4, 5 and 20 involve a relative clause, the clausal structure behind 10 is slightly different. This correlates with the verbs that these numerals involve (copula vs. existential verb). This discussion will also provide further evidence for the view that the simplex numerals 1, 4, 5, 10 and 20 are synchronically bimorphemic. In order to make this point, I will compare simplex numerals in the object counting use and the serial counting use. To set the scene for that discussion, however, we must first clarify the status of two suffixes, *-p(e)* and *-iw/-n*, which accompany numerals in certain contexts.

3.2. Pronominal numerals and the so-called classifiers

Numeral modification in Ainu is prenominal by default. When a numeral is not prenominal, it must bear a special suffix. If the counted noun is [+human], the suffix is *-iw* after consonants and *-n* after vowels, while if the quantified noun is [–human], the suffix is *-pe* after consonants and *-p* after vowels (see also Section 6). (25) and (27) show that NP ellipsis triggers these suffixes. (Bugaeva 2012 does not provide the ungrammatical version of (27) with the suffix dropped, but it is clear from her text that the suffix is obligatory here as well.)

(24) tu cise
two house
'two houses' (Bugaeva 2012)

(25) tu-*(p)
two-p
'two things' (Bugaeva 2012)

(26) tu aynu
two men
'two men' (Shibatani 1990, 42)

(27) tu-n
two-n
'two' (people) (Bugaeva 2012)

[8]The role of the nominalizer *-p(e)* that accompanies 'three' will be clarified later. To preview the discussion, *-p(e)* is a bound noun. In (22) it provides a morphological host for the clitic numeral *re* 'three'.

(29) and (31) show that numerals can occur postnominally, and in this case, too, the suffixes are obligatory.

(28) asikne suma
five rock
'five rocks' (Shibatani 1990, 42)

(29) suma asikne-p
rock five-p
'five rocks' (Shibatani 1990, 42)

(30) sine acapo
one uncle
'one uncle' (Shibatani 1990, 42, Ishikari Ainu)

(31) acapo sine-n
uncle one-n
'one uncle' (Shibatani 1990, 42, Ishikari Ainu)

Since nominal modification is otherwise exclusively prenominal and postnominal numerals bear special morphological marking, I concur with Refsing (1986, 115) that such examples involve appositive modification.

(32) $[_{NP}$ suma] $[_{NP}$ $[_{RC}$ asik-ne]-p] 'rock, five (ones)' (=(29))

In their adnumeral use, *-iw/-n* and *-p(e)* are characterized in the literature either as nominalizers (Shibatani 1990) or as (quasi-)numeral classifiers (Batchelor 1905 [1888]; Bugaeva 2012). The classifier analysis has likely been inspired by Ainuists' familiarity with classifiers in Japanese (an important contact language for Ainu), and at first sight, it has much to recommend it. Firstly, the [± human] distinction is the first, most basic categorization of nominals that classifier languages make (Adams & Conklin 1973). In a language with only two classifiers, this is the distinction that we expect to emerge. Secondly, we have seen that *-iw/-n* and *-p(e)* appear in noun phrases with ellipsis, as in (25) and (27). In the languages that have them, numeral classifiers play an important role as anaphoric devices in elliptical contexts, allowing the elided noun to be easily recovered from the context (Bisang 1999). In Ainu, too, *-iw/-n/p(e)* aid the recovery of the ellipted noun by making explicit its [± human] feature. Thirdly, while in many well-known classifier languages (e.g., Mandarin or Thai) classifiers are free morphemes, there are also languages that feature morphologically bound classifiers that cliticize onto the numeral. Examples include Bangla (Indo-European, Simpson & Syed 2016) and Newari (Tibeto-Burman, Bhaskararao & Joshi 1985), among others. Ainu would thus fit in the latter group.

On closer inspection, however, the classifier analysis runs into serious problems. In all numeral classifier languages I am familiar with, it is possible for the classifier to occur in non-elliptical NPs as well. That is, there are NPs which comprise a numeral, a classifier and an overt noun. Numerals bearing the *-iw/-n/-p(e)* suffixes, on the other hand, occur only

pronominally ((25) and (27)) or post-nominally ((29) and (31)), and I argued that post-nominal numerals are also pronominal numerals (in appositive position). In other words, the Num-Cl N pattern, expected from a classifier language, is ruled out.[9]

The form *-p(e)* is widely used in Ainu as a bound noun with the generic meaning 'thing' (Bugaeva 2012) (occasionally also 'person', mostly in a pejorative sense, Bugaeva 2017a). As a bound element, it cannot occur on its own; it must combine with a noun modifier. (33) shows it with a demonstrative, (34) with intransitive verbs ('adjectives') and (35) with a relative clause headed by a transitive verb.[10]

[9]Refsing (1986) shows that in the Shizunai dialect (Northeastern Hokkaido) Num+*iw/n* can be used adnominally.

(i) Tu-n katkemat ukoitak kane apkas.
two-n woman chat CONTEMPCONJ walk
'Two women walked while chatting.' (Refsing 1986, 115) Shizunai Ainu

Num+*n* also occurs in adnominal position in West Sakhalin Ainu (Dal Corso 2021, 72). It is therefore plausible to regard *-iw/-n* as a classifier in these dialects. A reviewer suggests that this pattern may have emerged from a structure similar to (22), i.e., an indirect modification structure in which the numeral combined with the human bound noun *-iw/-n* within a clause headed by *ne*. Later on, *ne* came to be dropped from the structure, resulting in the current surface-form.

(ii) tu-n ne katkemat
two-N.human COP woman
'two women' (hypothetical example)

It must be emphasized, though, that Num+*p(e)* cannot be used adnominally in Shizunai Ainu or West Sakhalin Ainu either, so there is no evidence for the classifier analysis of *-p(e)* from any dialect.

[10]A relative clause headed by *-p(e)* can take on an idiomatic meaning.

(i) ci-ronnu-p
1PL.EXCL-kill.PL-thing/person
'fox' lit. 'things (that) we kill' (Bugaeva 2017a)

Bugaeva (2017a, 237) comments on such examples in the following way: "some relative clause heads are bound, so they tend to cliticize and eventually turn into derivational suffixes which occur in a great number of lexical nominalizations" and Bugaeva (2017b, 884) writes that "even lexical nominalizations look like finite clauses". It has not been demonstrated that the *-p(e)* in these examples has different properties from the *-p(e)* in the examples in the main text, and that 'lexical nominalizations' merely look like finite clauses rather than being actual finite clauses embedded under a semantically light noun (here: *-p(e)*). That the combination of a particular subject, object and verb can take on an idiomatic meaning is not unknown (cf. *the shit has hit the fan* or *the cat is out of the bag* in English) and is thus independent of whether the predicate is nominalized. Therefore I do not see it warranted to assume two different *-p(e)* morphemes. (Cf. also Bugaeva (2016, 96), acknowledging that "most authors recognize that there is no clear-cut distinction between" clausal and lexical nominalization.) That there is a bound noun *-p(e)* and a separate nominalizer *-p(e)* is also maintained for West Sakhalin Ainu in Dal Corso (2021). Although Dal Corso argues that "[t]he dependent noun *-pe* should not be confused with the homographous nominalizer *pe/*p*" (p. 48), he acknowledges a diachronic connection between the two uses, with the nominalizer developing from the bound noun (p. 78). However, no arguments are presented for the view that we are looking at separate listemes of the lexicon. I suggest that in the appropriate contexts, *-p(e)* can be viewed as a 'nominalizer' to the extent that it provides a nominal head to which the modifiers can attach, and as such it gives the modifier+*p(e)* constituent an external nominal distribution.

(ii) $[_{xNP}\ [_{RC}\ pro\ \_\__i$ ci-ronnu]-p_i]
1PL.EXCL-kill.PL-thing/person
lit. 'things (that) we kill', i.e., 'fox'

(33) a. tan cise
this house
'this house'

b. tan pe
this pe
'this' (Bugaeva 2012)

(34) a. pirka
'good'

b. pirka-p
'good thing'

c. husko
'old'

d. husko-p
'old thing' (Shibatani 1990)

(35) [nea nispa kor pa] p
that rich.man have PL thing/person
'the wise man's belongings' (Nakagawa et al. 2016–2024, K8109171UP.229)

I suggest that the *-p(e)* of pronominal numerals is none other than this bound noun.

As for the *-iw/-n* suffix, it has no similar wide-ranging use as *-p(e)*. The bound noun for people is *kur*, as illustrated below.

(36) a. *kur
man, person (Bugaeva 2017a)

b. tan kur
'this man/person' (Bugaeva 2017a)

c. [pet etok un] kur
river front belong.to person
'the man from the source of the river' (Bugaeva 2017a)

I assume that similarly to *-p(e)*, *-iw/-n* is also a bound noun. Its range of use is probably restricted to numerals because it is in competition with another bound noun with a similar meaning, i.e., *kur*, but I will have to leave a deeper study of *-iw/-n* for future investigation.

With the background on *-p(e)* and *-iw/-n* in place, we can now turn to a comparison between object counting numerals and serial counting numerals in Ainu. This will allow me to fine-tune the structure of numerals containing the copula *ne* on the one hand and the structure of those containing the existential verb *an* on the other.

3.3. Two routes to indirect modification, as revealed by serial vs. object counting forms

In their so-called object counting use numerals combine with a noun to express the noun's quantity.

(37) three cats

Abstract numerals, on the other hand, are used without making reference to the quantity of a particular noun: serial counting numerals are used in the counting sequence (38a) and arithmetical numerals are used in abstract arithmetical operations (38b).[11]

(38) a. one, two, three, four, …
b. Two plus three is five.

There is no consensus in the literature on how object counting and abstract numerals are related to each other. According to Ionin & Matushansky (2018), the object counting function is primary and structurally simpler, with abstract numerals formed from their object counting counterparts via nominalization.

(39) Ionin & Matushansky (2018)

a.	[[numeral] N]	object counting
b.	[[numeral] NMN]	abstract counting

Wągiel & Caha (2021), on the other hand, argue that abstract counting is primary and structurally simpler; object counting numerals are derived from the abstract counting counterparts via an additional overt or covert classifier.

(40) Wągiel & Caha (2021)

a.	[[numeral+CL] N]	object counting
b.	[numeral]	abstract counting

As far as Ainu is concerned, no information is available about arithmetical numerals (if abstract mathematical operations were used by the Ainu at all), but the forms of object counting and serial counting numerals are known, and they differ from each other in interesting ways. Lower and higher numerals show a different pattern, so I will discuss them separately.

Starting with the numerals between 1 and 4, the serial counting numerals appear to morphologically contain the corresponding object counting numerals and the *-p(e)* morpheme introduced above (data from Tamura 1988/2000, 255).

[11]Some languages use different formatives in object counting and serial counting contexts for the same number concept, see Hurford (2003), He (2015), Žoha, Wągiel & Caha (2022).

(41)

	serial counting	object counting with *pa* 'year'
one	*sine-p*	*sine pa* ('one year')
two	*tu-p*	*tu pa* ('two years')
three	*re-p*	*re pa* ('three years')
four	*ine-p*	*ine pa* ('four years')

Tamura (1988/2000) uses the term 'nominalized' to describe the serial counting numerals in (41). As already pointed out in fn. 10, in certain examples the literature on Ainu considers *-p(e)* to be a lexical (derivational) nominalizer. This nominalizer *-p(e)* is assumed to occur in expressions with less semantic transparency/more idiomaticity. However, there is no **syntactic** difference between this use of *-p(e)* and the *-p(e)* in (33)–(35), and so there is no need to assume two different *-p(e)* morphemes in the grammar of Ainu, one of them functioning as a bound noun and the other as a nominalizer. As a bound noun, *-p(e)* provides a nominal head for NP modifiers to attach to. The NP projected by *-p(e)* 'envelopes' the modifier and provides it with an external nominal distribution (42). In this sense, it 'nominalizes' the modifier without being a derivational affix.[12]

(42) $[_{xNP}$ modifier $[_{NP}$ N=*-p(e)*$]]$ xNP=extended NP

I suggest that this is exactly what *-p(e)* does with the numerals in (41).

In the case of 2 and 3, I proposed that these numerals combine with the head noun via a direct modification structure, without the mediation of a verb. In the serial counting form they must be accompanied by *-p*. Kirikae (2005, 2006), cited in Ochiai (2021), points out that in contrast to the other numerals, 2 and 3 are clitics (see also Tamura 1970, 607). I suggest that this accounts for the appearance of *-p*: since these numerals require a morphological host, they must always be accompanied by a noun. The semantically underspecified bound noun *-p* surfaces to fill this function. This is reminisent of *do*-support in English to the extent that the morpheme which is used to support a morphologically bound element is the one with the least semantic-conceptual content within its own grammatical category. What this means is that these numerals are always in adnominal position: the object counting and serial counting forms differ only in the semantic load of the head noun.

(43) $[_{NumP}$ tu $[_{NP}$ pa$]]$
two year
'two years' (Tamura 1988/2000, 255) object counting

(44) $[_{NumP}$ tu $[_{NP}$ p$]]$
two thing
'two' (Tamura 1988/2000, 255) serial counting

To some extent, the structure in (44) can be taken to instantiate (39b). The bound noun *-p(e)* acts as the N head in the structure and so provides an external nominal distribution to the

[12]On this analysis of certain 'nominalizations' in other languages, see Dékány & Georgieva (2020).

phrase containing the numeral. The motivation for this structure, however, is rooted in the Ainu-specific morphological properties of these numerals rather than in universalist structural considerations.

As for *sine* 'one' and *ine* 'four', I argued that they involve a noun modifying clause. I suggest that this is specifically a relative clause, with the copula *ne* as the RC predicate. The two arguments of this copula are the numeral and a gap; the gap is co-referential with and bound by the head noun. On this analysis it is immediately clear why *sine* and *ine* are always accompanied by a noun, a common noun in the object counting use or a bound semantically light noun, *-p*, in the serial counting use. It is this noun that provides the content of the gap and hence saturates the RC predicate (which, as an unsaturated predicate, could not stand on its own). In other words, 1 and 4, too, are adnominal even in the serial counting forms, but this follows from their syntactic structure rather than their morpho-phonological properties.[13]

(45) [$_{RC}$ ___$_i$ si-ne] [$_{NP}$ pa$_i$]
truly/oneself-COP year
'one year' (Tamura 1988/2000, 255) object counting

(46) [$_{RC}$ ___$_i$ si-ne] [$_{NP}$ p$_i$]
truly/oneself-COP thing
'one', lit. 'one thing' (Tamura 1988/2000, 255) serial counting

With numerals above 4, we see a different pattern: object counting and serial counting numerals either have the same shape (6, 7, 10) or the object counting form contains the serial counting form (5, 8, 9, 20).

(47)

	serial counting	object counting with *pa* 'year'
five	*asik*	*asik-ne pa* ('five years')
six	*iwan*	*iwan pa* ('six years')
seven	*arwan*	*arwan pa* ('seven years')
eight	*tupes*	*tupes-an pa* ('eight years')
nine	*sinepes*	*sinepes-an pa* ('nine years')
ten	*wan* (also the Japanese loan *to)*	*wan pa* ('ten years')
twenty	*hot*	*hot-ne pa* ('twenty years')

[13] As pointed out before, demonstratives are also built with the help of verbs, mostly with the copula and the existential verb. I suggested that these also involve relative clauses. This captures the fact that demonstratives cannot be used on their own either: they are either adnominal to a common noun or the light noun *-p(e)*, but cannot "make up a complete NP on their own" (Bugaeva 2012, 471).

(i) tan cise, tan pe, *tan
this house this NMN this
this house, this, this (Bugaeva 2012)

Note that an account which takes demonstratives to contain the copula diachronically but not synchronically (i.e., an analysis assuming reanalysis into a monomorphemic expression) cannot account for this pattern.

In the case of 5 and 20, the extra morhological piece of object counting numerals is the copula *ne*. 8 and 9 are complex numerals whose internal structure is subject to debate in the literature; we will look at them in detail in Sect. 6. For now, it suffices to observe that in the object counting form they involve the additional piece *an*, which is also the singular form of the verb 'exist'.[14] 5, 8, 9 and 20 thus have in common that the object counting form contains an added verbal part. These stand in contrast to 10 (as well as 6 and 7, which transparently contain *wan* 'ten'), where the object counting and the serial counting forms are identical.

