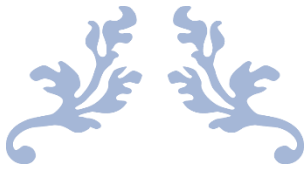
The Unseen Impact of Ainu on Japanese: The search for hidden Ainu grammar in Old Japanese

Mulder, Pepijn



The Unseen Impact of Ainu on Japanese

The search for hidden Ainu grammar in Old Japanese

Pepijn Mulder
MA thesis
30 June 2024
14939 words

Contents

“Like”, “basically”, “I mean”, and similar English catchphrases have become extremely common in the modern-day English language. This is because the younger generations have adopted them into most common conversations. It is however not only young English speakers who use them. In the Netherlands for example, it is common to hear the younger generations start many sentences with these words, and then utter the rest of the sentence in Dutch (see Onze Taal (2023) for details on Dutch “Anglicisation”). This phenomenon can largely be traced back to the rise of the internet, where people could connect to people from all over the world. The English language is very prominent on the internet, which means that people are exposed to the English language for large portions of the day. This is a form of language contact, which can result in a variety of outcomes, ranging from mutual coexistence to mutual destruction.

Language contact, in its simplest definition, is “the use of multiple languages at the same place at the same time”, as Thomason (2001, 1) so elegantly put it. However, as she notes too, it is usually more than just “using multiple languages”. In her example, when two groups of people are in the same room, they do not necessarily have to interact. If the groups do not interact, their languages do not make meaningful contact either. Therefore, for language contact to occur, there must be meaningful contact – there must be an interaction between these groups. Once meaningful contact has been established, there are multiple outcomes depending on the circumstances. The most important factor is the power balance between the two languages, that is, is the exchange equal, or does one language disproportionally influence the other? If both languages influence each other equally, there can either be mutual coexistence, or mixing. For mutual existence both languages can adapt features from the other (typically vocabulary), while mixing is the birth of a new language (Thomason 2001, 157). If there is unequal exchange – one language has more influence over another – the outcomes tend to be more extreme. In a mild case, one language simply adopts features from the more dominant language, whereas the dominant language does not adopt features from the other language. In an extreme case, however, the dominant language exerts such influence over the other language that this language is spoken by fewer and fewer people, until it ultimately goes extinct (see Thomason 2001, 223).

It is not a secret that the Ainu language is one of these critically endangered languages. While ethnic Ainu people are estimated to number in the lower ten thousand range (Nihon Keizai Shimbun, 2018), the number of native speakers is “a few” at most (Fukuzawa 2019, 3) and Ainu “as a first language is truly on the verge of extinction” (Bugaeva 2022, 1). A survey conducted by the Hokkaido government in 2017 showed that out of the 671 people of Ainu descend who responded to the survey, only 0.7% would "be able to have a conversation in the Ainu language", while 92.7% would "not be able to have a conversation” (Hokkaido government 2017, 49).

Ainu was once much more widespread. We know this because the Ainu language and its history have already been researched and documented extensively, as I will explain in the next section, but this is mostly limited to modern-day Ainu or reconstructions of Proto-Ainu (PA). By studying Old Japanese texts that were written during the Nara period (奈良時代 Nara jidai, 710-794), researchers have been able to reconstruct lists of Ainu loanwords in Old Japanese, but there seems to be a distinct lack of information about the grammar of Classical Ainu (CA) – the Ainu spoken in that time period. In this paper, I will attempt to expand the little information we have about CA by examining the closest thing we have to a written form of CA: sources in Eastern Old Japanese.

1. Background and context

Despite the Ainu people being so few in numbers nowadays, extensively has been written about their culture and history. As detailed in *The Ainu Language* by Shibatani (1990) and *The Handbook of the Ainu Language* edited by Bugaeva (2022), the Ainu people were historically distributed among the Sakhalin-Hokkaido-Kuril islands and even inhabited the Tohoku region of modern-day Japan and the southern part of the Kamchatka peninsula of modern-day Russia. However, after Russian and Japanese imperial ambitions and colonisation efforts in the 19th and 20th centuries, the areas of the Ainu peoples gradually declined, until the modern day where the last Ainu speaking people live on Hokkaido. It should be noted that the Ainu ethnicity has not been eradicated, but the Ainu people have assimilated into Russian and Japanese societies and usually are not even aware of their Ainu heritage. The Ainu did not develop a writing system and thus we do not have written records of the Ainu language, but the closest we come is via poetry like the *yukar* (ユカラ) – epics about the experiences of the *kamuy*, the gods (see Appendix B, 1). Most records of Ainu dialects are documentations of these *yukar* compiled in the 20th century by Japan and the USSR. Most historical records pre-17th century are in the form of archaeological sites in (former) Ainu territories, particularly the Hokkaido island in northern Japan. Tezuka (2009, 184) periodises the archaeological history of Hokkaido into the Epi-Jomon, Okhotsk, Satsumon, and Ainu periods. She dates the Satsumon culture from 1200 to 800 years ago (800 CE to 1200 CE), and the Ainu culture from 800 years ago to now (1200 CE to now). This dating seems to be generally accepted. Adachi et al. (2017) for example support this dating with genetic studies, claiming that the Ainu people arose from the Jomon people on Honshu and Siberian people on Sakhalin. It is therefore well-established that the Ainu people already inhabited northern Japan before the Nara period.

1.1 Developmental stages of Ainu and Japanese

Tracking the development of the Ainu languages is more difficult than that of more widespread language families. As a language isolate, that is, a language with no known related languages, the Ainu language can only be compared to itself. This is the internal reconstruction method: compare dialects to one another and older versions of the language in order to trace the language back through time. José Andrés Alonso de la Fuente compares the development of the Ainu dialects to the development of Slavic languages in chapter 5 of the *Handbook of the Ainu Language*: both underwent four distinct stages – the original language, the first language in the branch, the last common ancestor of the modern dialects, and the modern dialects – as shown in the table below:

Stage 1	Pre-Proto-Ainu	Proto-Indo-European
Stage 2	Proto-Ainu	Proto-Slavic
Stage 3	Common Ainu	Common Slavic
Stage 4	Ainu dialects (Hokkaido, Sakhalin, etc)	Slavic dialects (West, East, etc)

Table 1: Development stages of Ainu compared to Slavic (Alonso de la Fuente 2022, 179).

However, we must be careful not to equate the time periods despite the similar stages. All the stages of the Slavic languages have been dated, but the placement of Ainu stages is vaguer. We can only be confident about Common Ainu, which is the stage of Ainu development before the emergence of the Sakhalin and Kuril dialects in the 13th century (Alonso de la Fuente 2022, 178). The definition of Pre-Proto-Ainu is “the earliest stage of Ainu that can be reached by means of the comparative method and internal reconstruction” (Alonso de la Fuente 2022, 178). I would argue that these stages are

rather arbitrary, because we cannot establish a clear definition between Proto-Ainu and Pre-Proto-Ainu and Proto-Ainu and Common Ainu besides comparing them to other language families. I thus suggest that we simplify these stages to Proto-Ainu, Classical Ainu, and Modern Ainu, so we can define the stages more efficiently. Proto-Ainu (PA) would be the oldest form of the Ainu language family that can be reconstructed by means of the internal reconstruction method. Using Alexander Vovin's work *A Reconstruction of Proto-Ainu*, we can date PA to "the last centuries of the first millennium A.D., when Ainu began to move northward from northern Honshu under Japanese pressure" (Vovin 1993, 155). The Japanese started to push northward in the 8th century, so we take 700 CE as a soft boundary for PA. Common Ainu would be the stage of Ainu before it split into distinct dialects in the 13th century, and it would be what I call Classical Ainu in this research (the abbreviation of both being coincidentally CA). The last stage, Modern Ainu (MA), would be defined as the Ainu dialects that emerged after the split in the 13th century. Using this periodisation, we can define the following stages:

Stage:	**Name:**	**Approximate date:**
Stage 1	Proto-Ainu	before 700 CE
Stage 2	Common Ainu (= Classical Ainu)	about 700 to 1300 CE
Stage 3	Ainu dialects (= Modern Ainu)	about 1300 CE to now

Table 2: New periodisation of Ainu.

The Nara period spanned the 8th century, so from Table 2 we see that the contemporary Ainu dialect was Classical Ainu, but extremely close to Proto-Ainu. I will discuss this later. First, we should establish its Japanese counterparts, or the stages of the Japanese language. As established previously, the 8th century CE marked the transition of Proto-Ainu to Classical Ainu, but it also marked the end of Old Japanese. Old Japanese is a stage in the development of the Japanese language which is distinguished from later versions of Japanese by its grammar and phonology. The written Japanese language can be divided into three main stages:

Stage:	**Name:**	**Approximate date:**
Stage 1	Old Japanese (OJ)	800 CE and a few centuries prior
Stage 2	Middle Japanese (MJ)	800 to 1600 CE
Stage 3	Modern Japanese (NJ, 'new Japanese')	1600 CE to now

Table 3: Periodisation of Japanese (Adapted from Frellesvig (2010, 1))

These stages do not necessarily represent the *spoken* Japanese dialects, because those also include earlier stages like Proto-Japanese (the oldest language in the Japanese branch of the Japonic language family) and Proto-Japonic (the earliest language of the Japonic family and the ancestor of the Japanese and Ryukyuan languages). This table indicates that the language that was used in the Nara period was indeed Old Japanese, and Classical Ainu was its contemporary.

1.2 Historical contact between the Ainu and Japanese

With the periodisation in place, we can investigate the interactions between the Ainu and Japanese during the Nara period. The book *Ainu: Spirit of a Northern People* explains that there was no clear defining difference between the Ainu ethnicity and other people who lived in the Tohoku region of Honshu. Then Yamato state (the historic Japanese nation, centred around southern Honshu) started to expand into the Tohoku region in the 8th century. The people who inhabited Tohoku were referred to as Emishi by the Japanese, but this term seemed to define the "barbarians" who were not part of

the Yamato state, which means that they stopped being “Emishi” after they were assimilated into the Yamato state (Fitzhugh & Dubreuil 1999, 75).

Image 1: Approximate territorial extend of the Yamato state at the start of the Nara period in 710 CE. Note the Emishi living to the northeast of the state, in what is called the Tohoku region.[1]

While this establishes that the Japanese had contact with other people, evidence that (at least a number of) these people were Ainu can be found in the languages of that time. As we will discuss later in more detail, language contact promotes the adoption of loanwords, which was also the case with Japanese and Ainu. Fukazawa (2019), who has written a book chapter about phonological interchange between Ainu and Japanese called *Ainu Language and Ainu Speakers*, provides examples of Japanese loanwords in Ainu, with some of them starting with the sound /p/, for example:

pito ‘divine man’	from Japanese *hito* ‘person’
pone ‘bone’	from Japanese *hone* ‘bone’

[1] *Source: “History of Japan”, Wikipedia. Accessed on 11 June 2024.*

pukuru ‘bag’ from Japanese *fukuro* ‘bag’
(Fukuzawa 2019, 11)

These examples are interesting, because these words must have been borrowed from Japanese before the Japanese /p/ weakened to [h]. As it is well established that Japanese /p/ changed from [p] > [f] > [h ~ ɸ ~ ç], with the [f] stage dating to the second half of the Middle Japanese stage (Frellesvig 2010, 205). This means that these words must have been borrowed into Ainu before the 13th century. Fukuzawa also provides evidence that the Ainu were in contact with the Japanese during the Old Japanese period with these loanwords:

nomi ‘to pray’ from Japanese *nomi* ‘to pray’
kamuy ‘spirit, god’ from Japanese *kamï* ‘spirit, god’
muy ‘winnow’ from Japanese *mï* ‘winnow’
(Fukuzawa 2019, 12)

To understand this, we need to know the vowel inventory of Old Japanese. OJ had more vowels than later stages, distinguishing between /i/ and /ï/, also written as /i_1/ and /i_2/. These represented the values [i], the front high unrounded vowel, and [ɨ], the central high unrounded vowel respectively. These later merged into [i], but the difference seems to be retained in Ainu, because OJ /i/ corresponds to Ainu /i/, while OJ /ï/ corresponds to Ainu /uy/. This can only be explained if these words were borrowed before these two vowels merged in Japanese, which occurred in the transition to MJ (Frellesvig 2010, 26). We can thus be confident of the fact that the Japanese and Ainu people had frequent contact in the Nara period (8th century CE) already.

1.3 Contact-induced language change

This in turn opens the topic of language contact. I have already explained the basics of language contact in the introduction of this paper, so I will discuss contact between the Ainu and Japanese specifically. According to Thomason (2001, 60), there are broadly three main language contact typologies: change, mixture, and death. While the third one applies the most to modern-day Ainu, in the 8th century it was the first one: contact-induced language change. There are many factors that play a role in contact-induced language change – social factors and linguistic factors – that influence the effects on the receiving language and the mechanism that cause language change. Let us first examine the social factors: duration of contact and intensity of contact.

Duration is the easiest to answer. As established before, the Japanese came into contact with the Ainu people around 700 CE at the latest, and has stayed in contact with the Ainu until the present. This is a period of over a thousand years, which makes it practically impossible to have no linguistic influences between these two languages. The second factor, contact intensity, is more complex. Contact between the Japanese and Ainu during the northern wars in the 8th century was likely rather intense, when the Yamato state was subjugating the people in Tohoku, but thereafter contact intensity would differ between the Ainu in Tohoku and those on Hokkaido: the Ainu who remained on Hokkaido had less contact with both the Japanese people and language than those living in the newly colonised northern areas. The Ainu were thus split into two different groups: Ainu with intense contact with the Japanese (those in Tohoku) and Ainu with limited contact with the Japanese, at least for a few centuries (those in Hokkaido). Besides plain territorial expansion, one of the reasons for this northern expansion was the assimilation of the Emishi mentioned before into the Yamato state. This included abandoning their native languages in favour of Japanese, which can still be seen today as the

people In Tohoku speak Japanese dialects instead of Ainu dialects. This is one of the ways we can see that the language exchange between Ainu and Japanese was not equal: as Fukazawa (2019) notes, Classical Ainu was influenced more by Old Japanese than vice versa because the number of Japanese loanwords in Ainu far outnumbers the number of Ainu loanwords in Japanese.

Intensity of language contact could also act as a determining factor of the intensity of language change that results from this contact (Thomason 2001, 66). This intensity influences what types of features get borrowed; the more intense the contact, the more likely it is to borrow features that would typically be difficult to borrow. As Thomason notes, certain linguistic features are more likely to be borrowed than others: phonological features like stress and syntactic features like word order are far more likely to be borrowed than features that are heavily integrated into a complex system, like inflectional morphology (using affixes to modify existing words to form a ones). Based on Thomason's paper I will use the following scale from most likely to most unlikely to get borrowed:

Most likely	Non-basic vocabulary items
Less likely	Relatively superficial phonological features; simple structural influences; simple function words
Rather unlikely	Syntactic features; more basic vocabulary items; derivational and affixes; deeper (morpho)phonological features
Most unlikely	Inflectional morphology; fundamental changes to syntax

Table 4: the "likeliness scale" for the borrowing of language features based on language contact intensity (Adapted from Thomason (2001, 69-71)).

Let us apply this to contact between the Ainu and Japanese languages. As mentioned previously, the goal of conquering the Emishi was to *assimilate* them into the Yamato state, which by definition promotes the adoption of the Japanese language instead of simply borrowing from Japanese, so how did this impact the Japanese language? As noted previously, contact between the Japanese language and the Ainu language was not equal. Japanese impacted Ainu far more than Ainu impacted Japanese, which implies that the Ainu people had far more intense contact with the Japanese language than the Japanese people had with the Ainu language. As a result, we should expect that there was more borrowing into the Ainu language than into the Japanese language. If we apply Thomason's contact intensities (2001, 70-1), we should expect to see the effects of "*casual contact"* or "*slightly more intense contact"* when it comes to the influences of classical Ainu on Old Japanese, which includes *most likely* and *less likely* features to be borrowed, but probably not *rather unlikely* and *most unlikely* features.

As far as I can tell, however, no conclusive research has been done to determine to what degree the Ainu language has impacted the Japanese language during the period of intense language contact in the 8th century CE. Many lists of loanwords on both sides have been compiled, but no grammatical counterpart has been made. This study is therefore concerned with exactly that: to what degree has Classical Ainu influenced Old Japanese? I will attempt to answer this very question by determining the intensity of language contact between the Ainu and Japanese, as explained above.

2. Methodology

To answer this question, there are two smaller questions that must be answered. The first one is what the grammar of Classical Ainu was like, and the second one is whether this grammar is represented in Old Japanese. The first question is important, because we will not be able to find Classical Ainu influences if we do not know what CA grammar was like. The second question is necessary, because we will not be able to answer our main research question "to what degree has Classical Ainu impacted Old Japanese?" without having an idea of possible Ainu influences on Old Japanese.

2.1 Classical Ainu grammar reconstruction

The first question is "is it possible to reconstruct the grammar of Classical Ainu?". There exists no complete analysis or documentation of the grammatical structures of CA, and the reason as to why no such documentation exists is straightforward: no Classical Ainu was ever written down, making it nigh impossible to uncover aspects of the language except for clear traces, such as loanwords in OJ as discussed in the previous section. However, written sources are not the only way to uncover features of a language from the past. Traditional songs are one way we could find some features of an older version of a language. For this we can use their *yukar* – heroic stories passed down through generations – as mentioned previously. By analysing the structures of these *yukar*, we can discover some hints at what classical Ainu must have been like. Shibatani (1990) has already written an extensive documentation of the structure of Modern Ainu, while also mentioning how these structures are attested in *yukar*. Shibatani refers to the language in which the *yukar* are written as *classical Ainu*, as opposed to the spoken language which he calls *colloquial Ainu*. His definition of Classical Ainu thus differs from the one used in this study, but nonetheless, the *yukar* have retained many features of older versions of Ainu, which makes Shibatani's observations extremely valuable to this study.