How to make sense of this complex pattern, and square it with what we see with the numerals below 5? I suggest that in spite of appearances, the numeral expressions involving 1, 4, 5, 8, 9 and 20 exhibit two sides of the same coin in serial counting. In the object counting use they all involve a relative clause modifying the quantified noun. A relative clause is an unsaturated expression with a gap, and as such, it will not stand on its own. In serial counting, however, there is no specific noun to quantify, and so nothing to bind the RC gap, by definition. Southern Hokkaido Ainu solves this problem in two different ways. In the case of 1 and 4, the RC structure is retained and a semantically light, generic nominal, *-p*, is inserted as the head. In the case of 5, 8, 9 and 20, the problem of the unbound gap is solved by not projecting the RC structure to begin with. Here the quantity expression is used all by itself, without establishing any grammatical relation to a noun (be it contentful or light). This way neither a dummy N head, nor the copula or existential verb linking the quantity expression and the dummy N appears.

(48) $[_{RC}$ __$_i$ [asik]-ne$][_{NP}$ pa$_i$]
five-COP year
'five years' (Tamura 1988/2000, 255) object counting

(49) [asik]
five
'five' serial counting

It remains to be seen why low and higher numerals do not use the same strategy. Ochiai (2021) suggests that Proto-Ainu originally had expressions only for 2 and 3, with 1 and 4 added in the next stage, and the higher numerals emerging even later, in several steps. If this is on the right track, it may provide a partial understanding, if not explanation, of why there is a split between 1 and 4 and the higher numerals in serial counting.[15] It should also be pointed out that Ainu is

[14]Interestingly, Tamura (1988/2000) does not assume that the *an* in 8 and 9 is a morpheme (see the discussion in Sect. 6). In her analysis the serial counting forms of 8 and 9 are shorter but are not forms which "have dropped the suffix". In Sect. 6 I will argue for a structure in which the *an* form here is indeed a morpheme. Specifically, I will suggest that it is the very same *an* that also occurs in the serial counting forms of 6, 7, and 10.

[15]It is worth pointing out here that Pallas' (1977–1978 [1786–1789]) comparative dictionary registers the names of the numerals for 1 through 10 in Kurile Ainu, and they all end in *-p(e)*: *šinep, tup, rep, inep, ašinep, otъvamp, aravamp, tumisampe, šinimesampe*, and *oampe* (data published in Alonso de la Fuente 2021). This variety of Ainu thus applies a unified strategy to all numerals, construing them with a light noun whenever they do not quantify over a common noun. This confirms, to my mind, that there is no deep grammatical reason why Southern Hokkaido Ainu would make a split between 1 and 4 and the higher numerals.

not unique in low and high numerals having different grammatical properties. In languages with so-called quantificational case, numerals assign a particular case to the quantified noun, and a split between low and high numerals is not uncommon. In Inari Sami (Finno-Ugric) 2 through 6 assign accusative, while higher numerals assign Partitive to the noun (Nelson & Toivonen 2000); in Russian 2 through 4 assign genitive singular/paucal case while 5 and higher simplex numerals assign genitive plural (Babby 1985; Pesetsky 2013), in Czech only 5 and higher numerals assign case (genitive plural) to the noun (Caha 2015). A comparison of Ainu with other languages exhibiting a split between low and high numerals may bring rewards in future research.

To summarize, object quantification by 1, 4, 5, 8, 9 and 20 involves relative clause modification of N. With the numeral/quantifier expression sitting inside a clause, these examples instantiate neither the pattern in (39) nor the structure in (40) – both of which involve direct modification. In Ainu only 2 and 3 involve direct modification. For these numerals the NP projected by the light noun *-p* envelopes the numeral and so provides an external nominal distribution to it. In this sense, the serial counting forms are nominalized.

This leaves us to tackle the numerals 10, 6 and 7, which correspond to or contain *wan* 'ten', and for which the object counting and the serial counting forms are identical. As mentioned earlier, Kindaichi (1931/1993) traces *wan* 'ten' back to the reciprocal suffix (normally *u-*) and the existential verb *an*, designating that (the fingers of) both hands are to be counted. It is easy to understand why the serial counting form is not produced similarly to 5, 8, 9 and 20, by dropping the verbal part of the quantity expression: the reciprocal morpheme is a verbal prefix, which needs a supporting verb in syntax and at PF. But if the *an* 'exist' of *wan* cannot be dropped, why is the serial counting form not construed with the bound noun *-p(e)*, as in the case of 1 and 4?

With 1 and 4, I suggested that the quantity expression was embedded in a relative clause, and the gap in the RC had to be bound, hence, merger of an external nominal head was inevitable (45)–(46). If *wan* does not require such an external head, then it must be because it has no gap. This would be possible in two different ways. One logical possibility is that *wan* is synchronically not decomposable any more: it has become a monomorphemic numeral, similarly to 2 and 3. The numerals 2 and 3 require *-p* in serial counting because of their clitic status. *Wan* is not a clitic, however (either diachronically or synchronically), so there is no motivation for *-pe* to appear in serial counting. The other possibility is that *wan* in serial counting still involves a clause, but that clause is a complete, saturated expression, not a gapped RC. I follow the latter route and suggest that the *wan*-clause involves a predication structure between a dropped subject *pro* (they=both hands) and a silent PP THERE. This makes it a complete, saturated expression on its own, which may be linked to the noun in object counting by an understood *such that* relation (cf. *a house such that you like living in it*).[16]

(50) [$_{\text{PredP}}$ *pro* THERE w-an] [$_{\text{NP}}$ pa]
RECP-exist year
'ten years' (Tamura 1988/2000, 255, glosses and structure added)

[16]On silent THERE, cf. Kayne (2005) and Den Dikken & Dékány (2018).

This analysis is attractive because PP predicates generally take *an/oka* 'exist', not the copula *ne*, and so it gives us a handle on why this particular numeral (and the ones built on it) is not construed with *ne*. Note also that while the copula *ne* always takes two terms which it relates to one another, the existential verb *an/oka* is intransitive. Thus if it already has an argument in the form of a plural *pro*, it could not also take another, gapped argument. This means that the structure for 10 (and 6 and 7 built on 10) does involve clausal modification but it is not a relative clause.

In the serial counting version the PredP structure can stand on its own (51a), or alternatively, it can undergo zero nominalization (51b).

(51) a. [$_{\text{PredP}}$ *pro* THERE w-an]
RECP-exist
'ten', lit. 'both [hands] there' serial counting 1
(Tamura 1988/2000, 255, glosses and structure added)

b. [[$_{\text{PredP}}$ *pro* THERE w-an] ∅]
RECP-exist NMN
'ten', lit. 'both [hands] there' serial counting 2
(Tamura 1988/2000, 255, glosses and structure added)

The latter option would tie in *wan* 'ten' – and the numerals transparenty containing it, namely 6 and 7 – to what we see with 2 and 3. Zero (action) nominalization of verb phrases (transitive and intransitive alike) is independently attested in Hokkaido Ainu, and all intransitive verbs can undergo zero-nominalization to function as nouns with no concomitant morphological change (Bugaeva 2011a, 2016, 2017a, 2017b, 2018). At present, the available evidence is not sufficient to allow us to take a definite stand on (51a) vs. (51b).[17]

Let us summarize the discussion of simplex numerals. I suggested that 2 and 3 involve a direct modification structure, while the other simplex numerals are merged in an indirect modification structure. 1, 4, 5, and 20 are embedded in a gapped relative clause headed by

[17]As indicated in the main text, the data in this subsection come from Tamura's (1988/2000) grammar, which describes the Saru dialect. This is the only source, to my knowledge, which explicitly compares object counting and abstract counting numerals for 1 through 10 (and for some selected further numerals as well). Hamp (1968) also presents numerals from the Saru dialect, based on fieldwork he conducted in 1968. He does not discuss whether the numerals in his list are object counting or serial counting forms, but I presume that they are the latter type, as up to 9 they perfectly correspond to Tamura's serial counting forms (i.e., *-p(e)* up to 4, but not beyond that; missing *ne* in the case of 5). For 10, however, Hamp presents the form *waNpe*. There are also differences in the case of the odd multiples of 10: Tamura lists the *e-* forms to be discussed in Sect. 5, while in Hamp's list these are formed by addition (e.g., 30 is 'ten exceed twenty'). It is clear, then, that Tamura's and Hamp's informant(s) have slightly different grammars when it comes to numbers. For lack of evidence to the contrary, I will take this to be idiolectal variation. For Tamura's *wan* 'ten', I maintain the clausal analysis outlined in the main text. For Hamp's *waNpe*, on the other hand, I suggest that the informant has reanalyzed the clausal modifier into a monomorphemic word. This word could be a numeral proper, similarly to 2 and 3, with the light noun added due to analogy with these numerals. Alternatively, picking up on a suggestion made by a reviewer, the reanalyzed *wan* could be an intransitive verb 'be ten'. In this scenario, it modifies the noun in an RC structure, such that the single gapped argument of 'be ten' is coreferential with the nominal head: [$_{\text{xNP}}$ [$_{\text{RC}}$ ___$_i$ *wan*] N$_i$]. In this case *-pe* is added for the same reason as with other numerals expressed via RCs: it saturates the argument position of the RC-internal predicate. The proposed clause-to-morpheme reanalysis is a structural simplification. The extension of the additive (*ikasma* 'exceed') strategy to the odd multiples of 10 likewise reflects a somewhat simplified grammar for numbers.

the copula. 10 (and 6 and 7 transparently containing 10), on the other hand, has a structure in which the clausal modifier is a fully saturated expression: it is projected from the existential verb which takes a *pro* argument.[18] The bimorphemic, clause-based approach to quantity expressions involving indirect modification allowed us to capture the differences between the object counting and serial counting forms. It is hard to see how these differences could fall out from a synchronically monomorphemic approach to these numerals.

The analysis in this section corroborates the existence of the direct vs. indirect distinction among adnominal modifiers put forth by Cinque (2010). It further shows that this split applies to more types of modifiers than previously thought: in Southern Hokkaido Ainu, both numerals/quantity expressions and possessors avail themselves of the two distinct ways to combine with the head noun.

The simplex numerals discussed in this section form the building blocks of all other numerals. In the following sections we will examine complex numerals in detail. We will look at additive numerals as well as subtractives (6–7), and will revisit the debate on the make-up of odd multiples of 10 as well as 8–9.

4. ADDITIVE NUMERALS AND CONSTITUENCY

The constituency of additive complex numerals has sparked a lively debate in the literature. The classic WISYWIG approach is that the two components of the complex numeral, i.e., the augend and the addend, form a constituent to the exclusion of the noun. That is, additive numerals involve a relationship between individual numerals, and only one token of the head noun is present in the underlying structure (52). This is taken to be the default structure of additive numerals cross-linguistically in He et al. (2017), for instance.

(52) [twenty (connector) two] books

The opposite approach is that the components of additive numerals do not form a constituent to the exclusion of the quantified noun. Additives involve a relation between two NumPs, each NumP comprising an individual numeral (the augend or the addend) and a token of the quantified noun. This is adopted as the cross-linguistic underlying structure of additives in Ionin & Matushansky (2006, 2018).

(53) [twenty books] (connector) [two books]

Ionin & Matushansky (2006, 2018) argue that languages show variation in the i) (c)overtness of the connector, ii) the category of the connector (conjunction, P, or other category) and iii) which, if any, of the identical head nouns is deleted on the surface.

(54) a. [twenty books] (connector) [two books]
b. [twenty ~~books~~] (connector) [two books]
c. [twenty books] (connector) [two ~~books~~]

[18]I will defer a detailed discussion of 8 and 9 until Sect. 6; these are complex numerals in Ainu.

There are also approaches that acknowledge the validity of both analyses, suggesting that the choice between them may be a matter of cross-linguistic variation, and there may even be language-internal variation depending on the type of complex numeral or the diachronic stage of the language (Hurford 1987; Žoha, Wągiel & Caha 2022).

In Ainu, additive numerals feature the connector *ikasma* 'exceed' (to be discussed below), and, at least in their object counting use, their constituency is clearly that in (53). Firstly, it is possible and natural to spell out the quantified noun after both the augend and the addend.

(55) a. sine suma ikasma wan suma
one stone exceed ten stone
'eleven stones' (Tamura 1988/2000, 257)

b. sine pa ikasma wan pa
one year exceed ten year
'eleven years' (Tamura 1988/2000, 255)

Secondly, it is possible to spell out the quantified noun only at the end of the phrase, after the augend, but in this case the non-final component of the complex numeral, i.e., the addend, must be followed by *-p(e)* for non-human and *-iw/-n* for human nouns.

(56) a. sine-p ikasma wan sike
one-NMN exceed ten bag
'eleven bags' (Tamura 1988/2000, 257)

b. sine-n ikasma wan poyson
one-NMN exceed ten child
'eleven children' (Tamura 1988/2000, 257)

Importantly, this is so even for numerals which do not take *-p(e)* in the counting sequence, and which thus can, in principle, stand on their own, without being construed with a common noun or a light noun. A case in point is *iwan* 'six'. The object counting form and the serial counting form for this numeral are both *iwan* (47). Consider now (57), where the numeral 'seventy six' is expressed as 'sixteen toward eighty' (see Section 5 for discussion). The numeal 'sixteen' is lit. 'six exceed ten', and, crucially, the addend *iwan* 'six' is followed by the bound noun *-pe*.

(57) iwan-pe ikasma wan-pe e-ine-hot
six-NMN exceed ten-NMN to-four-score
'seventy six', lit. '[six exceed ten] to(wards) eighty' (Tamura 1988/2000, 255, glosses have been adatped)

As discussed above, numerals followed by *-p(e)* and *-iw/-n* are in adnominal position to a light noun, so Num+*p(e)* and Num+*iw/n* correspond to NumPs. Ainu thus consistently and

exclusively uses the strategy in (53) to form additive numerals; it does not appear to have the wherewithal to build a phrase like (52).[19]

Other languages in which the same pattern can be observed include Luvale and Biblical Welsh.

(58) a. mikoko makumi atanu na-mikoko vatanu
sheep ten five and-sheep five
'fifty-five sheep' (Ionin & Matushansky 2018, 122) Luvale (Bantu)

b. saith mlynedd ac wyth gan mlynedd
seven years and eight hundred years
'eight hundred and seven years' (Hurford 1975, 175, Genesis 5.7) Biblical Welsh

These findings do not mean, of course, that (53) is the only cross-linguistically valid structure underlying additives. They do show, however, along with languages having so-called discontinuous numerals (schematized in (54c) and exemplified below), that cross-linguistic unification in the direction of (52) is not possible.

(59) a. mærə́w médd-æn ə́d sæmmos
ten.M man-M.PL with five
'fifteen men' (Ionin & Matushansky 2018, 117) Tamashek (Berber)

b. seachd fir deug
seven man ten
'seventeen men' (Meinunger 2015) Scottish Gaelic

Let us now turn to the question of how the extended NPs (xNPs) corresponding to the augend and the addend are connected in Ainu. As mentioned above, additives feature the overt connector *ikasma*, which Tamura (1988/2000) glosses as 'exceed'. According to *a topical dictionary of conversational Ainu* (NINJAL 2015), this is a derived form: it comprises the verb *kasma* 'leave, exceed' and the antipassive prefix *i-*.

(60) tupesan pa i-kasma hotne pa ku=ne ruwe ne.
eight year APASS-exceed twenty year 1SG.A=COP INFREV COP
'I am twenty-eight.' (NINJAL 2015)

Since the majority of numerals are connected to the head noun via a verb (Sect. 3), it is not surprising that the connector within additives is also based on a verb. However, *ikasma* clearly does not head a gapped relative clause. As an antipassive form, it is intransitive and has only one argument: its subject is the xNP that corresponds to the addend and precedes it (in (60), *tupesan pa* 'eight years'). I suggest that the addend and the connector form a constituent. This

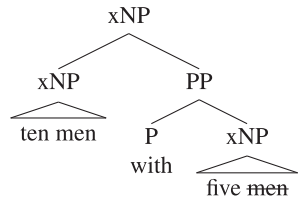
[19]Notice also that *-p(e)* or *-iw/-n* are not just blindly added after all non-final numeral words; rather, they are included after each component of additives. In multiplicative numerals the multiplier never takes these suffixes (cf. *ine* 'four' in (57)). This is expected based on the general consensus that the multiplier takes the multiplicand as its complement.

constituent is a full clause, in which the addend's xNP corresponds to the subject and the connector to the intransitive verbal predicate, and this clause is adjoined to the xNP of the augend (61).

(61) [$_{xNP}$ [$_{clause}$ [$_{xNP}$ tupesan pa] i-kasma] [$_{xNP}$ hotne pa]]
eight year APASS-exceed twenty year
'twenty-eight years' (lit. 'eight years be in excess, twenty years')

A similar adjunction structure has been proposed for additives connected by a P in Ionin & Matushansky (2018, 143). In their analysis, the Tamashek example in (59a) would have the structure in (62).

(62)



The Ainu structure differs from this in the direction of adjunction, the category of the adjunct (a clause rather than a PP), and in not deleting either of the nouns on the surface.

5. ODD MULTIPLES OF 10: OVERCOUNTING, NOT SUBTRACTION

Dialects of Ainu differ in whether they use a base-10 (decimal) or base-20 (vigesimal) numeral system (Laufer 1917; Refsing 1986, 113; Dal Corso 2021, 69; Ochiai 2023). West Sakhalin Ainu (Rayciska dialect), Kurile Ainu as well as the Asahikawa and the extinct Matsumae dialects of Hokkaido Ainu employ a decimal system. Other Hokkaido dialects and pre-18th century Kurile Ainu, on the other hand, are vigesimal varieties. The base-20 system has been reconstructed for Proto-Ainu in Ochiai (2021) (see also Laufer 1917). In this section we will examine the vigesimal system of Southern Hokkaido Ainu.

Even decimals (20, 40, 60, etc.) are formed by multiplication of 20 (Tamura 1988/2000, 255).

(63) hot, tu hot, re hot, ine hot
score two score three score four score
'twenty, forty, sixty, eighty' (Hattori 1964, Saru dialect)

Recall that *hot* 'score' is a noun. I assume with Ionin & Matushansky (2018) that in multiplicative numerals the multiplier takes the multiplicand as its complement. *Tu hot* 'forty' and *re hot* 'sixty' thus have the structure in (64).

(64) [$_{NumP}$ tu/re [$_{NP}$ hot]]
two/three score
'two/three scores'

When used for object counting, (64) fills one argument of the copula *ne*, and the gap (coreferential with the head noun) corresponds to the other argument.

(65) [$_{RC}$ ___$_i$ [tu-hot]-ne] [$_{NP}$ pa$_i$]
two-score-COP year
'forty years (Tamura 1988/2000, 255, glosses added)

In the case of *ine hot* 'eighty' and *asikne hot* 'one hundred' the multiplicand 20 likewise combines with the multiplier compositionally (66), and the resulting xNP functions as one of the arguments of the copula in object counting (67).

(66) [$_{RC}$ ___$_i$ [asik]-ne] [$_{NP}$ hot$_i$]
five-COP score
'five scores' (Ochiai 2021, glosses added)

(67) [$_{RC}$ ___$_i$ [asik-ne-hot]-ne] [$_{NP}$ pa$_i$]
five-COP-score-COP year
'one hundred years (Hokkaido University 2022, glosses added)

The odd multiples of 10 (i.e., 30, 50, 70, etc.), henceforth: odd decimals, are built on the template '*wan-pe* (= 'ten-NMN') + *e-* morpheme + next-higher multiple of 20'. That is, 30 is '10 *e*-40', 50 is '10 *e*-60', 70 is '10 *e*-80', etc.

(68) wan-pe e tu hot, wan-pe e re hot, wan-pe e ine hot, wan-pe
ten-NMN e two score ten-NMN e three score ten-NMN e four score ten-NMN
e ashikne hot
e four score
'thirty, fifty, seventy, ninety' (Hattori 1964, Saru dialect)

The identity of the *e-* morpheme in these numerals is subject to debate. *E-* occurs outside of numeral expressions as well, with several different uses. It functions as i) the second person singular verbal prefix cross-referencing A, S and O arguments, ii) an allomorph of the unproductive lexical causative suffix, iii) an applicative morpheme, and iv) a directional 'to' prefix. Previous accounts have pursued a unification of the *e-* in numerals with either the applicative *e-* or the directional *e-*. Here I will compare the two approaches in detail and provide novel syntactic arguments for unification with the directional prefix.[20]

[20]In his cross-linguistic overview of numerals, Greenberg (2000) states that *e-* in Ainu odd decimals expresses negation. It is unclear where this claim comes from, as *e-* has no link to negation. The negative particle of Southern Hokkaido Ainu is *somo*, which shares no phonetic material with *e-*. Given the lack of empirical foundations for Greenberg's claim, I will not discuss it any further.

5.1. Against the subtractive-applicative analysis

Ainu has three applicative prefixes: *e-*, *ko-*, and *o-*. These promote oblique PPs into core object NPs. As described in Bugaeva (2010, 757), "each prefix is highly polyfunctional, i.e., the actual referential range of the respective applicative constructions stretches over several different semantic roles and the exact role is attributed to the interaction between the semantics of the applicative and the semantics of the verb". The object of *e-* applicatives may have the semantic role of Content, Location, Instrument, Theme, Cause/Purpose, Comitative co-agent, Beneficiary, Path or Manner (see Bugaeva 2010 for details).

Kindaichi (1931/1993) and Chiri (1942/1973), both cited in Bugaeva (2010), take the *e-* of odd decimals to have a comitative (applicative) function, thus they interpret e.g., 30 as 'together with ten, forty'. Kindaichi & Chiri (1936/1974), cited in Bugaeva (2010), view the *e-* of odd decimals as an instrumental (applicative), that is, 30 is taken to be 'by 10, 40'. Tamura (1988/2000) translates and glosses *e-* as 'with (by adding)', while Refsing (1986) uses the gloss 'thereby'. They do not take a stand on whether *e-* is comitative or instrumental. *A topical dictionary of conversational Ainu* (NINJAL 2015) also glosses the *e-* of odd decimals as 'with.APPL', without coming down on the side of either the comitative or the instrumental interpretation. I will apply the umbrella term 'with-applicative' to these analyses. The comitative and instrumental applicative uses of *e-* are illustrated below.

(69) kira, e-kira
escape APPL-escape
'escape, escape with somebody' (Bugaeva 2010, 780)

(70) raige, tam e-raige
kill, sword APPL-kill
'kill, kill with a sword' (Batchelor 1905 [1888], 86)

(71) tek ari kar-pe
hand with make-NMN
'things made by hand' (Shibatani 1990, 66)

(72) tek-e-kar-pe
hand-APPL-make-NMN
'hand-made goods' (Shibatani 1990, 66)

All proponents of the 'with-applicative' analysis also hold that odd decimals involve the operation of **subtraction**.[21] Subtractive numerals are well-known from languages such as Latin, Biblical Welsh or Finnish (Hurford 1975). Complex numerals with this operation anticipate the **next-higher** multiple of the base and subtract a certain value from it. An example is given in (73).

[21]The subtractive view is extremely widespread; there are many works that simply take *e-* to be a subtractive particle without trying to establish a connection with the uses of *e-* outside of complex numerals. See Pott (1847), Chamberlain & Batchelor (1887), Batchelor (1905 [1888]), Laufer (1917), Peng & Brainerd (1970), Hurford (1975), Simeon (1968) and Patrie (1978).

(73) un-de-viginti
one-from-twenty
'nineteen' Latin

Subtractive numerals stand in contrast to additive numerals, which make reference to the **next-lower** multiple of the base and add a certain value to it.

(74) neun-zehn
nine-ten
'nineteen' German

It is indeed the case that Southern Hokkaido Ainu odd decimals anticipate the next-higher multiple of the base. 30, for instance, is expressed with the help of 40, the next-higher multiple of the base 20. That said, the 'with-applicative' analysis as well as the view that odd decimals are subtractive face serious challenges. In what follows, I will uncouple the 'with-applicative' analysis of *e-* from the subtractive view. First I will argue on semantic grounds that a *with*-connector cannot produce subtractive numerals. Therefore adopting the 'with-applicative' analysis of *e-* forces an analysis of odd decimals in which an operation other than subtraction is involved. Then I will level morpho-syntactic arguments against treating *e-* as an applicative morpheme in general (whether a 'with-applicative' or otherwise).

5.1.1. Against subtraction. Let me start by demonstrating that the 'with-applicative' analysis of *e-* is inherently incompatible with the view that odd decimals are subtractive. 'With'-phrases are cross-linguistically widely attested in complex numerals. Importantly, they always yield additive numerals rather than subtractive ones. This is because 'with' has an **inherently** additive semantics (cf. the cross-linguistically widespread use of *with*-coordination). Illustrative examples are given below.

(75) mærə́w médd-æn ə́d sæmmos
ten.M man-M.PL with five
'fifteen men' (Ionin & Matushansky 2018) Tamashek (Berber)

(76) thanl tapxa cöquiih
being.ten being.three is.with (= have)
'(be) thirteen' (Ionin & Matushansky 2018) Seri (isolate, Mexcio)

The connectors of subtractive numerals, on the other hand, are typically caritive ('without') morphemes or source ('from') Ps.

(77) onid un mlwydd cant
unless one year.of.age hundred
'ninety-nine years' (Ionin & Matushansky 2018) Welsh

(78) un-de-viginti pueri
one-from-twenty boy.PL
'nineteen boys' (Ionin & Matushansky 2018) Latin

If Ainu odd decimals contain a 'with' connector, it must be prevented from additively linking the two parts of the numeral. Otherwise the phrase in (79) would mean 50 rather than 30.

(79) wan-pe e tu hot
ten-NMN e two score
'thirty'

This problem has not been addressed in the literature. Bugaeva (2012, 498) writes that (79) has the "implied reading 'forty minus ten', but [it is] lit. 'together with ten forty/on ten – forty'", but she does not discuss the issue in more detail. The way the literal and implied meanings are assumed to part ways here seems to suggest that we have an idiomatic expression. Of course, it is not impossible that certain complex numerals may have a non-compositional meaning. However, ours is a peculiar scenario in which out of three ingredients, namely the numeral 10, *with*, and the numeral 40, only one, namely *with* does not appear with its literal reading; moreover, the idiomatic/implied reading ('from, minus') would be exactly the opposite of the literal, garden-variety reading ('with, adding').

I suggest that it is, in principle, possible to preserve the view that *e-* means 'with' and get a compositional semantics for (79), but it requires an *additive* analysis of odd decimals. On this view, Ainu odd decimals are unusal in that unlike Tamashek and Seri *with*-additives, they do not name the augend and the addend; instead, they name the addend and the **sum** of the addition (lit. '[the current number] with 10, [it] becomes 40'). In order for this meaning to emerge, we need a syntactic structure in which 10 and 40 do not form a constituent in (79) (otherwise 'with' will inevitably link them up in an additive manner). This can be achieved if the 'with'-phrase is left-extraposed, as in (80) from English.

(80) With (another) 10, you will/would have 40.

In (80) *have 40* is the main predicate, and the *with*-phrase is a modifier of this predicate. Although (80) indicates that 10 more of something is needed to get the desired number 40, it does not express subtraction. Instead, with its usual additive semantics, *with* links 10 and the phonologically silent but understood number or sum that one currently has. (80) is thus equivalent to (81): the part before the comma contains the augend and the addend, and the part after the comma names their sum. Crucially, 40 is not a subtrahend here.

(81) [what I have/the current value] with 10, will/would yield 40

The numerical value of odd decimals (i.e., 30, 50, 70, etc.) is lower than the next higher multiple of the base that they reference (e.g., the overall value of (79) is less than two scores). This should be indicated by placing the predicate (the sum of the addition) into an irrealis context: this is a value that is not currently available. As already said, all odd decimals are built on a multiple of 20, and 20 involves a relative clause structure with the copula *ne* (Sect. 3):

(82) [RC ___i hot-ne] [NP pa$_i$]
score-COP year
'twenty years'
(Tamura 1988/2000, 255, glosses and paraphrases added)

In addition to its stative reading 'be', the Ainu the copula *ne* also has the inchoative reading 'become'. The latter meaning is illustrated below from the Saru and Shizunai dialects.[22]

(83) suy aynu ne hine ki kor an
again human COP and do and exist.SG
'He had transformed into a human again.'
(Nakagawa et al. 2016–2024, K7708242UP.107)

(84) Kotan konnispa ne.
settlement chief COP
'Be/become the chief of a settlement.' (Refsing 1986, 145) Shizunai Ainu

I suggest that with its 'become' meaning mobilized, the copula could be the source of the irrealis mood we are looking for. Thus 30, for instance, could be envisioned as (85).[23]

(85) wan pa e tu hot ne pa
ten year with two score COP year
'thirty years', lit. '[the current value] with ten years, becomes 40 years'

To summarize, the 'with' analysis of *e-* and the subtractive view of the odd multiples of 10 cannot be reconciled. This is because 'with' is an inherently additive linker, and so it could not be pressed into service for subtraction.[24] On the 'with' analysis, *e-* could additively link the linguistically unexpressed 'current value' with 10. In this case the next-higher multiple of the base would name the value resulting from the addition (81).

5.1.2. Against the applicative analysis. In the foregoing discussion I argued against the subtractive analysis of odd decimals. I will now show what it would mean in syntactic terms to treat *e-* in the odd multiples of 10 as an applicative (whether a 'with'-applicative or other applicative) and why this analysis is untenable in light of what we know about applicatives in general and in Ainu in particular. Consider what the applicative analysis would entail for a phrase like (86).

[22]Bugaeva (2018) notes that in fact most Ainu stative vebs have an additional inchoative meaning.

[23]In addition, the RC structure itself could also be partly responsible for the irrealis mood, even without the copula. In English, for instance, *to* infinitives have an irrealis reading without involving a morpheme that the irrealis could be specifically tied to, e.g., *a man to fix the sink*. I thank Marcel den Dikken for bringing this to my attention.

[24]Note also that indubitably subtractive numerals in Ainu do not contain the *e-* linker, see Sect. 6.

(86) wan pa e tu hot ne pa
ten year with.APPL two score COP year
'thirty years'

Ainu applicatives are prefixes to verbs, which may potentially be morphologically quite complex. Odd decimals like (86) do contain a verb, namely the copula. Since the *e-* morpheme is not directly adjacent to *ne*, we must treat the *tu hot* 'two score' part of the numeral as incorporated. Incorporation as well as applicativization of verbs with incorporated Ns are independently attested in Ainu.

(87) u-ko-po-kor ka an wa
RECP-with.APPL-NI.child-have even exist.SG and
'Finally, we had children.' (lit. 'there was having children with each other')
(Nakagawa et al. 2016–2024, OI990302UP.334)

In the applicative analysis, 10 (*wan pa*) is the applicative object. Applicative objects come in two types: they are related either directly to an argument of the base verb (low applicative) or to the whole event (high applicative). As discussed above, 10 should not be related directly to 40, because connecting them by an instrumental or comitative would inevitably produce 50 rather than 30. Thus we cannot be dealing with a low applicative. This is confirmed by the observation that low applicatives are generally associated with a transfer-of-possession semantics (Pylkkänen 2008), which is lacking here. If (86) is to be treated as an applicative expression, it must be a high applicative, where the whole event, *become 40*, is related to 10 (in accordance with (80), (81) and (85) in the previous subsection). Thus the structure should be (88).

(88) $[_{\text{ApplP}}$ wan pa [Appl=e $[_{\text{VP}}$ tu-hot$_{\text{NI}}$+V=ne]]]

This view has four serious problems. Firstly, while applicativization and noun incorporation do indeed exist and may even combine in Ainu (87), Bugaeva's (2015) corpus study has found that neither applicativization nor noun incorporation is attested in copular VPs with *ne*. The only valency changing operation that can affect *ne* is direct causativization. As the applicative analysis posits applicativization by definition, and must also crucially rely on noun incorporation, it cannot be upheld.

Secondly, applicative phrases generally have full PP minimal pairs. It must be aknowledged that Ainu is exceptional in that certain applicatives have no PP paraphrases. However, these are restricted to Patient and Theme applicatives with *e-* and Recipient/Beneficiary applicatives with *ko-* (Bugaeva 2010). 'With'-applicatives do alternate with PPs headed by *tura* (comitative) or *ani* (instrumental).[25]

(89) oraun a=utar-i utar tura
then 4.(A)=kin/people-POSS PL together.with
'and together with the villagers' (Nakagawa et al. 2016–2024, K7908051UP.301)

[25]In Tamura (1988/2000) and Bugaeva (2012) the instrumental is *ani*, but in Shibatani (1990) it systematically takes the form *ari*. As (92) is cited from the latter work, it also features the *ari* form. This has no bearing on the issue at hand.