Shibatani (1990) is not the only point of reference for CA, however. Remember that, according to our definition, the version of the Ainu language that was in use during the Nara period was the earliest version of Classical Ainu, but it was very close to the "end" of Proto-Ainu. This means that Vovin's reconstructions of the phonology of Proto-Ainu and Alonso de la Fuente's reviews of Vovin's reconstructions and his own insights on PA phonology are the best estimation to reconstruct what the Ainu spoken in the Nara period sounded like. This study will thus combine the works of Shibatani and Vovin to make an estimation of what the main features of the Classical Ainu from the Nara period were. Once this reconstruction is complete, we will know what features to look for in Old Japanese. This study will compare the Yakumo, Horobetsu, Saru, Obihiro, Bihoro, Asahikawa, Nayoro, Soya, Rayciska (Sakhalin), and Kuril dialects where possible, for which Hattori (1964) is a wonderful source.

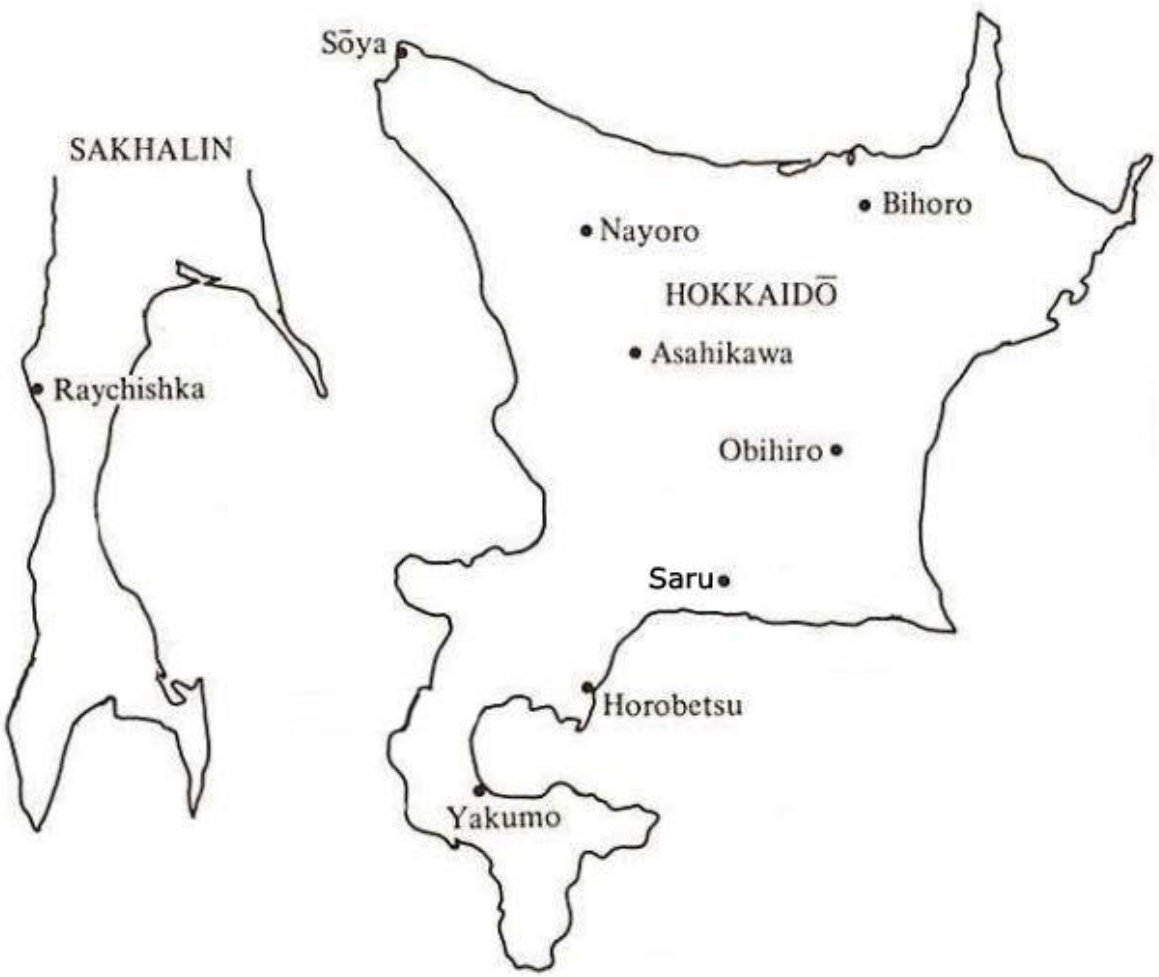


Image 2: The approximate locations of the Ainu dialects. Kuril Ainu is not shown. Adapted from Hattori (1964, 18).

2.2 Analysis of Ainu grammar in Japanese

The next question is whether we, based on the reconstruction of Classical Ainu grammar, can find influences of Classical Ainu grammar in Old Japanese. First of all, there were differences between Western Old Japanese (WOJ) and Eastern Old Japanese (EOJ), but also within these dialects. Kupchik emphasises that "Eastern Old Japanese" should not be regarded as one single dialect group, but rather as an umbrella term for two major dialect groups that were spoken in the Azuma region, which broadly was the area to the east of the Shizuoka-Nagano line but to the southwest of the region inhabited by the Emishi people. The two major dialects in this region are the western and eastern dialects, which Kupchik calls Töpo-Suruga Old Japanese (TSOJ) and Eastern Old Japanese respectively (Kupchik 2023, 1-2). This study will follow the same naming, so "Eastern Old Japanese" refers specifically to the most eastern dialect of Old Japanese, which logically had the most interactions with the Ainu living to its northeast of all Old Japanese dialects.

Sources written in WOJ far outnumber those written in EOJ. Famous sources in WOJ are the *Kojiki* (古事記), *Nihon Shoki* (日本書紀), and most books of the *Man'yōshū* (万葉集). The primary sources for EOJ are book 14 and 20 of the *Man'yōshū*, in which the poems are written predominantly in EOJ with only a handful poems having been written in WOJ (Vovin 2012, preface). This study will use

Man'yōshū book 14 (referred to simply as MYS henceforth), because analysing both book 14 and 20 would be too extensive for the scope of this study. The analysis of the MYS poems will focus on structural features that have not yet been extensively researched. This thus excludes Ainu loanwords, and will focus on grammar. I will focus my efforts on the features of the reconstructed CA grammar, because that will indicate what to look for and thus focus this study. I will note all potential CA grammar (be it either a direct copy or remnant of a copy), after which I will discuss the findings.

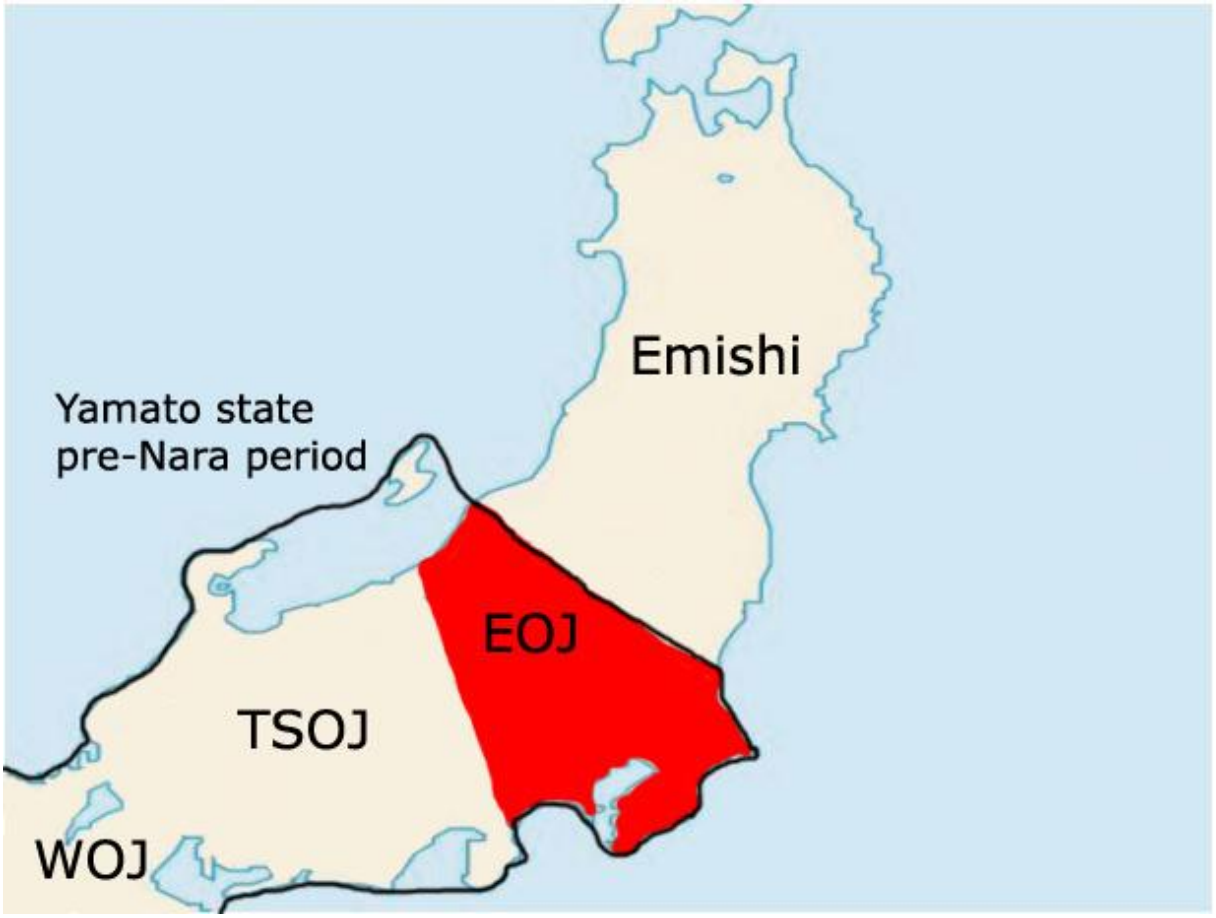


Image 3: Map of the approximate areas of the major Old Japanese dialects relative to the Yamato state and the Emishi. The red area shows the approximate area where the Eastern Old Japanese dialect of this study was spoken.[2]