(90) u-ko-po-kor ka an wa
RECP-with.APPL-NI.child-have even exist.SG and
'Finally, we had children.' (lit. 'there was having chidren with each other')
(Nakagawa et al. 2016–2024, OI990302UP.334)

(91) tek ari kar-pe
hand with make-NMN
'things made by hand' (Shibatani 1990, 66)

(92) tek-e-kar-pe
hand-APPL-make-NMN
'hand-made goods' (Shibatani 1990, 66)

No full PP alternates have been reported for odd decimals, however.

Thirdly, fractions also contain the *e-* morpheme. These comprise 'half' or 'middle, halfway', the *e-* morpheme and the next higher integer (93)–(94). Some sources report them with 'half' or 'middle, halfway' dropped (95).

(93) emko e tu-p, emko e re-p
half e two-NMN half e three-NMN
'one and a half, two and a half' (Batchelor 1905 [1888], 102)

(94) hontom e-tu-p
middle e-two-NMN
'one and a half' (NINJAL 2015)

(95) e-tu-p, e-re-p, e-ine-p
e-two-NMN e-three-NMN e-four-NMN
'one and a half, two and a half, three and a half'
(Batchelor 1905 [1888], 96, also confirmed in Simeon 1968, 46)

Odd decimals and these fractions have in common that they name a point halfway between two distinguished points in the counting series: odd decimals name the halfway point between multiples of the base 20 and the fractions at hand name the halfway point between individual steps (digits) in the counting series. As the reader will recall, 2 and 3 combine with the counted noun via direct modification, without an overt or covert verb mediating the relationship. Since these expressions do not involve a verb, they cannot feature an applicative morpheme either.

Finally, we have seen that in serial counting 20 is not accompanied by the copula (47). This also characterizes complex numerals built on 20.

(96) wan pa e-tu-hot-ne pa
ten year e-two-score-COP year
'thirty years' (Tamura 1988/2000, 255, glosses added) object counting

(97) wan-pe e-tu-hot
ten-NMN e-two-score
'thirty' (Tamura 1988/2000, 255, glosses added) serial counting

Similarly to fractions, the serial counting forms of odd decimals thus do not involve a verb, and so there is no host for an applicative here.

I conclude from the foregoing discussion that due to Ainu-internal morphosyntactic reasons, the *e-* of odd decimals cannot reasonably be analyzed as an applicative morpheme.[26] In the next section I will argue in favour of the 'toward' analysis of *e-* and support the view that odd decimals involve the operation of overcounting.

5.2. Overcounting in Ainu odd decimals

5.2.1. The operation of overcounting. We have seen that odd decimals are built on the template '*wan-pe* (= 'ten'-NMN) + *e-* morpheme + next-higher multiple of 20', thus 30, for instance, is '10 *e* 40'.

(98) wan pa e tu hot ne pa
ten year e two score COP year
'thirty years'

The fact that odd decimals anticipate the next-higher multiple of the base (e.g., in the case of 30, 40 is referenced) led most work on Ainu to the conclusion that these numerals are subtractive. However, cross-linguistically there exists another operation on complex numerals, too, which involves the next higher multiple of the base. So-called overcounting numerals take the next lower multiple of the base as their implicit reference. The complex number itself comprises i) a number naming the amount of steps we have taken on the counting sequence towards the next higher multiple of the base since our implicit starting point, and ii) the next *higher* multiple of the base. That is, (in decimal systems) "one [complex] number … is named by the corresponding unit and the following decade" (Marcos-Marín 1992, 759).[27] Consider the expressions for 21 in Northern Mansi: they comprise 'one' and the next higher multiple of the base, 'thirty', linked by a 'toward' P or a lative case suffix (lit. 'one toward thirty').

(99) wāt nupəl ak^{w}a
thirty toward one
'twenty one' (Riese 2001)

[26] *E-* applicatives can also have a locative ('on') reading, and Bugaeva (2010) tentatively suggests that odd decimals may involve this type of applicative (e.g., 'on ten – forty'). The morphosyntactic counterarguments presented above apply equally to this approach as well. The 'on'-applicative analysis does not present an improvement over the 'with'-applicative view on the semantic side either. 'On' – similarly to 'with' – is attested in complex numerals cross-linguistically, but always with an additive semantics.

[27] The next lower multiple of the base is usually not represented morphologically; it must be recovered from the context.

(100) wā̆t-n ak^{w}a
thirty-LAT one
'twenty one' (Riese 2001) Northern Mansi

While overcounting is cross-linguistically much rarer than additive, multiplicative or subtractive numerals, it can be found in various unrelated language families: it has been documented in some Germanic, Uralic, Turkic, Tibeto-Burman, Austronesian and Mayan languages (Menninger 1969; Hurford 2003; Marcos-Marín 1992; Hanke 2005; Mazaudon 2008; Coupe 2012; Ochiai 2014; He 2018; Dékány 2025a). Overcounting numerals may involve an overt connector between the parts of the complex numeral, as in Northern Mansi (with a 'toward' morpheme) or in earlier Chungli Ao (with a negated verb).

(101) metsɣ maben tenet
twenty not.brought seven
'seventeen' (Clark 1893) earlier Chungli Ao

In some languages the "corresponding unit and the following decade" are juxtaposed without an overt morphological connector. This is the case in Old Turkic (102), for instance. Based on the translation of this example given by Erdal (2004), it is clear that this is neither an additive nor a subtractive numeral ($4+40\neq34$, $40-4\neq34$). (102) expresses that since we have left 30, the next-lower multiple of the base, we have come four steps toward 40, the next higher multiple of the base.

(102) tört kïrk
four forty
'thirty(!) four' (Erdal 2004) Old Turkic

Greenberg (2000) points out that overcounting exists in time-telling expressions of European languages as well, citing German *halb vier* lit. 'half four', meaning 03:30/15:30 as an example. While this example is potentially ambiguous between a subtractive ('half from four') and an overcounting ('half towards four') reading, clear examples are furnished by the Hungarian examples below. Here subtraction is not an option: 'quarter from four' would be 03:45/15:45, but the meaning of lit. 'quarter four' is 03:15/15:15 'quarter towards four' (when counting from three; the starting point of the counting again remains implicit).

(103) a. negyed négy
quarter four
'03:15/15:15'

b. három-negyed négy
three-quarter four
'03:45/15:45' Hungarian

5.2.2. The advantages of the overcounting view for Ainu. A small minority of researchers considers Ainu odd decimals to involve overcounting. In his cross-linguistic survey, Menninger (1969) includes the relevant Ainu numerals in the section on overcounting, and Arapu (1977, 77)

calls the pattern of odd decimals "système vigésimal anticipatif". Both Menninger (1969) and Arapu (1977) rely on earlier Ainuist literature which treats odd decimals as subtractive and/or applicative. They do not engage with this literature, however: it is not acknowledged that their stance on odd decimals differs from that of all earlier literature, no arguments are levelled against subtraction (confusingly, Menninger 1969 even glosses *e-* as *von* 'from'), and no support is offered for the overcounting view. Menninger (1969) and Arapu (1977) do not try to establish a connection between the *e-* of complex numerals and one of the other uses of *e-* either.[28] In a recent paper Ochiai (2021) also supports the overcounting analysis. She engages with earlier literature and acknowledges how her view is different from the dominant stance, but she does not offer morpho-syntactic arguments against subtraction or for overcounting either.

I would like to argue that adopting the overcounting analysis for odd decimals has four immediate payoffs. Firstly, this analysis avoids positing applicativization of and noun incorporation into the copula, which falls in line with the fact that these operations do not seem to apply to this verb. Secondly, the analysis easily handles fractions with 'two' and 'three', which are numerals not construed with the copula.[29]

[28] Greenberg (2000) also relies on earlier work that views Ainu odd decimals as subtractive. He maintains the subtractive analysis for odd decimals, but claims – again without arguments – that fractions with *e-* involve overcounting. On p. 779. he writes: "Overcounting is sometimes found in fractions. An example is Ainu, in which *e*, expressing negation, is used for subtraction in the cardinal system. In similar fashion, $1\frac{1}{2}$ is expressed as *e-tup* 'not-two', $2\frac{1}{2}$ by *e-rap* 'not-three', etc." As pointed out earlier, there is no evidence for *e-* being a negative marker, and it remains unclear why (or compositionally how) the same morpheme would signify subtraction with cardinals but overcounting with fractions.

[29] Fractions are formed by overcounting in some Indo-European and Finno-Ugric languages as well (data from Menninger 1969, 78).

(i) dritte-halb, vierte-halb
third-half fourth-half
'two and a half, three and a half' German

(ii) derde-half
third-half
'two and a half' Dutch

(iii) pol-tora
half-second
'one and a half' Russian

(iv) semis-tertius
half-third
'two and a half' Latin

(v) puoli-kolmatta
half-third
'two and a half' Finnish

(vi) harmad-fél
third-half
'two and a half' Hungarian

What is common to these forms and the Ainu expressions in the main text is that a fraction is named by 'half' and the next *higher* integer. In their morphological make-up, however, (i)–(vi) are significantly different from their Ainu counterparts. In (i)–(vi) the next higher integer is in its ordinal form, and the understood relation between 'half' and the ordinal is either possession ('half of the third') or identity ('the third is half'). The Ainu fractions are not marked in the same way as possessive expressions, and the integer is not in an ordinal form either. As will become clear later on, in my analysis the Ainu equivalent of these fractions has the form '[halfway] to(wards) three'.

(104) e-tu-p, e-re-p, e-ine-p
e-two-NMN e-three-NMN e-four-NMN
'one and a half, two and a half, three and a half'
(Batchelor 1905 [1888], 96, see also Simeon 1968, 46)

(105) hontom e-tu-p
middle e-two-NMN
'one and a half' (NINJAL 2015)

Unlike applicatives, overcounting expressions are not necessarily built around a verb (see the Old Turkic (102) and Northern Mansi (99)–(100) examples above), thus 'no verb' poses no problem. The same applies to the serial counting forms of odd decimals, which likewise have no copula (see (97)). Finally, the overcounting approach does not predict the existence of full PP paraphrases for odd decimals (with comitative *tura* or instrumental *ani*), therefore the lack of such paraphrases is not an explanandum.

Based on the foregoing arguments, I conclude that the overcounting analysis is a better fit for the data than the (subtractive-)applicative approach. Mathematically, the two analyses are equally adequate (for instance, with 10 and 40, both subtracting '10 from 40' and overcounting '10 towards 40' yield 30). However, the overcounting view does not lead to the morphosyntactic problems discussed above. I will therefore adopt it in the rest of the paper.

5.2.3. *E-* as a 'to' connector. Let us now turn to the question of what role the *e-* particle plays in odd decimals. As pointed out before, outside of the numeral system *e-* functions as i) the second person singular verbal prefix cross-referencing A, S and O arguments, ii) an allomorph of the unproductive lexical causative suffix, iii) an applicative morpheme, and iv) a directional 'to' prefix. I concur with Ochiai (2021) that the *e-* of odd decimals is none other than the directional 'to' prefix, thus the full gloss of an odd decimal like 30 is the following:[30]

(106) wan pa e tu hot ne pa
ten year to(ward) two score COP year
'thirty years', lit. 'ten years toward forty years'

Let us first illustrate the directional use of *e-* outside of complex numerals. According to Nakagawa (2022a, 505), the directional prefix *e-* 'to(ward)' and its opposite *o-* 'from' occur in the frames in (107), with a noun flanked by the prefix on one side and by a verb (the copula *ne* or the verb *un*, indicating stative location or origin[31]) on the other.

(107) a. *e*-N-*ne* 'to-N-COP'
b. *o*-N-*ne* 'from-N-COP'

[30] Arapu (1977) also provides the literal translation "dix vers deux vingtaines" 'ten towards two twenties' for 30, though he does not comment on this.

[31] The verb *un* is also the diachronic source of the allative P.

c. *e*-N-*un* ‘to-N-attach.to/belong.to/live.at’

d. *o*-N-*un* ‘from-N-attach.to/belong/live.at’ Nakagawa (2022a, 505, glosses added)

The nouns occurring in these frames are inherently locative nouns that come in “pairs indicating opposite directions along the same axis” (Nakagawa 2022a, 504) such as *kim* ‘(the direction of a) mountain’ ~ *pis* ‘(the direction of a) shore’, *rep* ‘(the direction of the) ocean’ ~ *ya* ‘(the direction of the) coast’, *cupka* ‘east’ ~ *cuppok* ‘west’, etc. Some examples are given below.[32]

(108) e-kim-un
to-mountain-belong.to
‘toward the direction of a mountain’

(109) o-kim-un
from-mountain-belong.to
‘(come) from the direction of a mountain’

(110) e-rep-un
to-ocean-belong.to
‘toward the direction of the ocean’

(111) o-rep-un
from-ocean-belong.to
‘(come) from the direction of the ocean’

[32]While Ochiai (2021) characterizes the directional *e-* outside of numerals as rare, its existence is not in question: it has been documented in various expressions in Batchelor (1905 [1888], 140), Simeon (1968, 56), Tamura (1988/2000, 223–224), Sato (2002), Nakagawa (2022a) and the corpus of Nakagawa et al. (2016–2024). Refsing (1986, 158) described these particles in the Shizunai dialect as well; in this variety directional interrogative pronouns (and the indefinite pronouns built on them) are also formed with these prefixes (Refsing 1986, 102–103; 107–108):

(i) neyta
‘where?’

(ii) enun
‘to where?’

(iii) onun
‘from where?’

(iv) neyta-ka
‘somewhere’

(v) enun-ka
‘to somewhere’

(vi) onun-ka
‘from somewhere’ Shizunai Ainu

(112) e-cupka-un, e-cuppok-un
to-east-belong.to to-west-belong.to
'toward the east, toward the west' (Nakagawa 2022a, 506–507, glosses added)

Such prefixed forms typically modify verb phrases, and so are characterized by Ainu grammars as adverbs.

(113) E-kim-ne paye.
to-mountain-COP go.PL
'They went into the mountains.' (Tamura 1988/2000, 93, glosses adapted)

(114) O-kim-ne san.
from-mountain-COP descend
'He/she/it came down from the mountains.' (Tamura 1988/2000, 93, glosses adapted)

In some cases, however, they are inflected like verbs themselves.[33]

(115) […] ekimne-an yakka
go.to.the.mountain-4.S though
'[Two people, my elder brother and me] went to the mountain.' (Nakagawa 2022a, 508)

It should be pointed out that the frames in (107) mesh well with the idea that odd decimals contain the directional prefix *e-*. Firstly, numerals have a close semantic connection to locative nouns: they name positions or locations in the counting sequence. Furthermore, a shown earlier in (105), fractions can be formed with *hontom*, which is a genuine locative noun meaning 'middle, in the midst, halfway'. Finally, odd decimals involve the copula *ne*, which is one of the verbs that the directional prefix *e-* collocates with. Odd decimals thus arguably instantiate the frame in (107a).[34] This view provides a pleasing direct parallel with Northern Mansi and with the Tibeto-Burman language Ntenyi (aka Northern Rengma), which likewise use a directional morpheme as a linker in overcounting expressions (see Dékány 2025a for further examples):

(116) wāt nupəl ak^{w}a
thirty toward one
'twenty one' (Riese 2001)

[33] Alonso de la Fuente (2022, 175) notes that *ne* is also used to derive dynamic verbs, cf. the inherently locative noun *soy* 'outside' and *soy-ne* 'to go outside'. (115) may contain this use of *ne*.

[34] The fact that fractions and the serial counting form of odd decimals do not involve a verb does not constitute a problem. I will argue below that *e-* is a functional adposition. Its occurrence thus does not depend on the presence of verbs.

(117) wāt-n ak^{w}a
thirty-LAT one
'twenty one' (Riese 2001) Northern Mansi

(118) chaa she tüo
thirty towards six
'twenty six' (Mills 1937, 291, glosses mine) Ntenyi

Let us now turn to the *e-* appearing in complex numerals. With *e-* identified as a directional particle, odd decimals involve a (Path) PP: 'ten towards forty/two scores'. In Sect. 3 I suggested that 'twenty' involves a relative clause structure headed by the copula:

(119) [$_{RC}$ _$_{i}$ [hot]-ne] [$_{NP}$ pa$_{i}$]
score-COP year
'twenty years' (Tamura 1988/2000, 255, glosses mine)

As discussed before, even decimals involve the same structure, with an additional multiplier of 'twenty':

(120) [$_{RC}$ _$_{i}$ [tu hot]-ne] [$_{NP}$ pa$_{i}$]
two score-COP year
'forty years' (Tamura 1988/2000, 255, glosses mine)

PPs are relational phrases that relate a Figure and a Ground. In Talmy's (1978) words, the Figure is an object "whose path or site is conceived as a variable the particular value of which is the salient issue", while the Ground is "a reference-point, having a stationary setting within a reference-frame, with respect to which the FIGURE's path or site receives characterization" (p. 419). In the PP at hand, 'two scores/forty' functions as the reference-point or Ground. Since in PPs adpositions take the Ground as their complement, I suggest that the directional particle *e-* froms a constituent with the extended NP in (120).

(121) [$_{PP}$ e- [$_{xNP}$ tu hot ne pa]]
toward two score COP year
'toward forty years'

I further suggest that the PP in (121) modifies the NP *wan pa* 'ten years' (or the NP with the bound noun, *wan-pe* 'ten-NMN') as an adjunct. The overall meaning is 'ten years/ten ones [that are] towards two score years'.

(122) [$_{xNP}$ [$_{xNP}$ wan pa] [$_{PP}$ e tu hot ne pa]]
ten year to(ward) two score COP year
'ten years toward forty years'

(123) [$_{xNP}$ [$_{xNP}$ wan-pe] [$_{PP}$ e tu hot ne pa]]
ten-NMN to(ward) two score COP year
'ten ones toward forty years'

This PP adjunction structure is similar to (62), repeated here as (124a).

(124) Ionin & Matushansky's (2018) proposal for additives connected by a P
a. [$_{xNP}$ [$_{xNP}$ twenty books] [$_{PP}$ P=with [$_{xNP}$ three books]]]
b. [$_{xNP}$ [$_{xNP}$ three books] [$_{PP}$ P=on [$_{xNP}$ twenty books]]]

That in overcounting numerals the (NP containing the) smaller numeral is the head of the construction and the goal PP containing the next-higher multiple of the base is a modifier receives interesting confirmation from Lhōtā Nāgā. This language exhibits a pattern that, to my knowledge, is unique: the overcounting pattern explicitly includes the next-lower multiple of the base as well as an overt connector between this base and the rest of the numeral. In (125), for instance, 21 is literally 'twenty and one toward thirty'.

(125) mekwü sü thamdro-we ekhā
twenty and thirty-toward one
'twenty one' (Witter 1888, 27) Lhōtā Nāgā

The operation in this numeral is addition, with two basic components: the augend 20 and the addend 1. The phrase 'toward thirty' is a modifier of the addend, which, in this particular construction, appears to provide additional but not crucial information about 1. Most overcounting languages simply drop the augend and the connector and feature only what in (125) corresponds to the addend: 'one toward thirty' for 21. With the next-lower multiple of the base not being overt, the PP modifier provides essential information in order to identify which numeral is being referred to; this is the only device that distinguishes between 21 ('one toward thirty'), 31 ('one toward forty'), etc.[35]

With the structure of odd decimals in place, let us now fine-tune the position of *e-* within the PP. Southern Hokkaido Ainu is an exclusively postpositional language. The inventory of lexical postpositions comprises locative *ta*, allative *un* (inanimate Goal), dative *eun* (animate Goal),

[35]It is an interesting question whether all overcounting numerals should be thought of as covert additives, with the addend having a PP modifier. That is, one may think of the syntax of overcounting numerals in one of two ways. It may be the case that the next-lower multiple of the base is not present in the syntax, and e.g., 21, 31, etc. are all all different varieties of 'one': '[the] one [that is] toward thirty', '[the] one [that is] toward forty', etc. In this case the PP modifier is restrictive. This simpler syntax is what I pursued in the main text. Alternatively, overcounting could always involve addition in its underlying syntax, and be a subtype thereof, as in (125), with the augend and the connector remaining unpronounced on the surface. In this case the PP modifier is non-restrictive. As this possibility involves several structural components which do not receive a surface-realization, it would have to be carefully argued for.

ablative *wa*, instrumental *ani*, comitative *tura/tura-no*, traversal *peka* ('over') and mutative *ne* ('as'), see Bugaeva (2012).[36]

(126) kim **ta** paye=an orano,
mountains at/to go.PL=4.S then
'We went to the mountains' (Nakagawa et al. 2016–2024, K8303243UP)

The morphemes *e-* and *o-* thus cannot be lexical adpositions: they must be higher, functional adpositions, much like the particles of Germanic languages. That the directional *e-* is higher than lexical postpositions is also confirmed by the form of the dative postposition *eun*: it contains the allative *un* used for inanimate Goals plus the directional *e-* (Bugaeva 2011b, 237). Functional adpositions are known to combine with lexical adpositions in other languages as well, cf. Tortora (2005) on Spanish.

The directional particles *e-* and *o-* are diachronically related to two lexical prefixes, *e-* 'head/top' and *o-* 'buttocks/tail' (on these, see Tamura 1988/2000, 196 and Bugaeva 2010, 762), which, in turn, come from the relational nouns *he* 'head/top' and *ho* 'buttocks/bottom' (Bugaeva 2010, 762). These are illustrated below.[37]

(127) a. he-puni
head-lift.up
'to raise (one's) head' (Tamura 1988/2000, 197)

b. e-toy-poso
head-soil-pierce.through
'(a sprout) breaks ground' (Tamura 1988/2000, 196)

(128) a. ho-puni
bottom-lift.up
'to get up, to fly away' (Tamura 1988/2000, 197)

b. o-nis-poso
bottom-cloud-pierce.through
'(a meteor or lightning) falls' (Tamura 1988/2000, 196)

[36]Of these, the formative that has been identified as mutative *ne* 'as/(turn) into' is likely not a real postposition but simply the copula (in its 'become' meaning) which is used to expone the head of a Small Clause:

(i) ku-poro yakun isa **ne** k-an kusu-ne
1.(S)-be.big if.so doctor ne 1.(S)-exist intention-COP
'I'm going to become a doctor when I grow up.' (Tamura 1988/2000, 134, glosses adapted)

[37]Two of the three applicative morphemes, namely applicative *e-* and applicative *o-* have also developed from these relational nouns (Bugaeva 2010). Thus there is a distant diachronic connection between the *e-* of numerals and applicatives: these have a common ancestor but cannot be conflated synchronically.

Relational nouns and PPs are tightly linked in Ainu: apart from place-denoting proper names and a handful of intrinsically place-denoting nouns (such as 'mountain', 'village', etc., a group that is reminiscent of Matushansky's 2019 so-called L-nouns in Latin), the Ground and the relevant postposition are always linked by a relational noun (Bugaeva 2012; Nakagawa 2022a; Dékány 2025b). This is exemplified below.

(129) a. nisat-ta tunas-no tan **kotan** **wa** k=arpa kusu ne na
dawn-at be.early-ADV this village ABL 1SG.S=go.SG intention COP FIN
'I will leave this village early tomorrow.' (Bugaeva 2012) intrinsically locative N

b. **kuca** **or** **ta** arpa=an
hunting.cabin place LOC go.SG=INDF.S
'I went to the hunting cabin.' (Bugaeva 2012) intrinsically non-locative N

The 'relational noun-to-directional prefix/functional P' transition at hand thus instantiates a cross-linguistically well-documented grammaticalization path: semantic bleaching coupled with 'moving up the tree'.

In their discussion of the functional structure of PPs, Den Dikken (2010) and Den Dikken & Dékány (2025) draw a parallel between the internal functional structure of extended PPs and VPs (as well as NPs), and argue that the following functional layers are present in all types of extended projections. i) A lower aspect projection encodes grammatical aspect in the clause and the Place/Path distinction in the PP. ii) A higher aspect projection is the locus of viewpoint aspect in the clause and the bounded/unbounded contrast in the PP. iii) A projection for deixis and perspective hosts tense morphology in the clause and perspective or orientation markers in the PP, such as German *hin-* or *her-* in complex PPs (cf. *auf das Dach hinauf* lit. 'on the roof HIN.on', 'onto the roof' or *wo gehst du hin?* 'where are you going?' vs. *aus dem Haus heraus* lit. 'out.of the house HER.out.of', 'out of the house' or *wo gehst du hin?* 'where are you going?'). iv) A layer that houses discourse particles. v) The maximal extension of the phrase, which constitutes a locality domain and provides an escape hatch for extraction (CP in clauses and an analogous projection in PPs, sometimes called $CP^{Path/Place}$).

Nakagawa et al. (2016–2024) gloss directional *e-* and *o-* as 'head.POSS.PF' and 'bottom.POSS.PF', respectively, indicating that much of the original meaning is retained in these formatives. I think that this is right, and directional *e-* is probably best rendered as 'heading towards, (with the) head towards, facing' (hence, 'to'), while *o-* is best captured as 'bottom towards, (with the) buttocks toward, buttocks facing' (thus 'from'). Based on this, I tentatively propose that directional *e-* and *o-* are located in the functional layers of the PP that encode perspective or orientation. The way they differ from German *hin-* and *her-* mentioned above is that the German particles encode the speaker's perspective, while the Ainu particles are linked to the perspective of the Figure: the Figure is envisioned as an object with a front and a back,

and *e-* and *o-* encode the Firgure's orientation towards the Ground (i.e., which part of the Figure is facing the Ground).[38]

To summarize the discussion, I suggested that Southern Hokkaido Ainu odd decimals instantiate the overcounting operation with *e-* functioning as a 'to' connector. In the next section we will look at complex numerals built on odd decimals (31–39, 51–59, 71–79, etc.) and discuss their internal structure.

5.3. Complex numerals built on odd decimals

Before I conclude this section, I would like to briefly address the structure of complex numerals built on odd decimals, that is, non-round complex numerals containing *e-*. We have seen that even decimals are formed by multiplication of the base, 20. We have also seen that Ainu has additive numerals formed with the *ikasma* 'exceed.INTRA, remain, left' connector. The relevant examples are repeated below.

[38]As alluded to in the main text, *un* does double duty as both the allative postposition and a verb meaning 'be at, belong to, live in, fit'. (This ambiguity is not accidental: most postpositions originate in a verb; Bugaeva 2017b.) Bugaeva (2011b, 252) characterizes the verb *un* as a monotransitive locative copula. She notes that synchronically it is "predominantly (if not solely) used in adnominal clauses", as in (i), though it is also retained in a number of place names.

(i) nea kotan pa un kur
that village upper.part live.at/belong.to person
'the man from above the village' (Nakagawa et al. 2016–2024, K8010281UP.212)

As for the combination of the directional particles and *un*, it is not entirely clear that this use of *un* must be a verb, as suggested in Nakagawa's (2022a) frames in (107). Nakagawa et al. (2016–2024) gloss the *un* in (ii) as a verb, in the a) example as 'fit' and in the b) example as 'belong.to/live.in'.

(ii) a. e-kim-un
to-mountain-un
'toward the mountain' (Nakagawa et al. 2016–2024, K7908032UP.366, glosses adapted)

b. o-kim-un
from-mountain-un
'from the mountains' (Nakagawa et al. 2016–2024, OI990914UP.014, glosses adapted)

On the other hand, Tamura glosses the *un* of (iii) as 'towards', thus clearly sees the allative postposition in these forms.

(iii) a. e-koypok-un
to-west-un
'towards the west' (Tamura 1988/2000, 223, glosses adapted)

b. o-koypok-un
from-west-un
'from the west' (Tamura 1988/2000, 224, glosses adapted)

It is difficult to ascertain whether such examples feature the verb or the postposition, and it is also possible that the relevant examples are ambiguous. To the extent that the *un* that combines with the directional particles can be the allative adposition proper, it provides support for the proposed semantics of *e-* and *o-*. If *o-* were simply 'from', then its combination with the allative postposition would be an anomaly. However, if it is indeed 'bottom towards, (with the) buttocks toward', as suggested in the main text, it easly fits with the 'to' semantics of the allative.

(130) a. sine suma ikasma wan Suma
one stone exceed ten Stone
'eleven stones' (Tamura 1988/2000, 257)

b. sine-p ikasma wan Sike
one-NMN exceed ten bag
'eleven bags' (Tamura 1988/2000, 257)

This strategy applies to 11–19, 21–29, 41–49, 61–69, etc. The remaining complex numerals, namely 31–39, 51–59, 71–79, etc., are built on the following template:

(131) single-digit numeral < additive marker *ikasma* < *wan e* < next higher multiple of the base

Compare the forms for 'thirty' and 'thirty six':[39]

(132) wan e tu hot ne
ten to two score COP
'thirty' (Batchelor 1905 [1888], 95, glosses added)

(133) iwan ikasma wan e tu hot ne
six exceed ten to two score COP
'thirty six' (Patrie 1978, 115 and Batchelor 1905 [1888], 95, glosses added)

Previous literature treated examples like (133) as involving a combination of multiplication (2×20), subtraction ((2×20)−10) and addition (((2×20)−10)+6), see e.g., Peng & Brainerd (1970), Patrie (1978, 115) and Tamura (1988/2000, 255). In other words, the constituency in (134) has been assumed.

[39] Batchelor (1905 [1888]) and Patrie (1978) present abstract forms of numerals, without the bound noun *-pe* after the non-final numerals. This is presumably because instead of *-pe*, it is also possible to employ *-iw*/*-n* with human nouns as well as a full double of the head noun. Strictly speaking, the formative *-pe* is not part of the numeral; it is only one possible exponent of a position that is part of the numeral. Compare (132) with the examples below (according to Batchelor, *niu* is a longer form of *-n* for people):

(i) wan to e tu hot ne to
ten day toward two score COP day
'thirty days' (Batchelor 1905 [1888], 101, glosses added)

(ii) wa niu e tu hot ne niu
ten person toward two score COP person
'thirty persons' (Batchelor 1905 [1888], 99, glosses added)

(iii) shine-p iashima wan-be e tu hot ne-p
one-NMN exceed ten-NMN toward to score COP-NMN
'thirty one things' (Batchelor 1905 [1888], 100, glosses added)

This detail need not concern us here.

(134) six exceed [ten *e*=minus forty] (lit. six exceed thirty)

With the overcounting analysis of odd decimals, we could adapt this as (135).

(135) six exceed [ten *e*=toward forty] (lit. six exceed thirty)

With *e-* treated as a 'to' expression, however, another perspective opens up regarding the constituency of numerals between odd decimals and even decimals. As discussed in Ochiai (2021), it is possible to treat such numerals as involving overcounting with a 'teen'-numeral:

(136) [six exceed ten] *e*=to forty (lit. [counting from twenty,] sixteen toward forty)

Compare (133) with the numeral 16:

(137) iwan ikasma wan
six exceed ten
'sixteen' (Ochiai 2023, 101, glosses added)

That 'teen'-numerals can be involved in overcounting is shown by the Mayan language Ch'ol. Ch'ol is also a vigesimal language, which allows numerals above 20 to be expressed either with the additive or the overcounting method (Aulie 1957).[40]

(138) a. hun-k'al yik'ot hum-p'ehl
one-twenty with one-CL.general
'twenty one' (lit. 'twenty with one')

b. hum-p'ehl i čaʔ-k'al
one-CL.general ASERIES.3 two-twenty
'twenty one' (Aulie 1957, 282, glosses mine) Ch'ol

In the overcounting form in (138b), 'forty' is referenced: 21 is 'one in the group of two-20s'. In direct parallel to the suggested parse in (136), 'thirty six' relates 'sixteen' to 'forty'.

(139) a. wɬk-luhúm-p'ehl
six-ten-CL.general
'sixteen' (Merrifield 1966, glosses added)

b. [wɬk luhúm-p'ehl] i-čáʔ k'ál
six ten-CL.general ASERIES.3-two twenty
'thirty six' (Merrifield 1966, glosses added) Ch'ol

[40]Examples are presented as in the original, which is why the transliterations and hyphenations in (138) and (139) are slightly different.