In the discussion I will analyse the findings from the MYS and attempt to determine which findings are likely CA borrowings and which are unlikely to be CA borrowings, based on the CA reconstruction in this study and EOJ documentations like the one from Kupchik (2023). This brings us to the final part of this study. I will compare the findings that are most likely to be CA borrowings according to the "likeliness scale" mentioned in the section 1.3, so I can answer the main question of this study: "to what degree has Classical Ainu impacted Eastern Old Japanese?" As a bonus result to answering this question, this study will also indirectly determine how intense the contact between the Ainu and Japanese was as seen from a Japanese point of view. As explained in section 1.3, a plethora of

[2] *Image is adapted from figure 1.1 in Kupchik (2023) and the one presented on* https://commons.wikimedia.org/wiki/File:Yamato_en.png

loanwords have already been found in both Japanese and Ainu, confirming that contact did indeed occur and that "*casual contact*"-level language change did happen.

3. Reconstructing the structure of Classical Ainu

Let us first reconstruct the phonology of Classical Ainu. Thankfully, this is rather easy to do for two reasons: 1) extensive research has been done on reconstructing the phonology of Proto-Ainu, the stage of Ainu before Classical Ainu, and 2) the internal reconstruction method, the only possible method for language isolates, allows us to reconstruct PA phonology for a thousand to 1500 years ago, with Vovin and Alonso de la Fuente dating it to around the Nara period to ~1000 CE (Vovin 1993, 155; Alonso de la Fuente 2022, 148). This means that we can simply apply their reconstruction of PA phonology to our reconstruction of CA. In this study I will use the reconstruction of Alonso de la Fuente because it is the most recent one, in which he reviewed the data from earlier studies to reconstruct the best possible phonology of PA. He reconstructs the following consonants:

	Labial	Alveolar	Palatal	Retroflex	Velar	Glottal
Stop	p, p^j	t, t^j		ʈ	k, (k^j) (k^w)	ʔ
Fricative		s, (s^j)				(h), (h^j)
Nasal	m	n				
Approximant		r/l	j		w	

Table 5: Proto-Ainu reconstructed consonants

There is still some uncertainty about the existence some of these, which are shown here in brackets. There is also some uncertainty about whether PA had an /r/ or /l/. These might have been allophones of one phoneme, or separate phonemes that merged, or incorrect reconstructions of one single phoneme. However, as Alonso de la Fuente notes, "there is no material possibility to reconstruct medial *-l-" (Alonso de la Fuente 2022, 157), so I will use <r> to refer to this sound in this study. I will avoid using the other uncertain sounds unless they make more sense than alternative reconstructions. The vowel inventory of PA is reconstructed as a simple five-vowel system (Alonso de la Fuente 2022, 150):

	Front	Mid	Back
Close	i		u
Mid	e		o
Open		a	

Table 6: Proto-Ainu reconstructed vowels

Without historical documentation such as a written text, the grammar is far more difficult to reliably reconstruct than phonology, which is why no reconstruction of Nara-period Ainu grammar exists yet. This study will thus be able to reconstruct a very basic grammar based primarily on MA; the reconstruction of CA grammar presented below is merely an informed estimation based on the limited information we have. This study is predominantly based on the work of Shibatani (1990) and Hattori (1964).

In the following sections, I will provide my reconstruction in the following pattern:
*[*reconstruction*] (<MA [*modern Ainu word*])

While Shibatani has already provided some information of CA grammar, he writes these words and affixes in the modern pronunciation. For example, he writes the CA verb prefix **eti-* as *eci-*, because the MA pronunciation of /t/ followed by /i/ is /tɕ/. In these instances, I will use the CA reconstruction instead of Shibatani's modern spelling. This also means that I will use <j> for /j/ instead of the more common <y>. MA will be transliterated in the common method, which includes <c> and <y>.

3.1 The basics of Ainu grammar

We can make the strongest predictions about the basic, fundamental structure of Ainu. Modern Ainu is a head-final language in most ways, meaning that it places the core part of a phrase (typically the noun, pre/postposition or verb) after the additions:

- Noun before postposition;
- Attribute before noun;
- Relative clause before noun;
- Genitive before noun;
- Demonstrative before noun;
- Quantifier before noun;
- Adverb before main verb;
- Verb before auxiliary verb;
- Sentence-final particles (except the negation particle, which comes before the verb).

This makes the language an SOV language too, placing the subject first, object second, and verb last. While it is possible for languages to change these listed word orders (see Hawkins Word Order Universals), Shibatani seems to claim that CA featured roughly the same word order as Modern Ainu. Therefore, we will reconstruct CA with this same word order. The second point is the language's degree of synthesis. Modern Ainu is polysynthetic, glueing a variety of affixes to the verb stem. According to Shibatani, Ainu verb affix order is as follows:

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
subject	object	applicative 1	generalised object/ reflexive/ reciprocal	applicative 2	verb stem	plural	causative	iterative

Table 7: Ainu affix order (Shibatani 1990, 76)

For this reconstruction we will adopt this pattern, mainly for a lack of information. These affixes are explained in more detail in the Verbs section.

3.2 Nouns

Nouns are relatively straightforward in Modern Ainu, because there is barely any morphology. This is not to say that nothing happens with nouns, on the contrary, what it lacks in morphology it makes up for with particles.

3.2.1 Case particles

MA features case particles for the following noun cases: dative, locative, allative, ablative, instrumental, and comitative, as well as other similar particles.

3.2.1.1 Dative ‘to, for’

Y *orun; oroun*
H *orun*
S *eun*
O *eun*
B *oren*
A -
N -
S -
R *otta*
K -

The variants *orun* and *oroun* can be explained as: *oroun* was first and /ou/ > /u/ which leads to *orun*. The variant *eun* shares *-un*, so *oroun* and *eun* can be analysed as *oro-un* and *e-un*. I suspect the *e-* is a locative or allative applicative (see section 3.3.3) because the only other *e-* prefix is the second person verb prefix, which would not logically explain *eun*. If true, this would mean that *un* is a verb. This is tentatively supported by the Ainu dictionary, where *un* indicates a person’s or object’s location or reaffirms a statement, and in both cases it seems to function like a verb (Batchelor 1905, 476). The combination *e-un* ‘LOC.APPL-to.be’ thus seems plausible. Batchelor (1905, 320) gives *oro* ‘to be in; at, by; inside of’ as MA word, which would give *oro-un* ‘to be at/in’ and thus indicates essentially the same as *e-un*. *Otta* is a locative particle (see below), which strengthens the analyses of *oroun* and *eun* as locative constructions and indicates that Ainu does not differentiate much between the dative and locative. *Oren* is the only variant I cannot explain, so I tentatively assume it is a case of the irregular sound change /ou/ > /e/. I will therefore reconstruct the dative particles as **oro-un*; **e-un*.

3.2.1.2 Locative ‘in, at’

Y *ta ~ orota; otta*
H *ta; otta*
S *ta*
O *ta; otta*
B *ta*
A *-ta; -otta*
N *-ta*
S *ta; otta*
R *ohta*
K -

The dialectal variants here can all be relatively easily explained. Batchelor (1905, 435) offers *ta* ‘at; to; in’, and *oro-* is likely the same as in the dative *oroun*, which gives *oro-ta*. *Otta* is a contraction of *orota* similar to *oroun* > *orun*. This is supported by Vovin (1993, 39) who reconstructs the PA cluster -rt- from MA -tt-, so *orota* > *orta* > *otta*. *Ohta* is a case of p/t/k > h in coda position as is common in Rayciska (Sakhalin) Ainu (Shibatani 1990, 11-2). We can thus confidently reconstruct **oro-ta* as the CA locative particle.

3.2.1.3 Allative ‘(in)to’

Y *orun ~ un*
H *un*

S *un*
O *en*
B *ene*
A *ekota*
N *orun*
S *onne*
R *onne*
K -

Contrary to the locative, the variants here cannot be so easily explained. *Orun* ~ *un* seems to have the same etymology as the dative particle, but *en* and *ene* are a mystery. Batchelor does not provide a definition for *en* and *ene* that would make sense (1905, 103-4). *Onne* is similarly mysterious, because it seems to be *on-ne* with *ne* 'to become', but I cannot find an etymology for *on* that makes sense. The only analysis I can think of for *ekota* is *e-ko-ta* 'LOC.APPL-to-LOC' (*ko* 'to' from Batchelor 1905, 232), and given that it is only attested in one dialect this must be a later invention. The only particle we can confidently reconstruct is **oro-un*, the same as the dative particle. It seems that the dative particle broadened in usage to include the allative function in some dialects, while other dialects formed seperate allative particles. I will thus reconstruct **oro-un* as the original allative particle.