(136) would mean that numerals between even decimals and odd decimals (21–29, 41–49, etc.) employ addition (to 20 or its multiple), while numerals between odd decimals and even decimals (31–39, 51–59, etc.) employ overcounting with a(n internally additive) 'teen'-numeral. In other words, halfway between two multiples of the base, there is a change in the strategy of composing complex numerals. While this may look like a complication of the system, it only uses components that we already have in place: additives (necessary for 11–19, 21–29, 41–49, etc.) and overcounting (necessary for 30, 50, 70, etc.). As discussed in Ochiai (2021), there is also a logic to such a system: interval numerals refer to the *closest* multiple of the base. In the interval between 20 and 40, for instance, the closest multiple of the base for 21–29 is 20, but from 31 on the closest multiple of the base is 40.

The plausibility of this scenario is supported by the old counting system of Ao (Tibeto-Burman), which employs a switch such as this. Ao uses a decimal numeral system. Until the early 20th century, the Mongsen, Chungli, Changki, and Longla dialects referenced the next lower multiple of ten if the complex numeral ended in 1–5 (140). These were additive numerals.

(140) ter, teri ka, teri ana, teri asym, teri pezy, teri pungu
ten ten one ten two ten three ten four ten five
'ten, eleven, twelve, thirteen, fourteen, fifteen' (Clark 1893, 44) earlier Chungli Ao

Complex numerals ending in 6–9, however, referenced the next higher multiple of ten, with overcounting, as in (141) (Avery 1886; Coupe 2007). These numerals used a negated verb as a connector.[41]

(141) metsy maben trok, metsy maben tenet, metsy maben ti, metsy maben tvko, metsy
twenty not.bring six twenty not.bring seven twenty not.bring eight twenty not.bring nine twenty
'sixteen, seventeen, eighteen, nineteen, twenty' (Clark 1893, 45) earlier Chungli Ao

The examples above illustrate the 'teen'-series, but the split applies to the higher series as well. Although due to the efforts of Christian missionaries Ao has changed to a fully additive system around the early 20th century, its original way of counting shows that proximity to a round numeral may influence whether a particular number is expressed by addition or overcounting. A system like 19th century Ao's is also found in other Tibeto-Burman languages of Nagaland, such as Rengma, Ntenyi, Pochury and Meluri (Mills 1937; Matisoff 1997).[42]

That increasing closeness to the next higher numeral can boost overcouting is also shown by Sumi (aka Sema, Tibeto-Burman) and Central Mansi (Uralic). In Sumi, complex numerals ending in 9 always employ overcounting, while numerals ending in 7 and 8 (and in the Tizu

[41]The Ao glosses have been added based on the discussion and glossaries in Clark's (1893) grammar. Clark paraphrases 16 as 'twenty not brought six, the six before twenty'. Coupe (2007) uses the gloss 'NEG-complete'. According to his informants, 16 is literally "(the) twenty not completed, (the) six".

[42]In Eastern Rengma, also known as Anyo, there was a split among the 'teens': only 16–19 employed overcounting, while 11–15 were additive (Coupe 2007, 120). Between 20 and 100, there was a fully overcouting system; see also Mills (1937, 293).

Valley dialect, also 6) have both additive and overcounting variants (Hutton 1916, 4). Among the teens of Central Mansi, 11–17 are additive, while 18 and 19 also have overcounting variants (Honti 1993, 298). Further examples of this type are discussed in Dékány (2025a).

I currently do not see morpho-syntactic arguments that could adjudicate for Ainu between the old parse in (134)/(135) and the new parse in (136). It is possible that the matter could be settled on the basis of prosody, but I am not in a position to conduct the relevant experiments, and as Ainu is a moribund language, it may not be feasible for anyone to check this with speakers who have good competence. The reason why it is important to discuss the issue of parsing is that if persuasive arguments for (136) would surface in the future, it would provide additional evidence for the overcounting analysis of odd decimals. This is because the parse in (136), repeated here as (142), is possible only on the overcounting view.

(142) [six exceed ten] *e*=to forty (lit. [counting from twenty,] sixteen toward forty)

While 'sixteen towards forty' is indeed 36, if *e-* signified either subtraction or a 'with'-applicative, we would wrongly predict this numeral to mean 24 (40−16, or 'with 16, [it] becomes forty'). The parsing generally adopted in the literature, on the other hand, is compatible with both the overcounting and the subtractive-applicative view: (134) and (135).

This concludes my discussion of odd decimals. To sum up, I suggested that odd decimals involve modification of 10 by a path PP, where *e* is a functional Path adposition. I raised the possibility that overcounting applies not only to odd decimals but also to 31–39, 51–59, etc., but a more qualified assessment of these numerals will have to remain for future work. In the next section I will turn to the last set of numerals that we have not covered so far: those between 6 and 9.

6. TRUE SUBTRACTIVE NUMERALS

Let us now turn to the numerals 6 through 9. Morphologically, 6–7 and 8–9 constitute separate groups.[43] There is consensus in the literature that 6 and 7 are formed by subtraction from 10 (Pott 1847; Laufer 1917; Arapu 1977; Patrie 1978; Tamura 1988/2000), while the structure of and the operation involved in the formation of 8 and 9 is debated.

Starting with 6 and 7, these numerals are formed by juxtaposition of the subtrahend (3 or 4) and the minuend (10).

(143) a. i-ne
four-COP
'four'

b. i-wan
four-ten
'six'

[43]A similar split is also attested in some Altaic languages, see Hamp (1968), Patrie (1982).

(144) a. re
‘three’

b. ar-wan
three-ten
‘seven’

Notice that 4 loses its *ne* ending here, confirming that *ne* is a separate morpheme from the root *i-*; the latter expresses the quantity by itself.[44] 7 contains a special allomorph of 3. Why this should be so is not at all understood, but there is broad agreement in the literature that at the core, (144a) and (144b) involve the same morpheme.[45]

The morphological composition of 8 and 9 is more complex, and while everyone agrees that these are also subtractive numerals, there is more than a little controversy over their exact morphological decomposition. Below they appear in an unsegmented form.

(145) tu
‘two’

(146) sine
‘one’

(147) tupes
‘eight’
(serial counting)

(148) sinepes
‘nine’
(serial counting)

(149) tupesan N
‘eight Ns’
(object counting)

(150) sinepesan N
‘nine Ns’
(object counting)

[44]In addition to corresponding to 6, *iwan* is also used to mean ‘a bunch, a lot’ (Tamura 1988/2000, 260); in this capacity, it can refer to a number smaller or bigger than 6. Tamura relates this to the fact that it expresses a number that is greater by one than the number of one hand’s fingers.

[45]The exception is Ochiai (2021), who suggests that this is originally *ar-i-wan* lit. ‘other-four-ten’, that is, ‘the other six’, and the present form is a result of contraction. A reviewer points out that the copula *ne* and the existential verb *an* both contain the *n* segment and so these two verbs may be bimorphemic, both containing the same morpheme. The link between the two allomorphs of 3, namely *re* and *ar*, could then be the same. Such a link between the copula and the existential verb is plausible. Pursuing this line of thinking would take me too far afield, however, therefore I leave a more thorough investigation of the *ne* ~ *an* (and *re* ~ *ar*) connection to future work.

These numerals have received altogether four different analyses in the literature, which have not been systematically compared before. Many works simply do not engage with the competing proposals. Those that do typically mention only one of the alternative analyses (no work discusses all existing views). In general, the proposals are put forth without support from morphosyntactic arguments, and counterarguments against alternative views, rooted in the morphosyntax of the language, are also missing. The treatment of 8 and 9 in Refsing (1986, 111) is illustrative here: she mentions Analysis 1 and Analysis 4, and simply calls the latter a "more likely interpretation". Below I will lay out all four views on 8 and 9, and provide arguments for or against these views in light of how they fit into my analysis of odd decimals and the numerals in the counting sequence.

Analysis 1, suggested in Laufer (1917), assumes three morphemes in 8 and 9: i) the subtrahend, ii) the *-pe* nominalizer, and iii) *san*, which is taken to be an allomorph of *wan* 'ten'.

(151) tu-pe-san
two-NMN-ten
'eight'

(152) sine-pe-san
one-NMN-ten
'nine' Analysis 1

It must be emphasized that *w*~*s* is an otherwise unattested alternation in the (morpho)phonology of Ainu (Laufer 1917), thus from this perspective, the forms of 8 and 9 are somewhat unexpected.

Analysis 2, pursued in Pfizmaier (1851) and Batchelor (1905 [1888]), assumes the same decomposition as Analysis 1, but it identifies *san* as the verb 'descend'.

(153) tu-pe-san
two-NMN-descend
'eight'
lit. 'two comes down'

(154) sine-pe-san
one-NMN-descend
'nine'
lit. 'one comes down' Analysis 2

The use of *san* 'descend' is illustrated om (155). Nakagawa et al. (2016–2024) further decompose *san* into two morphemes: 'front place' and a (singular) intransitivizing marker. As shown in (155), *san* takes the subject agreement characterisitc of intransitive verbs.

(155) sa-n=an ayne
front.place-INTRA.SG=4.S finally
'As I traveled down' (Nakagawa et al. 2016–2024, K7807151UP.060)

Analyses 1 and 2 are similar in several respects. Firstly, they both assume a trimorphemic analysis of 8 and 9. Secondly, neither approach bears on the analysis of *e-* as an applicative or a 'toward' P, as they do not postulate this *e-* in 8 and 9. Finally, they suffer from the same problems. The first problem is that the *-p* allomorph of the bound noun appears after vowels, while *-pe* surfaces after consonants. As *sine* and *tu* end in vowels, we would not expect them to be followed by the *-pe* from (cf. their serial counting forms: *sinep* and *tup*, (41)). The second problem is that Analyses 1 and 2 cannot give a principled account of why *an* disappears in the serial counting form, as the *an* string does not correspond to a morpheme in these analyses. I consider this to be a strong reason to reject these approaches.

Analysis 3, pursued in slightly different forms in Tamura (1988/2000) and Ochiai (2021), takes 8 and 9 to comprise four morphemes: i) the subtrahend, ii) the *-p* allomorph of the bound noun/nominalizer *-p(e)*, iii) the same *e-* that appears in odd decimals, and iv) *san*, which is taken to carry the meaning of 'ten' (normally: *wan*). For Tamura, the *e-* in odd decimals has the meaning 'with' (that is, it is an applicative morpheme), thus the *e-* of 8 and 9 also means 'with'. (This decomposition, along with the identification of *e-* as an applicative, is also adopted in Bugaeva 2012.)

(156) a. tu-p-e-san
two-NMN-APPL-ten
'eight'
lit. 'two-thing-with(by adding)-ten'

b. sine-p-e-san
one-NMN-APPL-ten
'nine'
lit. 'one-thing-with(by adding)-ten' Analysis 3, as per Tamura (1988/2000)

I argued that the odd decimals feature a directional *e-*. If 8 and 9 indeed feature the same *e-* morpheme as odd decimals, (156) cannot be correct: no applicative morpheme is involved. In principle, it would be possible to maintain that the odd decimals contain the directional *e-*, while 8 and 9 feature an applicative *e-*. Thus, for instance, 30 is 'ten towards 40', but 8 is 'with/by (another) two, ten'. In this case 8 and 9 would be the only numerals built with the applicative *e-*. This possibility, however, can be ruled out on the grounds that the structures in (156) do not contain a verb, hence, they cannot feature an applicative morpheme either.

For Ochiai (2021), the *e-* of odd decimals has a 'toward' meaning. As for the *e-* in 8 and 9, she considers it "to be the same word, in terms of its meaning", but at the same time she states that in 8 and 9 *e-* "functions as representing subtraction". That is, *e-* is "multifunctional as it is used for both overcounting and subtraction".[46]

[46]She also mentions that *san* could be derived from *sam* 'none', but adds that there is no "definitive argument" either way.

(157) a. tu-p-e-san
two-NMN-SUBTRACT-ten
'eight'

b. sine-p-e-san
one-NMN-SUBTRACT-ten
'nine' Analysis 3, as per Ochiai (2021)

A single listeme of the mental lexicon cannot plausibly encode two distinct meanings. Therefore the only way I can understand the referenced 'sameness' is in terms of surface exponence: we are dealing with homophones, such that one *e-* means 'towards' and the other is a subtractive prefix.

The reason why Ochiai (2021) posits a difference between the *e-* of odd decimals and the *e-* of 8 and 9 is the following. When expressing a numeral value, the *e-* of odd decimals clearly takes the next-lower round numeral as its reference point, that is, 30 is 'ten toward forty, *when counting up from 20*'. In the numeral sequence 1 through 10, 5 is the last and highest simplex numeral, so it may be expected that an overcounting form would take this as the reference point. But it is clear that counting up from 5, 8 is not 'two towards ten', and 9 is not 'one towards ten'. Thus the exact same 'toward' strategy that we see with odd decimals clearly cannot be at work here.

Two paths can be taken if we want to avoid positing homophones with opposite meanings, that is, a 'toward' *e-* as well as a 'from (= subtract)' *e-*. One possibility is to adopt a structure which does not posit the *e-* of odd numerals in 8 and 9. Analyses 1 and 2 were such approaches, but we saw that they face some problems. Analysis 4, to be outlined below, likewise does not posit an *e-* morpheme in 8 and 9. The other possible path is to posit the same directional 'to' *e-* in both cases, and find an explanation for why 8 and 9 are different, i.e., why they are not expressed as '[from 5], three toward ten', and '[from 5], four toward ten', respectively. I suggest that a plausible way to think about the difference between odd decimals on the one hand and 8 and 9 on the other is the following. As described before, odd decimals take the next-lower round numeral, the next-lower multiple of the base as the starting point of counting: 30 is 'ten toward forty, *when counting up from 20*'. In the case of 8 and 9, however, there is no next-lower round numeral, or next-lower multiple of the base. So instead of taking 5, the next-lower simplex numeral as their reference point, these numerals take the number value to be expressed, or the present point, as their reference point. In other words, 8 is 'two [more] toward ten, *when counting from here*', and 9 is 'one [more] toward ten, *when counting from here*'. Viewed this way, 8 and 9 are not subtractive. They involve a 'toward' expression rather than the 'from' or 'without' that characterizes subtractives cross-linguistically; what is special about them is the reference point at which the counting starts.

Earlier on, I illustrated overcounting with time expressions such as German *halb vier* lit. 'half four', meaning 03:30/15:30 or Hungarian *három-negyed négy* lit. 'three-quarter four', 03:45/15:45. There is a time-expression equivalent of 'two [more] toward ten, *when counting from here*' as well, namely English 'five to 10', 09:55/21:55. It is clear that in 'five to 10', we have the same 'to' as in 'to the bridge' or 'five [more] kms to the bridge'. 'To' has not turned into a subtractive expression or a Source ('from') preposition here: it remains a Path preposition. What is special is that in contrast to Hungarian *három-negyed négy* 'three-quarter four, 03:45/15:45', where the reference point is the next-lower full hour, here the Ground or reference point is the 'here and now'.

While it is possible to unify the quadrimorphemic decomposition of 8 and 9 with the over-counting analysis of odd decimals this way, the analysis shares the drawback of Analyses 1 and 2: it cannot explain why it is precisely the *an* segment of 8 and 9 that disappears in the serial counting form. I will therefore put this approach aside and argue in favour of a fourth analysis.

Analysis 4, proposed in Patrie (1982) and taken on board in Refsing (1986), assumes a different trimorphemic analysis of 8 and 9. It is suggested that these numerals contain the subtrahend, the *upes* allomorph of 10, and the existential verb *an*.

(158) tu-upes-an
two-ten-exist
'eight'

(159) sine-upes-an
one-ten-exist
'nine' Analysis 4

Since the word for 10 is *wan* in all other numerals, the success of this analysis turns on to what extent *upes* can also be tied to 10. The word *uupis* is contained in Krašeninnikov's word list from 1738 as the Kurile Ainu form for 10 (Alonso de la Fuente 2021) and according to Refsing (1986, 111), it was independently registered with this meaning by Georg Wilhelm Steller as well.[47] While Kurile Ainu and Hokkaido Ainu are different major dialect(group)s, there is a diachronic connection between them. Alonso de la Fuente (2021, 8) writes about Kurile Ainu that "it seems beyond reasonable doubt that the language, while distinctive (especially on lexical grounds), was an offshoot of some varieties of Hokkaidō Ainu".

While it cannot be explained why Southern Hokkaido Ainu would use a different allomorph for 10 precisely here, exceptions in complex numerals are not unheard of, and the morphological decomposition in (158) and (158) has much to recommend it. Firstly, it provides a parallel with the morphological composition of 6 and 7. In all cases, the minuend and the subtrahend are juxtaposed without an overt connector, and without the subtrahend being nominalized. Secondly, this way all subtractive numerals contain the existential verb *an* in some form. This analysis can also straightforwardly model the loss of the *an* string from 8 to 9 in serial counting. *An* is dropped because it corresponds to a morpheme, and its loss is in direct parallel to the loss of *ne* 'COP' in 5. In all cases, the verbal part of the numeral disappears because the quantity expression is normally construed with the head noun via a noun-modifying clause, and when the head noun is not projected, the quantity expression stands on its own, without being embedded in a clause. For these reasons, I support Analysis 4 for the numerals 8 and 9.

(160) [$_{\text{PredP}}$ [$_{\text{NumP}}$ tu upes] an] N
two ten exist N
'eight Ns' (i.e., two from ten Ns)

[47] In connection with this claim, Refsing (1986, 111) refers to Murayama (1971, 19, 66). Steller published a book (Steller 1774) and left a Kurile Ainu vocabulary in a manuscript form (Alonso de la Fuente 2021). His book does not contain information about the Ainu, thus I assume that the source of *uupis* 'ten' is the vocabulary, but I have not been able to confirm this.

To summarize, the Ainu numerals 6 through 9 are subtractive, and they all involve the existential verb in some form. 6 and 7 do so via *wan* 'ten', which itself contains this verb. 8 and 9 contain a special allomorph of 'ten' as the minuend, followed by the existential verb. The clause projected by the existential verb is a noun-modifying clause which is linked to the counted head noun via an understood 'such-that' relation, exactly as in the case of *wan* 'ten'.[48]

7. CONCLUSIONS

This paper examined the syntax of simplex and complex numerals in Southern Hokkaido Ainu. As for simplex numerals, I argued that that only 2 and 3 are merged in a direct modification structure to the noun. The other simplex numerals involve clausal modification of the noun, either in the form of a relative clause or in the form of a modifying clause with a 'such that' interpretation (Table 1).

I proposed that depending on the particular numeral value, abstract counting and object counting forms for 1 through 5 entertain two different structural relationships. Abstract counting numerals can be followed by a bound noun that effectively acts as a nominalizer, in which case they involve the same amount of structure as the corresponding object counting numerals. This is the case for 1, 2, 3 and 4. When abstract numerals are not followed by a bound noun, they involve a smaller structure than object counting numerals, as in the case of 5.

As far as complex numerals are concerned, I showed that additive numerals involve a structural relationship between two identical, quantified nominals; Southern Hokkaido Ainu does not avail itself of a structure in which the parts of the additive complex numeral form a constituent to the exclusion of the head noun. I also suggested that all of the subtractive numerals have a structure in which the subtrahend and the minuend are juxtaxposed without an overt connector or further overt material between them. The resulting complex numeral is in a fully saturated clause headed by the existential verb.

Table 1. Simplex numerals

direct modification		**2, 3**
indirect modification		
	relative clause with the copula	**1, 4, 5, 20**
	fully saturated clause with an existential verb	**10**

[48]The subtractive numerals can also function as multipliers of *hot* 'twenty', e.g., *iwan hot* 'six scores' (Ochiai 2021). As with lower multipliers, I assume that first the multiplier combines with the noun *hot* 'score', and the resulting xNP functions as one of the arguments of the copula *ne* in object counting, cf. (67). Subtractive numerals are rarely used in this construction, however. Ochiai (2021) points out that starting with 120, the majority of dialects switch to addition instead: *hot-ne-p ikasma asik-ne hot* lit. 'twenty (things) exceed five scores', i.e., 'twenty and one hundred'. In the Saru dialect, both *i-wan hot* and *hot-ne-p ikasma asik-ne hot* are possible (Ochiai 2021, 104). Tamura (1999, 58) further notes about the numeral system of Ainu that "[b]y using this system it was possible to count to two hundred, but because there was little reason to ever count to such a large number, Ainu elders in the 1950s did not understand what "two hundred" meant; it was not customary to count large numbers, including people's age."

Finally, I argued that odd decimals involve the operation of overcounting rather than subtraction. I suggested, in line with Ochiai (2021), that odd decimals in Ainu have a directional 'to' connector, and explored the internal composition of these numerals. I suspect that the reason why the subtractive analysis has been so influential for these numerals in previous literature is that both Japanese and Western scholars of Ainu were unaware of the existence of overcounting, but they were familiar with the Latin pattern in which numerals referencing the next higher multiple of the base are indeed subtractive (161).

(161) duo-de-viginti
two-from-twenty
'eighteen' Latin

Ainu would not be the first language in which an overcounting pattern has been misidentified or mislabelled as subtractive. For instance, in his discussion of the numerals in six Tibeto-Burman languages of Nagaland, Matisoff (1997, 33) calls the higher teens subtractive. An example is given from Khonoma Angami in (162).

(162) meku-pomo-thetha
twenty-fall.short-eight
'eighteen' (Matisoff 1997, 33, gloss based on McCabe 1887) Khonoma Angami

The claim that (162) involves subtraction is mathematically clearly incorrect ($20-8\neq18$), and Matisoff himself explicitly notes that the pattern is different from subtraction in Latin. Similarly, Egli (1990) calls Paiwan (Formosan, Austronesian) overcounting numerals subtractive, noting, however, that their composition is not quite like that of Latin subtractives. It is perhaps for lack of a better term that they both stick with the expression 'subtractive'.

Viitso (1998) also mislabelled the overcounting numerals of Estonian (Finno-Ugric) as subtractive. In Estonian, overcounting is used from 11 through 19, with 'twenty' expressed as the 'second (ten)', and no overt connector between the single-digit numeral and 'twenty'.

(163) kaheksa-teist(kümmend)
eight-second.PART(ten)
'eighteen' (Anne Tamm, p.c.) Estonian

In spite of the use of the label 'subtractive', a genuine subtraction operation is not at play between the parts of the complex numeral in this case either.[49]

[49]Amis (Formosan, Austronesian) may form 11 and 999 in a similar fashion.

(i) tʃətʃaj (ku) ʃakatuʃa
one PRT second
'eleven' (He et al. 2017) Amis

While the Amis forms have additive alternatives, Estonian 11–19 can only be formed with overcounting. The reason Amis is interesting here is that even though He et al. (2017, 300) recognize that 11 and 999 involve overcounting, they also claim that this is "a less commonly used method to form additive numerals". It thus seems to be the case that overcounting has been mistaken in the literature not only for subtraction, but also for addition.

Table 2. Complex numerals

subtractives		
	fully saturated clause w. an existential verb, *wan*-based	6, 7
	fully saturated clause w. an existential verb, *upes*-based	8, 9
additives	relate two quantified NPs	11-19, 21–29, 41–49, etc.
multiplicatives	vigesimal system	40, 60, 80, etc.
overcounting	involve a toward PP and an RC	30, 50, 70, 90 (31–39, 51–59, 71–79, etc.)

The analysis of complex numerals argued for in this paper is summarized in Table 2. The structure of multiplicative, additive and subtractive numerals has been examined in previous literature in great detail (see Hurford 1975, 2003; Greenberg 2000; Ionin & Matushansky 2018; Wągiel & Caha 2021, to mention a few works on the topic). Overcounting numerals have been treated only on the descriptive-observatory level (Marcos-Marín 1992; Greenberg 2000; Ochiai 2021); their syntactic composition has not been examined from a formal perspective before. Comparing the structure of overcounting numerals with a 'to' connector, to those with a verbal connector and without an overt connector will be a worthwhile project in the future.

In the picture that emerges from this paper, Southern Hokkaido Ainu uses all possible operations in the formation of complex numerals: multiplication, addition, subtraction as well as overcounting. In this regard, it shows parallels with Finnish (Finno-Ugric) and Lhōtā Nāgā (Tibeto-Burman), which likewise take advantage of all operations in their numeral system.

ACKNOWLEDGEMENTS

I thank Izumi Ochiai for sending me an English translation of Ochiai (2021) and Hideki Kishimoto for transliterating the titles of articles and books published in Japanese. I am also grateful to the anonymous reviewers as well as Irina Burukina and Marcel den Dikken for constructive comments. All errors are mine. This research was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, the ÚNKP-23-5 New National Excellence Program and grant NKFIH FK 145985.

ABBREVIATIONS

A	subject of transitive
ABL	ablative
ADV	adverbial
APASS	antipassive
APPL	applicative
ASERIES	prefix in the so-called series A of Mayan linguistics ('set A')
CL	classifier

CONTEMPCONJ	contempral conjunctionalizer
COP	copula
EXCL	exclusive
FIN	final particle
INDF	indefinite
INFREV	inferential evidential
INTRA	intransitivizer
LAT	lative
LOC	locative
M	masculine
NEG	negative
NI	incorporated noun
NMN	nominalizer
O	object
PART	partitive
PF	prefix
PL	plural
POSS	possessive
PRT	particle
RECP	reciprocal
REFL	reflexive
S	argument of intransitive verb
SG	singular
SUBTRACT	subtractive particle

REFERENCES

Adams, Karen L. and Nancy F. Conklin. 1973. Toward a theory of natural classification. In C. Corum, T. Cedric Smith-Stark and A. Weiser (eds.) Papers from the 9th Regional Meeting of the Chicago Linguistic Society. Chicago, IL: Chicago Linguistic Society. 1–10.

Alonso de la Fuente, José Andrés. 2021. Kuril Ainu zoonyms and phytonyms in Pallas's *Zoographia* and *Flora Rossica*. International Journal of Eurasian Linguistics 3. 7–49.

Alonso de la Fuente, José Andrés. 2022. The Ainu language through time. In A. Bugaeva (ed.) Handbook of the Ainu language. Berlin & Boston, MA: De Gruyter. 147–183.

Arapu, Bogdan Daniel. 1977. Quantite et predication de la quantite en Ainu. Cahiers de Linguistique Asie Orientale 2. 77–85.

Aulie, Wilbur. 1957. High-layered numerals in Chol (Mayan). International Journal of American Linguistics 23. 281–283.

Avery, John. 1886. The Ao Naga language of Southern Assam. The American Journal of Philology 7. 344–366.

Babby, Leonard. 1985. Noun phrase internal case agreement in Russian. Russian Linguistics 9. 1–15.

Bartos, Huba. 1999. Morfoszintaxis és interpretáció: A magyar inflexiós jelenségek szintaktikai háttere. [Morphosyntax and interpretation: The syntactic background of Hungarian inflexional phenomena]. Doctoral dissertation. Eötvös University, Budapest.

Bartos, Huba. 2000. Az inflexiós jelenségek szintaktikai háttere [The syntacic background of inflexional phenomena]. In F. Kiefer (ed.) Strukturális magyar nyelvtan 3: Morfológia [Hungarian structural grammar 3: Morphology]. Budapest: Akadémiai Kiadó. 653–761.

Batchelor, John. 1905 [1888]. Ainu-English-Japanese dictionary (including a grammar of the Ainu language), 2nd edn. Tokyo: Methodist Publishing House.

Bhaskararao, Peri and Sunder Krishna Joshi. 1985. A study of Newari classifiers. Bulletin of the Deccan College Post-Graduate and Research Institute 44. 17–31.

Bisang, Walter. 1999. Classifiers in East and Southeast Asian languages: Counting and beyond. In J. Gvozdanović (ed.) Numeral types and changes worldwide (Trends in Linguistics: Studies and Monographs 118). Berlin: Mouton de Gruyter. 113–186.

Bugaeva, Anna. 2004. A grammar and folklore texts of the Chitose dialect of Ainu: Idiolect of Ito Oda. Endangered Languages Project Publication Series A2-045. Suita: Osaka Gakuin University.

Bugaeva, Anna. 2010. Ainu applicatives in typological perspective. Studies in Language 34. 749–801.

Bugaeva, Anna. 2011a. A diachronic study of the impersonal passive in Ainu. In A. Malchukov and A. Siewierska (eds.) Impersonal constructions: A cross-linguistic perspective (Studies in Language Companion 124). Amsterdam & Philadelphia, PA: John Benjamins. 517–546.

Bugaeva, Anna. 2011b. Ditransitive constructions in Ainu. STUF – Language Typology and Universals 64. 237–255.

Bugaeva, Anna. 2012. Southern Hokkaido Ainu. In N. Tranter (ed.) The languages of Japan and Korea (Routledge Language Family Series). London: Routledge. 461–508.

Bugaeva, Anna. 2015. Valency classes in Ainu. In A. Malchukov and B. Comrie (eds.) Valency classes in the world's languages, Vol. 1: Introducing the framework, and case studies from Africa and Eurasia (Comparative Handbooks of Linguistics [CHL]). Berlin: De Gruyter Mouton. 807–854.

Bugaeva, Anna. 2016. On the innovative nature of Sakhalin Ainu: Focusing on nominalization. Studia Orientalia Electronica 117. 95–117.

Bugaeva, Anna. 2017a. Noun-modifying clause constructions in Ainu. In Y. Matsumoto, B. Comrie and P. Sells (eds.) Noun-modifying clause constructions in languages of Eurasia: Rethinking theoretical and geographical boundaries (Typological Studies in Language 116). Amsterdam & Philadelphia, PA: John Benjamins. 203–250.

Bugaeva, Anna. 2017b. Polysynthesis in Ainu. In M. Fortescue, M. Mithun and N. Evans (eds.) Handbook of polysynthesis. Oxford: Oxford University Press. 882–905.

Bugaeva, Anna. 2018. Ainu complex predicates with reference to Japanese. In P. Pardeshi and T. Kageyama (eds.) Handbook of Japanese contrastive linguistics. Berlin & Boston, MA: De Gruyter Mouton. 247–272.

Bugaeva, Anna. 2024. The Ainuic language family. In E. Vajda (ed.) The languages and linguistics of Northern Asia (The World of Linguistics 10.1). Berlin & Boston, MA: De Gruyter Mouton. 541–632.

Bugaeva, Anna, Johanna Nichols and Balthasar Bickel. 2022. Appositive possession in Ainu and around the Pacific. Linguistic Typology 26. 43–88.

Caha, Pavel. 2015. Czech numerals and no bundling. In U. Shlonsky (ed.) Beyond functional sequence. New York, NY: Oxford University Press. 173–196.

Chamberlain, Basil Hall and John Batchelor. 1887. The language, mythology, and geographical nomenclature of Japan viewed in the light of Aino studies: Including "An Ainu grammar" by John Batchelor and a catalogue of books relating to Yezo and the Ainos and a catalogue of books relating to Yezo and the Ainos (Memoirs of the Literature College). Tokyo: Imperial University.

Chiri, Mashiho. 1942/1973. Ainu gohō kenkyū: Karafuto hōgen o chūshin to shite. [Studes on Ainu grammar: With an emphasis on the Sakhalin dialect]. In M. Chiri. Chiri Mashiho chosakushū [Selected writings of Chiri Mashiho], Vol. 3. Tokyo: Heibonsha. 455–586.

Cinque, Guglielmo. 2010. The syntax of adjectives: A comparative study (Linguistic Inquiry Monograph 57). Cambridge, MA: MIT Press.

Clark, Mrs. E. W. (Mary M.). 1893. Ao Naga grammar with illustrative phrases and vocabulary. Shillong: Assam Secretariat Printing Office.

Coupe, Alexander R. 2007. A grammar of Mongsen Ao (Mouton Grammar Library 39). Berlin: Mouton de Gruyter.

Coupe, Alexander R. 2012. Overcounting numeral systems and their relevance to sub-grouping in the Tibeto-Burman languages of Nagaland. Language and Linguistics 13. 193–220.

Dal Corso, Elia. 2021. The language and folklore of West Sakhalin Ainu: A re-edition of Murasaki Kyōko's 'Karafuto Ainugo' with translation and grammatical notes. Munich: Lincom.

de Graaf, Tjeerd. 2016. The Ainu in Japan. Online publication at https://catedra-unesco.espais.iec.cat/en/2016/01/19/10-the-ainu-in-japan/. Càtedra UNESCO de Diversitat Lingüística i Cultural, Barcelona.