3.2.1.4 Ablative '(away) from'

Y *orowa*; *kari*
H -
S *or(o)wa*
O *wa*
B *wa*
A *orowa*; *kari*
N -
S -
R *orowa*
K -

Here too we see a transparent *oro-wa*, so it seems *wa* is a later simplification similar to *orota* > *ta*. *Kari* is noted by Batchelor (1905, 212) as 'by; through', but I am unable to reconstruct its original CA form. *Orowa* can be reconstructed in two ways: **oro-wa* and **oro-k*^{w}*a* where PA **k*^{w} > MA *w* (Alonso de la Fuente 2022, 160-1). Shibatani (1993, 35) claims that an optional *-no* can be added to *(oro)wa*, but this is not supported by Hattori (1964, 319). See 3.2.1.8 for more details on this. We can thus reconstruct the ablative particle as **oro-wa*/*-k*^{w}*a*.

3.2.1.5 Instrumental 'with; using'

Y *ari*
H *ari*
S *ari*
O *ari*
B *ani*
A *ari* ~ *ani*
N *ani*
S *onne*

R *ani*
K -

Except for *onne* we find the same particle in all dialects: *ari* ~ *ani* (for r > n see 3.2.4). Alternatively, *ani* might be derived from *ani* 'to hold', although 'to hold' being *anpa* instead of *ani* in Asahikawa and Nayoro dialects weakens this possibility. I will thus reconstruct **ari* as the instrumental particle. For *onne* see 3.2.1.3.

3.2.1.6 Comitative 'with; accompanying'

Y *tura*
H *tura*
S *tura*
O *tura*
B *tura*
A *tura*
N *tura*
S *tura*
R *tura*
K -

The most straight-forward particle, we can confidently reconstruct the comitative particle as **tura*. Shibatani (1993, 35) and Bugaeva (2022, 35) claim that an optional -*no* can be added to *tura*, but this is not supported by Hattori (1964, 320). See 3.2.1.8 for more details.

3.2.1.7 Translative 'into (a shape or form)'

Y *ne*
H *ne*
S *ne an*
O *ne*
B *ne*
A *ne*
N *ne*
S *nean*
R *ne an*
K -

This particle is transparently derived from the verb *ne* 'to be, to become', complemented by *an* 'to be' in certain dialects.

3.2.1.8 Absessive 'without'

Y *sak*
H *sak*
S *sak*
O *sak*
B *sak*
A *sak*
N *sak*
S *sak*

R *e … sah*
K -

This particle is transparently derived from the verb *sak* ‘to lack; to not exist’, with the expected k > h in the Sakhalin dialect (see 3.2.1.2). Shibatani (1993, 35) claims that an optional *-no* can be added to *tura*, but this is not supported by Hattori (1964, 320). If we follow Shibatani and Bugaeva and assume that *-no* can indeed be added to *oro-wa*, *tura* and *sak*, there is one simple explanation: *-no* is an adverb suffix. I say this because *sak* is transparently a verb, and *tura* also takes person marking like a verb (Shibatani 1993, 36), and both Shibatani (1990) and Bugaeva (2022) also analyse *-no* as an adverb suffix. *Orowa* is the exception, however. Batchelor (1905, 322) only provides non-verb definitions of *orowa* and *wa*, so while this analysis of *-no* fits for *tura* and *sak*, it does not fit with *orowa*. It is therefore not clear what *-no* adds to *orowa*, and whether it can be added to other case particles.

3.2.1.9 Topic

Y -
H -
S *anak(-ne)*
O *anak*
B *anakne*
A -
N -
S -
R -
K -

This particle is limited to three dialects (Saru, Obihiro, and Bihoro) and I suspect the reason is that it has fallen out of use in most dialects in MA, because it was also not common in CA (Shibatani 1993, 38). The *-ne* suffix seems to be *ne* ‘to be, to become’, which is supported by Bugaeva (2022, 50) who analysed *anak-ne* as TOP-COP. I will thus reconstruct it as **anak(-ne)*.

3.2.1.10 Additive ‘and’

Y *newa; neun*
H *newa*
S -
O *newa*
B *newa*
A *newa*
N *newa*
S -
R *neewa … tura*
K -

This is likely a contraction of the copula *ne* ‘to be, to become’ and the particle *wa* ‘and’ that connects clauses. Similarly to **oro-wa*, there are two plausible reconstructions: **ne-wa* and **ne-k^{w}a*. The Rayciska *neewa* is explained by “allophonic vowel length […] on stressed syllables” (Alonso de la Fuente 2022, 149).

3.2.2 Possession

Notably there is no dedicated genitive particle, which is odd given that Ainu has many other cases. According to the theory of the case hierarchy (see Appendix B, 2) we would expect there to be a genitive particle since it has cases from lower on the hierarchy. However, while there is no genitive *particle*, there is a genitive *construction*. As one of the only forms of morphology nouns can take, Ainu employs possessive prefixes to indicate a genitival relation. In MA, these possessive prefixes are the same as transitive subject prefixes on verbs. If we assume this also was the case for CA, we get:

	Singular	Plural
1st	**a-*	**a-*
2nd	**e-*	**eti-*
3rd	*-V(-*ʔi/hʲi*)	*-V(-*ʔi/hʲi*)

The third person possessive suffix is the result of an old possessive suffix in the form of a vowel followed by the word *ʔi/hʲi* 'thing, time, place'. In MA the /i/ has been re-analysed as the preceding vowel: PA **tek-e ʔi/hʲi* -> **tek-e-hi* -> MA *tekehe* (Alonso de la Fuente 2022, 151).

3.2.3 Nominalisers

The other morphology CA nouns take is nominalisers. Verbs can be nominalised through a variety of suffixes:

**-pe* (<MA -*p(e)*) person/thing that is characterised by X

**-i* (<MA -*i*) nominaliser for place or time verbs

**-ike* (<MA -*ike*) person/thing that is X

*-*kur* (<MA -*kur*) person with the trait X[3]

This seems be related to *kuru* 'person'.

3.2.4 Pronouns

The personal pronouns were:

	Singular	Plural
1st	**a-sir-oma*	**a-oka(-i)*
2nd	**e-sir-oma*	**eti-oka(-i)*
3rd	**sir-oma*	**oka(-i)*

The *-i* seems to be optionally added to *oka* (but seemingly not to *oma*), and is the nominaliser suffix explained above. It nominalises the plural verb *oka* 'to exist (plural)'. Thus: **a-oka-i* 1st-exist.pl-NOM for 'we'. Shibatani gives *sinuma* for the third person and *asinuma* and *esinuma* for first and second person singular pronouns, which may come from **sir-oma* where *sir* may be an 'ambient prefix' with an existential verb *oma* (Shibatani 1990, 31). This is supported by Batchelor's dictionary, where we find *shiroma* 'she, he, it' (Batchelor 1905, 419), and he gives 'to be inside' for *oma* (page 313), so **sir-oma* might have originally meant 'that which is inside it' (possibly referring to one's soul?), which was used to indicate an object or person, and then the first and second person prefixes were added to

[3] Shibatani does not elaborate too much on the nuances between *-pe*, *-ike,* and *-kur*, so I am unable to reconstruct their original meanings either.

give all these personal pronouns. Remember the variation of the instrumental particle *ari*, which was *ani*. This gives two instances of *r* > *n* between vowels.

3.3 Verbs

Contrary to nouns, verbs take a variety of affixes in the form of personal affixes, tense-aspect-mood modifiers, and some other affixes. First will be the personal affixes.

3.3.1 personal agreement

CA featured polypersonal agreement and to some degree tripartite alignment. This means that the intransitive subject, transitive subject, and transitive object affixes are different, unlike a nominative-accusative system like English where the transitive subject and intransitive subject are the same, or an ergative-absolutive system where the intransitive subject and transitive object are the same (see Dixon 1993 for a detailed explanation). The affixes are:

	Transitive subject		Intransitive subject		Transitive object	
	Singular	Plural	Singular	Plural	Singular	Plural
1st	**a-*	**a-*	**-an*	**-an*	**i-*	**i-*
2nd	**e-*	**eti-*	**e-*	**eti-*	**e-*	**eti-*
3rd	**-*	**-*	**-*	**-*	**-*	**-*

Table 8: Classical Ainu personal prefixes

From this table we can see that CA features tripartite alignment in the first person only, with the second and third persons having the same affixes in all three categories. These affixes can be combined to create polypersonal agreement, which is done by placing affixes in a subject-object-verb pattern (see verb affix order above). Plurality of the third person subject can optionally be indicated by means of the verb suffix **-pa*: **kor-pa* 'they speak'.

3.3.2 TAM-system

Next, we will analyse the tense-aspect-mood system, starting with tenses. This is very easy, because CA most likely did not have grammatical tense. This is most evident from MA, which also does not have grammatical tense and instead uses a large inventory of aspectual and modal affixes to indicate the flow of events and, in combination with context, temporal relations. While we again cannot say with any certainty what exact aspects and moods CA had, using MA as a basis we can reconstruct the following aspectual auxiliary verbs:

3.3.2.1 Causative 'make someone do'

**-ʈe* (<MA *-re, -e, -te*)
(plural causee) **-(y)ar* (<MA *-(y)ar*)

The suffixes **-ʈe* is represented by three variants in MA: (vowel-final verb +) *-re*, (r-final verb +) *-e*, (other verbs +) *-te*. I suspect the original suffix was **-ʈe* and /ʈ/ was lost after /r/ while it changed to /r/ after vowels and to /t/ after other consonants. This sound change is tentatively supported by Alonso de la Fuente (2022, 158-9), so I have reconstructed it as such.

3.3.2.2 Perfective 'have done'

**a* (sg.), **rok* (pl.) (<MA *a, rok*)

Original meaning: 'to sit (down)'. Rayciska dialect has *aa*, *roh* because of the expected variations (see 3.2.1.10 and 3.2.1.2), but all other dialects have *a, rok*. An alternative reconstruction for **rok* is **ʈok*.

3.3.2.3 Inchoative ‘about to’

**o-asi* (<MA *oasi*)

Shibatani (1993, 79) claims it comes from *o-asi* ‘hip-arise’. Hattori (1964) however does not mention this auxiliary as an aspect marker, so I can only tentatively reconstruct this auxiliary.

3.3.2.4 Progressive ‘be …ing’

Y	*kor*
H	*kor*
S	*kor*
O	*kan*
B	*kane*
A	*kor; kane*
N	*kor*
S	*kanne*
R	*kanne*
K	-

Shibatani only mentions *kor*, which is transparently derived from the verb *kor* ‘to have’ (Shibatani 1993, 79). Hattori (1964) does not provide a translation of *kane/kanne* and Batchelor (1905) does not provide a definition for *kane/kanne* to plausibly explain it. I suspect *kane* and *kanne* are variants which might come from *kare* ‘to cause to do’ with /r/ > /n/ as attested before (see 3.2.1.5 and 3.2.4). Batchelor (1905, 208) also notes that *kan* is short for *kare*, which would thus explain every dialectal variant. The problem with this etymology is that a causative is a far stretch for the basis of a progressive construction. I can thus only reconstruct **kor* with any confidence.

3.3.2.5 Terminative ‘to finish’

**oke-re* (<MA *okere*)

The original meaning is the same as the auxiliary verb. The *-re* suffix may be a causative suffix, and *oke* may thus be an intransitive verb like ‘to end’. Batchelor’s Ainu dictionary lists o*kese* ‘at the end of’ (Batchelor 1905, 310), which supports **oke* being CA for ‘(to) end’.

3.3.2.6 Instantaneous ‘at once’

**-ko-san-u* (<MA *-kosanu*)

This word may be the result of a contraction of multiple suffixes, although I do not know which ones exactly. I can think of only one somewhat plausible etymology: *ko-san-u* ‘APPL-descend/flow.along-POSS’. This combination starts with an applicative suffix (see 3.3.3 below) with the verb ‘to descend/flow along’ (Batchelor 1905, 388) and a possessive suffix. The combination would thus mean ‘descend into’ or ‘flow along’. I am not sure how this got reinterpreted as an instantaneous action. This could be explained by a shift in meaning in the transition to MA, which will be difficult to uncover. I will tentatively reconstruct it this way.

Hattori (1964, 246) gives *nani* and *tane* as particles that mean ‘at once’ and does not mention *-kosanu*. I can thus only tentatively reconstruct **kosanu*.

3.3.2.7 Successive ‘occur successively’

**-rototke* (<MA *-rototo/-rototke*)

This suffix is not mentioned in Hattori (1964), so Shibatani (1993) is the only source. As Vovin (1993, 40) notes, the *-tk-* cluster was present in all dialects of MA, however these types of clusters sometimes simplify to *-tt-*. Geminates like *-tt-* can also simplify further, so *-rototo* is most likely a more recent innovation with *-rototke* being the original form. The final *-e* shifting to *-o* is likely a re-analysis based on the previous *-o-* in the suffix, similarly to the third person possessive suffix from section 3.2.3. An alternative reconstruction is **-ʈototke*.

3.3.2.8 Durative 'continually'

**-ʔi-tara/-na-tara* (<MA *-hitara/-natara*)

This suffix is not mentioned in Hattori (1964), so Shibatani (1993) is the only source. These seem to be the result of *-hi/na* + *tara*, although I do not know what these originally meant. *-na-* is a particle after verbs that indicated a completion of some kind (Batchelor 1905, 274) and *-hi-* can be added to verbs or adjectives to change them into substantives (Batchelor 1905, 148), which supports this hypothesis, although the meaning of *-tara* in this context is not clear. An alternative reconstruction for **ʔi* is **h^ji*.

3.3.2.9 Momentary 'suddenly'

**-osma* (<MA *-osma*)

Same form as *osma* 'begin', and probably the same origin. This suffix is not mentioned in Hattori (1964) so I can only tentatively reconstruct this suffix.

3.3.2.10 Trivial 'a little'

**-tek* (<MA *-tek*)

Hattori (1964, 246) only mentions *tek* in this function in the Horobetsu and Saru dialects, although he notes it as a separate word instead of a suffix. Because of this uncertainty I will only tentatively reconstruct this suffix.

3.3.2.11 Confirmatory

**ru-w-e* (<MA *ruwe)*

There are three alternative reconstructions: **ʈu-w-e*, **ru-k^w-e*, and **ʈu-k^w-e*. The construction originally meant 'the track of it' (Bugaeva 2022, 50).

3.3.2.12 Witnessed

**sir(-i)* (<MA *sir(-i)*)

An alternative reconstruction is **wir* through w > h and h > s /_i. Bugaeva (2022, 50) gives *siri* < *sir-i* 'its sight' (sight-3.POSS) as alternate form.

3.3.2.13 Hearsay-reportive

**wak^w-e-ʔi* (<MA *hawe(he)*)

Bugaeva (2022, 50) analyses this as *haw-e* 'its voice', and Hattori gives *haw-ehe* for the Rayciska (Sakhalin) and Kuril dialects. This can be reconstructed as *wak^w-e-ʔi/h^ji* (see 3.3.2). An alternative reconstruction is **saw-e-h^ji*. See Alonso de la Fuente (2022, 159-62) for an in-depth analysis of these two reconstruction patterns.

3.3.2.14 Semblative

**sum-i* (<MA *humi*)

The MA *h* might have come from PA **w*, **s*, or **ʔ* depending on which reconstruction pattern one follows (again, see Alonso de la Fuente (2022, 159-62) for this problem). The combination *wu* is not too stable (Japanese for example lost *w* before *u* in all environments) so I suspect it was either **s* or **ʔ*, leaving **ʔumi* as the only alternative reconstruction. Bugaeva (2022, 50) analysis this as *hum-i* 'its sound'.

3.3.2.15 Expectative

**kuni* (<MA *kuni*)

This word means 'probably; in order to' in MA.

3.3.2.16 Intentional

**kusu* (<MA *kusu*)

From the noun *kusu* 'reason'. This particle also indicates reasons and explanations ('because').

3.3.2.17 Conditional '(even) if'

Y *-ika; yakne ~ yakun ~ yak*
H *yakka; yakun*
S *yakka; yakun*
O *yakkay; cik*
B *akka(y); ciki*
A *yakka; yak*
N *yakka; ciki ~ yakun*
S *yakka; ciki*
R *-yahka ~ nahka; anah*
K -

Sakhalin *yahka ~ nahka* can be explained by the regular weakening of the coda stop (see 3.2.1.2) and /j/ - /n/ matching (Shibatani 1993, 13). I believe *yak* is the original form, which was later suffixed with *-un*, *-ka,* and *-ne*. *Yak* is a postposition (Batchelor 1905, 518-9), so *-un* as a copula would be possible (see 3.2.1.1 for *un*), as would *ne* 'to become' for *-ne*. Hattori (1964, 325) gives *ka* 'even; also, too' which seems a plausible origin for *-ka*. I will thus reconstruct it as **jak* with optional *-un/-ka/-ne*. *Ciki ~ cik* would have been **ʧik(i)* or **ʧʲik(i)*.

3.3.2.18 Juxtaposal 'but; even though'

Y *koroka*
H *korka*
S *korka*
O *korkay*
B *korka(y)*
A *korka*
N *korka*
S *korka*
R *ayyahka ~ anahka*
K -

This seems to be *kor(o)* ‘have, possess’ + *ka* ‘even; also, too’, which is supported by *ayyahka* ~ *anahka* from the Rayciska dialect which is transparently related to *yahka* ~ *nahka* from above. The only uncertainty is whether it was **kor-ka* with *koro-ka* being the variant, or *koro-ka* with *kor-ka* as the variant. I assume the first option, because *koroka* is the only exception, similarly to *pirika* ‘good’ in Rayciska compared to *pirka* in every other dialect. I will thus reconstruct it as **kor-ka*.