Dékány, Éva. 2015. The syntax of anaphoric possessives in Hungarian. Natural Language and Linguistic Theory 33. 1121–1168.

Dékány, Éva. 2021. The Hungarian nominal functional sequence. Cham: Springer.

Dékány, Éva. 2025a. Anatomy of a complex numeral: Overcounting, with special attention to Ch'ol. In V. Gribanova, L. Clemens and G. Scontras (eds.) Syntax in uncharted territories: Essays in honor of Masha Polinsky. Oakland, CA: UC eScholarship Repository. 125–150.

Dékány, Éva. 2025b. Strategies for PPs in Ainu. In S. Akolkar, A. Galvano and A. Ismael, K. F. Liu and L. Mikkelsen (eds.) Proceedings of WCCFL 42. Somerville, MA: Cascadilla Press. 104–111.

Dékány, Éva and Ekaterina Georgieva. 2020. Three ways of unifying participles and nominalizations: The case of Udmurt. In A. Alexiadou and H. Borer (eds.) Nominalizations: 50 years on from Chomsky's Remarks. Oxford: Oxford University Press. 169–202.

Derbyshire, Desmond C. 1985. Hixkaryana and linguistic typology. SIL International Publications in Linguistics 76. Dallas, TX: Summer Institute of Linguistics & the University of Texas at Arlington.

Dikken, Marcel den. 2010. On the functional structure of locative and directional PPs. In G. Cinque and L. Rizzi (eds.) Mapping spatial PPs (The Cartography of Syntactic Structures 6). New York, NY: Oxford University Press. 74–126.

Dikken, Marcel den and Éva Dékány. 2018. Adpositions and case: Alternative realization and concord. Finno-Ugric Languages and Linguistics 7. 39–75.

Dikken, Marcel den and Éva Dékány. 2025. Functional architecture of prepositional phrases. In S. Barbiers, N. Corver and M. Polinsky (eds.) The Cambridge handbook of comparative syntax. Cambridge: Cambridge University Press. 426–463.

Dougherty, Thomas. 2017. Ainu. In L. Campbell (ed.) Language isolates. London: Routledge. 100–116.

É. Kiss, Katalin. 2002. The syntax of Hungarian. Cambridge: Cambridge University Press.

Egli, Hans. 1990. Paiwangrammatik. Wiesbaden: Otto Harrassowitz.

Erdal, Marcel. 2004. A grammar of Old Turkic. Leiden: Brill.

Greenberg, Joseph. 2000. Numeral. In G. Booij, C. Lehmann, J. Mugdan, S. Skopeteas and W. Kesselheim (eds.) Morphologie / Morphology: Ein internationales Handbuch zur Flexion und Wortbildung. An international handbook on inflextion and word-formation, Vol. 1. Berlin & New York, NY: Walter de Gruyter. 770–783.

Hammarström, Harald. 2010. Rarities in numeral systems. In J. Wohlgemuth and M. Cysouw (eds.) Rethinking universals: How rarities affect linguistic theory. (Empirical Approaches to Language Typology [EALT] 45). Berlin & New York, NY: De Gruyter Mouton. 11–60.

Hamp, Eric P. 1968. On Proto-Ainu numerals. In B. J. Darden, C.-J. N. Bailey and A. Davison (eds.) Proceedings of the Fourth Regional Meeting of the Chicago Linguistic Society. Chicago, IL: Department of Linguistics, University of Chicago. 337–342.

Hanke, Thomas. 2005. Overcounting in numerals. LINGUIST List 16. 2448.

Hattori, Shirō (ed.). 1964. Ainugo hōgen jiten [An Ainu dialect dictionary]. Tokyo: Iwanami Shoten.

He, Chuansheng. 2015. Complex numerals in Mandarin Chinese are constituents. Lingua 164, A. 189–240.

He, Chuansheng. 2018. Overcounted numerals in West Yugur and Old Turkic and their compositional semantics. Journal of Minzu University of China (Philosophy and Social Sciences Edition) 45. 161–168.

He, Chuansheng, One-Soon Her, Xiaoshi Hu and Weijing Zhu. 2017. Overt coordination in additive numerals of minority languages in South China. Syntax 20. 292–316.

Hokkaido University. 2022. Hokkaido University campus guide map. https://www.math.sci.hokudai.ac.jp/sympo/fpsac2025/public/campusplan.pdf. Public Relations Division, Department of Public Engagement, Hokkaido University.

Honti, László. 1993. Die Grundzahlwörter der uralischen Sprachen. Budapest: Akadémiai Kiadó.

Hurford, James R. 1975. The linguistic theory of numerals (Cambridge Studies in Linguistics 16). Cambridge: Cambridge University Press.

Hurford, James R. 1987. Language and number: The emergence of a cognitive system. Oxford: Basil Blackwell.

Hurford, James R. 2003. The interaction between numerals and nouns. In F. Plank (ed.) Noun phrase structure in the languages of Europe. Berlin & New York, NY: Mouton. 561–620.

Hutton, James Henry. 1916. Rudimentary grammar of the Sema Naga language with vocabulary. Shillong: Assam Secretariat Print Office.

Ionin, Tania and Ora Matushansky. 2006. The composition of complex cardinals. Journal of Semantics 23. 315–360.

Ionin, Tania and Ora Matushansky. 2018. Cardinals: The syntax and semantics of cardinal-containing expressions. Cambridge, MA: MIT Press.

Kayne, Richard S. 2005. *Here* and *there*. In R. S. Kayne (ed.) Movement and silence. New York, NY: Oxford University Press. 65–84.

Kindaichi, Kyōsuke. 1931/1993. Ainu yūkara gohō tekiyō [An outline grammar of the Ainu epic poetry]. In K. Kindaichi. Ainu jojishi yūkara no kenkyū 2. Tokyo: Tōyō Bunko. 3–233. Reprinted in 1993: Ainugogaku kōgi 2 [Lectures on Ainu studies 2] (Kindaichi Kyōsuke zenshū: Ainugo I, Vol. 5). Tokyo: Sanseidō. 145–366.

Kindaichi, Kyōsuke. 1960. Ainugo kenkyū [Ainu language studies]. Tokyo: Sanseidō.

Kindaichi, Kyōsuke and Mashiho Chiri. 1936/1974. Ainu gohō gaisetsu [An outline of Ainu grammar]. In M. Chiri. Chiri Mashiho chosakushū [Selected writings of Chiri Mashiho], Vol. 4. Tokyo: Heibonsha. 3–197.

Kirikae, Hideo. 2005. Ainu-go-no mō-hitotu-no iti [Another 'one' in Ainu]. Hokkaidō minzoku gakkai kaihō 1. 3–11.

Kirikae, Hideo. 2006. Ainu-go-no iti-o simesu sūsi [Numerals indicating 'one' in Ainu]. Gengo Kenkuyu 129. 227–242.

Laufer, Berthold. 1917. The vigesimal and decimal systems in the Ainu numerals: With some remarks on Ainu phonology. Journal of the American Oriental Society 37. 192–208.

Marcos-Marín, Francisco A. 1992. Typology of numerals: Minority languages of China. In S. Luksaneeyanawin (ed.) Pan-Asiatic linguistics: Proceedings of the Third International Symposium on Language and Linguistics, Vol. II. Bangkok: Chulalongkorn University Printing House. 754–763.

Matisoff, James A. 1997. Sino-Tibetan numeral systems: Prefixes, protoforms and problems (Pacific Linguistics, Series B, Vol. 114). Canberra: Research School of Pacific and Asian Studies, The Australian National University.

Matushansky, Ora. 2019. The case of restricted locatives. In M. T. Espinal, E. Castroviejo, M. Leonetti, L. McNally and C. Real-Puigdollers (eds.) Proceedings of Sinn und Bedeutung 23, Vol. 2. Barcelona: Universitat Autònoma de Barcelona, Bilbao & Leioa: Ikerbasque & UPV/EHU, Alcalá de Henares: Universidad de Alcalá & Barcelona: Universitat Pompeu Fabra. 161–177.

Mazaudon, Martine. 2008. Number building in Tibeto-Burman languages. In S. Morey and M. Post (eds.) North East Indian linguistics, Vol. 2. New Delhi: Cambridge University Press India. Halshs-00273445. 117–148.

McCabe, Robert Blair. 1887. Outline grammar of the Angami Naga language. Calcutta: Superintendent of Government Printing.

Meinunger, André. 2015. Complex numeral expressions: A plea for grafts. Syntax 18. 103–123.

Menninger, Karl. 1969. Number words and number symbols: A cultural history of numbers. Cambridge, MA: MIT Press. English translation of the revised German edition "Zahlword und Ziffer: Eine Kulturgeschichte der Zahlen".

Merrifield, William R. 1966. Number names in four languages of Mexico. In H. Brandt Corstius (ed.) Grammars for number names (Foundations of Language Supplementary Series 7). Dordrecht: Springer. 91–102.

Mills, James Phillip. 1937. The Rengma Nagas. London: Macmillan and Co.

Murayama, Shichiro. 1971. Kita Chishima Ainugo: Bunkengakuteki kenkyū [Philological studies on the Northern Kurile Ainu language]. Tokyo: Yoshikawa Kōbunkan.

Nakagawa, Hiroshi. 1999. Ainu language: Present and future. In W. W. Fitzhugh and C. O. Dubreuil (eds.) Ainu: Spirit of a northern people. Washington, D.C.: Arctic Studies Center, National Museum of Natural History, Smithsonian Institution in association with University of Washington Press. 371–373.

Nakagawa, Hiroshi. 2022a. Parts of speech – with a focus on the classification of nouns. In A. Bugaeva (ed.) Handbook of the Ainu language. Berlin & Boston, MA: De Gruyter. 473–514.

Nakagawa, Hiroshi. 2022b. Verbal number. In A. Bugaeva (ed.) Handbook of the Ainu language. Berlin & Boston, MA: De Gruyter. 573–612.

Nakagawa, Hiroshi, Anna Bugaeva, Miki Kobayashi and Yoshimi Yoshikawa (eds.). 2016–2024. A glossed audio corpus of Ainu folklore. Tokyo: National Institute for Japanese Language and Linguistics, English HP: https://ainu.ninjal.ac.jp/folklore/en/.

Nelson, Diane and Ida Toivonen. 2000. Counting and the grammar: Case and numerals in Inari Sami. Leeds Working Papers in Linguistics 8. 179–192.

NINJAL, National Institute for Japanese Language & Linguistics. 2015. A topical dictionary of conversational Ainu. With audio and video recordings by Mrs Setsu Kurokawa [software]. Available from https://ainu.ninjal.ac.jp/topic/en/, edited by A. Bugaeva and S. Endō, speaker: Setsu Kurokawa, design and development by Shiro Akasegawa (Lago LWP).

Ochiai, Izumi. 2014. Overcounting in North-Central Vanuatu languages and its distribution. Kyoto University Linguistic Research 33. 229–252.

Ochiai, Izumi. 2021. Ainu-go sūsi saikō: Nizyussin-hō-niokeru kahou-sanpou-kara zyōhō-sanpō-e [Ainu numerals revisited: The shift from undercounting to overcounting in the vigesimal system]. Hoppō gengo kenkyū [Northern Language Studies] 11. 99–121.

Ochiai, Izumi. 2023. 'Ezo-Koku Hōkokusyo'-no Ainu-go sūsi-niokeru zyōi-kisanhō-no syōsitu [Ainu numerals in *Relazione del Regno di Iezo*: Disappearance of overcounting numeral system]. Hoppō gengo kenkyū [Northern Language Studies] 13. 283–296.

Pallas, Peter Simon. 1977–1978 [1786–1789]. Linguarum totius orbis vocabularia comparativa. Hamburg: Helmut Buske Verlag.

Patrie, James. 1978. The genetic relationship of the Ainu language. Doctoral dissertation. University of Hawaii, Honolulu, HI.

Patrie, James. 1982. The genetic relationship of the Ainu language. Oceanic Linguistics Special Publications 17. Honolulu, HI: The University Press of Hawaii.

Peng, Fred C. C. and Barron Brainerd. 1970. A grammar or Ainu number names. Lingua 25. 381–397.

Pesetsky, David. 2013. Russian case morphology and the syntactic categories (Linguistic Inquiry Monograph 66). Cambridge, MA: MIT Press.

Pfizmaier, August. 1851. Untersuchungen über den Bau der Aino–Sprache. Vienna: Kaiserlich-Königliche Hof- und Staatsdruckerei.

Pott, August Friedrich. 1847. Die quinare und vigesimale Zählmethode bei Völkern alle Weltheile: Nebst ausführlicheren Bemerkungen über die Zahlwörter indogermanischen Stammes und einem Anhange über Fingernamen. Halle: C. A. Schwetschke und Sohn.

Pylkkänen, Liina. 2008. Introducing arguments (Linguistic Inquiry Monograph 49). Cambridge, MA: MIT Press.

Refsing, Kirsten. 1986. The Ainu language: The morphology and syntax of the Shizunai dialect. Aarhaus: Aarhaus University Press.

Refsing, Kirsten. 2014. From collecting words to writing grammars: A brief history of Ainu linguistics. In M. J. Hudson, A.-E. Lewallen and M. K. Watson (eds.) Beyond Ainu studies: Changing academic and public perspectives. Honolulu, HI: University of Hawaii Press. 185–199.

Riese, Timothy. 2001. Vogul. Languages of the world materials 158. Munich: Lincom Europa.

Sato, Tomomi. 2002. A basic vocabulary of the Samani dialect of Ainu. Hokkaido University Collection of Scholarly and Academic Papers 106. 91–126.

Shibatani, Masayoshi. 1990. The languages of Japan (Cambridge Language Surveys). Cambridge: Cambridge University Press.

Simeon, George John. 1968. The phonemics and morphology of Hokkaido Ainu. Doctoral dissertation. University of Southern California, Los Angeles, CA.

Simpson, Andrew and Saurov Syed. 2016. Blocking effects of higher numerals in Bangla: A phase-based analysis. Linguistic Inquiry 47. 754–763.

Steller, Georg Wilhelm. 1774. Beschreibung von dem Lande Kamtschatka dessen Einwohnern, deren Sitten, Nahmen, Lebensart und verschiedenen Gewohnheiten. Frankfurt & Leipzig: Johann Georg Fleischer.

Talmy, Leonard. 1978. Figure and Ground in complex sentences. In J. H. Greenberg (ed.) Universals of human language. Stanford, CA: Stanford University Press. 625–649.

Tamura, Suzuko. 1970. Personal affixes in the Saru dialect of Ainu. In R. Jakobson and S. Kawamoto (eds.) Studies in general and oriental linguistics: Presented to Shiroō Hattori on the occasion of his sixtieth birthday. Tokyo: TEC Company. 577–611.

Tamura, Suzuko. 1988/2000. The Ainu language (ICHEL Linguistic Studies 2). Tokyo: Sanseidō. Originally published in Japanese under the title "Ainugo" [The Ainu language] in T. Kamei, R. Kōno and C. Eiichi (eds.) Gengogaku daijiten [Encyclopaedia of linguistics]. Tokyo: Sanseidō. 6–94.

Tamura, Suzuko. 1999. Ainu language: Features and relationships. In W. W. Fitzhugh and C. O. Dubreuil (eds.) Ainu: Spirit of a northern people. Washington, D.C.: Arctic Studies Center, National Museum of Natural History, Smithsonian Institution in association with University of Washington Press. 57–65.

Tortora, Christina. 2005. The preposition's preposition in Italian. In R. Gess and E. J. Rubin (eds.) Theoretical and experimental approaches to Romance linguistics (Current Issues in Linguistic Theory 272). Amsterdam & Philadelphia, PA: John Benjamins. 307–327.

Viitso, Tiit-Rein. 1998. Estonian. In D. Abondolo (ed.) The Uralic languages (Routledge Language Family Descriptions). London & New York, NY: Routledge. 115–148.

Witter, William Ellsworth. 1888. Outline grammar of the Lhōtā Nāgā language: With vocabulary and illustrative sentences. Calcutta: Superintendent of Government Printing.

Wągiel, Marcin and Pavel Caha. 2021. Complex simplex numerals. Acta Linguistica Academica 68. 470–515.

Žoha, Lukáš, Marcin Wągiel and Pavel Caha. 2022. The morphology of compound numerals: A cross-linguistic case-study. Linguistics Beyond and Within 8. 200–217.