3.3.2.19 Negative ‘not’

Y	*somo*
H	*somo*
S	*somo*
O	*somo*
B	*somo; komo*
A	*somo*
N	*somo*
S	*somo*
R	*hannehka*
K	*nebe; neben; nebam*

In all likelihood **somo* was the original particle, and *komo*, *hannehka*, and *nebe(n)/nebam* are later inventions. Depending on the reconstruction pattern, an alternative is **s*^{j}*omo*. This is the only example of a particle that modifies a verb or noun that is placed before that verb or noun. See Shibatani (1990, 24-5) for an explanation why this is the case.

3.3.3 Applicatives

Continuing our list of affixes and particles, MA uses applicative prefixes to encode valency-related and directional information. Languages with applicatives tend to hold on to them because of their usefulness, so it is expected CA featured them too (Shibatani 1990, 64):

Dative-goal:	**ko-* (<MA *ko-*)
Locative:	**e-, *o-* (<MA *e-, o-*)
Allative:	**e-, *ko-* (<MA *e-, ko-*)
Ablative:	**o-* (<MA *o-*)
Instrumental:	**e-* (<MA *e-*)
Comitative:	**ko-* (<MA *ko-*)

It is noticeable that all these different functions are all expressed via only three forms: *e-, o-, ko-*, meaning that the intended meaning is largely left to context.

3.3.4 Noun incorporation

What is not left to context, is noun incorporation. Both MA and CA enjoyed incorporating nouns into the verb, CA more than MA according to Shibatani (1990, 61). The incorporated noun is typically the direct object of a transitive sentence or an intransitive subject. In the case of transitive objects, incorporation into the verb turns the verb into an intransitive one, which is reflected in the intransitive subject person marking that is now applied to the verb:

1 *cise-kar-as*
house-make-1pl
‘We made a house’ (lit. ‘we house-made’)

(Shibatani 1990, 61)

If we apply the reconstructed grammar, we get:

2 *tise-kar-an*
house-make-1pl.INTR
'We made a house' (lit. 'we house-made')

Shibatani also gives these examples (both from page 63):

3a	*inaw*	*a-ke*		*inaw-ke-an*
	wooden symbol	1sg-make	->	symbol-make-1sg
	'I made a wooden symbol'			'I made a wooden symbol'
4a	*wakka*	*a-ta-re*		*wakka-ta-re-an*
	water	1sg-draw-CAUS	->	water-draw-CAUS-1sg
	'I make [someone] draw water'			'I make [someone] draw water'

If we apply the reconstructed grammar, we get:

3b	*inau*	*a-ke*		*inau-ke-an*
	wooden symbol	1sg.TRN-make	->	symbol-make-1sg.INTR
	'I made a wooden symbol'			'I made a wooden symbol'
4b	*k^{w}akka*	*a-ta-ʈe*		*k^{w}akka-ta-ʈe-an*
	water	1sg.TRN-draw-CAUS	->	water-draw-CAUS-1sg.INTR
	'I make [someone] draw water'			'I make [someone] draw water'

Note that the transitive first person prefix becomes an intransitive first person suffix in both cases, although Shibatani also gives a variation of 4a where the transitive subject prefix is maintained: *a-wakka-ta-re* 1sg.TRN-water-draw-CAUS. This is presumably because the causative opens a new slot for the direct object, the causee. Modern Ainu can also incorporate non-direct objects, like locatives and datives, but only if they have been promoted to the status of direct object (Shibatani 1990, 62). The main way to achieve this is by using an applicative construction. For example:

5	*tek-ari*	*kar-pe*		*tek-e-kar-pe*
	hand-INST	make-NML	->	hand-APPL-make-NML
	'things made by hand'			'things made by hand'

(Shibatani 1990, 66)

3.4 Notable features

So, with these aspects explored, we now have an estimation of the grammatical structure of Classical Ainu. As stated before, this reconstruction is not meant to be an exhaustive and perfect reconstruction, rather it is a basic framework to guide us while analysing the poems of Old Japanese. This information has now informed us on what we should be looking for in the MYS:

- Prefixes on verbs and nouns;
- Negation particle before the verb;
- Case particles that differ from Japanese particles;

- Tense-aspect-mood suffixes and particles that differ from Japanese ones;
- Applicative constructions;
- Noun incorporation.

4. Analysis of MYS

In this section, I will present the findings from the MYS. This study is concerned with the poems that are written in EOJ, but I have not completely ignored the poems that were written in WOJ (see Vovin 2012, 6 for a list of poems in WOJ and EOJ). I have compiled potential CA influences into the categories presented in section 3.4, as well as a category for features that do not fit well into any other categories. I will present my findings in their relevant sections, and discuss their likeliness of being of CA origin. Vovin's analysis of the poems in the MYS are presented as 14.XXXX with XXXX ranging from 3349 to 3577. For simplicity I will omit the '14.' from the poem's designation in this section, so 'poem 3351' for example corresponds to poem 14.3351 in Vovin's book. Some notes on romanisation: Vovin (2012) indicates prenasalised consonants with a capital N (example: Nka (possessive suffix) for [*ᵑga*]), but I will indicate this with a ^{m} or ^{n} before the consonant because it is in line with the International Phonetic Alphabet (IPA) and thus more accurate.

4.1 Noun incorporation

I have attested many instances of obvious incorporation and potentially incorporated words in the MYS. First we will examine the obvious cases:

6 *puna-^{m}-pitö*
boat-GEN-person
'boatmen'
(Poem 3349)

7 *ma^{n}tara-^{m}-pusuma*
spotted-DV-cover
'multicolour covers'
(Poem 3354)

8 *minanöse-^{n}-kapa*
Minanöse-GEN-river
'The Minanöse river'
(Poem 3366)

9 *niko-^{n}-kusa*
soft-DV-grass
'soft grass'
(Poem 3370)

10 *irima-^{n}-ti*
Irima-GEN-road
'road to Irima'
(Poem 3378)

11 *sasa-^{m}-pa*
bamboo.grass-GEN-leaf
'leaves of bamboo grass'
(Poem 3382)

12 *pa-^{n}-kata*
leaf-GEN-vine
'vines of leaves'
(Poem 3412)

13 *iya-^{n}-sakari*
more.and.more-DV-become.distant
'growing more and more distant'
(Poem 3412)

14 *ura-^{n}-kare*
top.branch-GEN-wither
'top branches wither'
(Poem 3436)

15 *wo^{n}kusa-^{n}-suke-wo*
Wokusa-GEN-assistant-man
'assistant man from Wokusa'
(Poem 3450)

16 *yo-^{n}-tati*
night-LOC-depart
'departing at night'
(Poem 3480)

17 *omop-i-^{n}-kurusu*
think-NML-GEN-hard
'it is painful to think (lit.: hardness of thinking')
(Poem 3481)

18 *ne-^{n}-to*
sleep-GEN-place
'place to sleep'
(Poem 3489)

All of the examples above follow the same pattern: word-$^{m/n}$-word, where $^{m/n}$ represents either a genitive particle, attributive form of the defective copula *n-*, or the locative particle. The genitive particle is *nö* (Kupchik 2023, 168), and the attributive form of *n-* is also *nö* (Kupchik 2023, 278). The locative particle is the only exception, being *ni* (Kupchik 2023, 173). This makes it very obvious when the previous word – always a noun in this case – is incorporated into the following word, because the particle *nö/ni* is reduced to ^{n} and fuses with the following word to form a prenasalised consonant. Phonologically it looks like this: (original) *puna nö pitö* -> (*nö/ni* is reduced) *puna-n-pitö* -> (*n* cliticises onto following consonant) *puna-^{n}pitö* -> (assimilation into following consonant) *puna-^{m}pitö*.

Now we will investigate the less obvious examples of (potential) incorporation.

19 *korö-^{n}ka nino pos-ar-u kamö*
girl-POSS cloth dry-PROG-ATTR DBT
"whether the girl is drying cloth"
(Poem 3351)

20 *sipo mit-unam-u*
tide rise-TENT-ATTR
"The tide could rise"
(Poem 3366)

21 *töki na-ki*
time not.exist-ATTR
"it is timeless (let.: time does not exist)"
(Poem 3379)

22 *wo^{n}kusa kat-i*
Wo^{n}kusa win-INF
"Wokusa wins"
(Poem 3450)

23 *nipi kusa ma^{n}sir-i*
new grass get.mixed-INF
"new grass gets mixed"
(Poem 3452)

24 *ka^{n}kami kake*
mirror hang.INF
"hang the mirror"
(Poem 3468)

25 *uma kos-i-^{n}kane-te*
horse make.cross-INF-NEG.POT-SUB
"failing to make the horse cross"
(Poem 3538)

26 *kökörö sir-a^{n}s-u-te*
heart know-NEG-INF-SUB
"[they] will not know [my] heart"
(Poem 3566)

These examples can be categorised into two types: object incorporation and subject incorporation. Examples 19, 24, 25 and 26 feature an unmarked direct object followed by a transitive verb and examples 20, 21, 22 and 23 feature an intransitive subject followed by an intransitive verb or a transitive verb that is used intransitively. There is another categorisation: the form of the verb. The verbs in examples 19, 20 and 21 are in the attributive form and the rest is in the infinitive form or the subordinate form (which itself is a modification of the infinitive form, as seen in the examples). We can thus place the examples in the following table:

	Intransitive subject	Transitive object
Infinitive verb	22, 23	24, 25, 26
Attributive verb	20, 21	19

Table 9: examples 19-26 categorised.

This categorisation is important because there are rules as to what can be incorporated. In his 2020 descriptive grammar of OJ, Vovin (2020) details the usage of OJ case particles and explains that OJ had active-stative tendencies (see Appendix B, 3) because it marked inactive intransitive subjects with *-wo* (NJ *-o*) as well as transitive objects, but it marked the transitive subject and active intransitive subject with *-^{n}ga/-nö* (NJ *-ga/-no*):

Transitive subject	Active intr. subject	Inactive intr. subject	Transitive object
-^{n}ga/-nö	*-^{n}ga/-nö*	*-wo*	*-wo*

Table 10: OJ core role marking based on Vovin (2020).

However, all of these roles could be unmarked too. In example 19 we see that the transitive subject is marked with *-^{n}ka (= -^{n}ga)* and the object is unmarked. In fact, examples 19-26 all feature unmarked objects or unmarked inactive subjects, although (22) could be considered an active subject. The fact that these nouns are unmarked is the important factor here, because Yanagida (2006, 192) explains that incorporation only occurs when a noun that occupies a position where it could be marked with *-wo* is rendered without *-wo* and is placed directly adjacent to an attributive verb. This is important, because it suggests that verbs in the infinitive form are not incorporated. According to this analysis, example 19 can thus be translated more literally as “the cloth-drying of the girl”.

	Intransitive subject	Transitive object
Attributive verb	20, 21	19

Table 11: examples that feature possible noun incorporation.

To summarise this section, we have found the following types of incorporation in the MYS:

- Nouns or nominalised verbs that are incorporated into nouns or nominalised verbs by means of a genitive, locative, or copula construction;
- Unmarked patient-like nouns (transitive objects and inactive intransitive subjects) that are incorporated directly into attributive verbs.

This is different from incorporation in Ainu based section 3.3.4. Ainu does not restrict noun incorporation to attributive verbs, thus differing from the second point, and Ainu does not require a genitive, locative or copula construction in order to incorporate nouns into other nouns. This indicates that EOJ did not develop noun incorporation because of CA influence, but it most likely developed this through a different, internal development related to the interactions of attributive verbs and its arguments and its rules for phrase formation.

4.2 Prefixes

I have found the following prefixes in the MYS: *sa-*, *ma-*, *wo(^{n})-*, *kaki-*, *uti-*, *y-*, and *na-*.

4.2.1 The prefix *ma-*

The *ma-* prefix is glossed as ‘INT’ by Vovin, short for ‘intensifier’, and is not rare, as it occurs seventeen times. It occurs both in WOJ and EOJ poems in MYS, which sheds doubt on the possible Ainu origin. There are two other factors that make it highly unlikely that this is of Ainu origin: 1) we

have no *ma-* prefix in the reconstruction of CA, and 2) this could be explained from a Japanese perspective as being the grammaticalisation of the OJ word *ma* 'true' (Kupchik 2023, 153). I therefore reject the CA origin of *ma-*.

4.2.2 The prefix *sa-*

The prefix *sa-* occurs almost exclusively on the verb *ne* 'to sleep' or the noun *ne* 'sleep' in MYS, and only three times on another verb. One variation has been attested: *sa^{n}-*. It occurs fourteen times on *ne*, and four times on other words:

sa-yo '*sa*-night' (3348)
sa-nar-aku '*sa*-rumble-NML' (3358a)
sa^{n}-töpo-mi '*sa*-far-GER' (3426)
sa-wo-sika '*sa*-male deer' (3530)
sa-wo-n-piki '*sa*-reins-LOC-pull-INF' (3536)

It also occurs both in WOJ and EOJ poems in MYS, which sheds doubt on the possible Ainu origin. According to Yanagida & Whitman (2009, 119-20), the *sa(^{n})-* prefix is used on inactive nouns and inactive verbs as agreement marking in OJ, and they tentatively present a hypothetical proto-form of *sa^{n}(V)* (from *sa-nö*) which "may be related to the mesial pronouns *sa* 'thus', *so* 'that', and *si* 's/he it'". The *nö* is the genitive particle. Based on our reconstruction of CA, the only particle that might be able to explain this form is the absessive particle **sak-no*. We would expect speakers of Japanese to remove the coda (C_1) in a C_1C_2 cluster (Kupchik 2023, 132), which would make **sakno* > **sano* > *sa^{n}* plausible on a phonological basis. However, there is no reason to assume this etymology because the meaning of neither the verb **sak* 'to lack' nor the particle **sak(no)* 'without' explains the meaning of *sa-* as a marker for inactivity better than the etymology offered by Yanagida & Whitman, and would thus be a case of finding patterns that do not exist. Therefore, I reject *sa(^{n})-* as a borrowing from CA for semantic reasons.

4.2.3 The prefix *uti-*

The prefix *uti-* is attested 3 times in MYS:

uti-kapë '*uti*-cross.over' (3482)
uti-kap-i '*uti*-cross.over-INF' (3482a)
uti-na^{m}pik-i '*uti*-stretch.out-INF' (3562)

Both Vovin (2020) and Kupchik (2023) gloss it as 'PREF' for 'prefix', noting that the meaning of this prefix is unknown to us. It is attested on only two verbs: *kap-* 'to cross over' and *na^{m}pik-* 'to stretch out', which do not appear to have much in common in form and meaning. As such we indeed cannot make claims about its meaning. This prefix also occurs both in WOJ and EOJ poems in MYS, which sheds doubt on the possible Ainu origin, but there is one prefix in our reconstruction that is very similar in form to *uti-*: the second person plural verb prefix **eti-*. So could this prefix explain *uti-*? Let us look at the three instances in the MYS:

27 *suso-nö* *uti-kapë*
hem-GEN *uti*-cross.over(INF)
"like the hems, that cross over [each other without meeting]"
(Poem 3482)

28 *suso-nö* *uti-kap-i*
hem-GEN *uti*-cross.over-INF
"like the hems, that cross over [each other without meeting]"
(Poem 3482a)

29 *uti-na^{m}pik-i*
uti-stretch.out-INF
"stretching out [herself]"
(Poem 3562)

In order to make **eti*- fit, we would need to add 'you (plural)' to the sentences. For example (27) and (28) this leads to two options: "like the hems, which you cross over" or "like the hems, which cross over you". Neither option holds much ground, because the poems are written from a first person plural inclusive ('you and I') perspective, and because the hems, who cannot meet, are analogous to this perspective of the poem, similarly to the way hems do not meet. Example (29) would result in either "while you are stretching out" or "stretching out yourself", which could fit the context, because the topic of the poem is not stated. Vovin assumed it would be 'my beloved', but if we see *uti*- as a second person plural marker the poem would still make sense:

"Will you all sleep alone, stretching out (yourself) like jewel seaweed that grows at the rough rocky shore, because you all cannot wait for me?" (adapted from Vovin 2012, 248)

The main problem with this idea, however, is that the person marking would be applied to a subordinate verb *na^{m}piki* 'stretch out' instead of the main verb *nuramu* 'sleep'. We would expect either only the main verb to have person marking (like in Indo-European languages) or the main verb and subordinate verbs (like in Ainu), but not only subordinate verbs (see Appendix B, 4). An additional problem the change of /e/ to /u/ that would have occurred when OJ borrowed **eti*- from CA. A common alteration is /u/ > /o/ or vice versa, but /e/ > /u/ would be unlikely (compare the borrowings between Ainu and Japanese that Fukuzawa (2019, 9-12) gives, where there are three instances of /u/ > /i/ and two of /i/ > /u/, but all other vowel alterations are /u/ > /o/ or /o/ > /u/ and no example of /e/ > /u/). I suspect *uti*- is not borrowed from CA because of these problems.

4.2.4 The prefix *kaki-*

The prefix *kaki*- occurs only once in MYS:

30 *kaki-mu^{n}tak-i*
kaki-embrace-INF
"embracing [her]"
(Poem 3404)

There is some disagreement about the translation, because Vovin takes *kakimu^{n}taki* alone, while Kupchik takes the preceding phrase as the direct object:

31 *kamitukëno* *aso-nö* *ma-so mura* *kaki-mu^{n}tak-i*
Kamitukëno Aso-GEN INT-hemp group *kaki*-embrace-INF
"holding a fine bundle of hemp from Aso [in] Kamitukeno close to [my] bosom"
(Kupchik 2023, 293)

Regardless of the context, the verb indicates a physical action using one's arms, but it is not clear what *kaki-* adds to the phrase. Kupchik claims this is a 'tangible prefix' that indicates that the verb is executed with one's hands that is "a grammaticalization of the infinitive or nominalized form of the verb kak- 'scratch'" (Kupchik 2023, 292). There is no prefix in our reconstruction of CA that could offer an alternative, CA etymology for *kaki-*, so it is highly unlikely this prefix is of CA origin.

4.2.5 The prefix *y-*

The prefix *y-* is also attested only once in MYS:

32 *na-y-ömöp-ar-i-sö-n-e*
NEG-y-think-INF-do-DES-IMP
"[I] wish [you] are not thinking"
(Poem 3526)

Vovin suspects this prefix is the Ainu third person indefinite object prefix *i-*, which has a few problems. First, according to our reconstruction, the prefix **i-* indicates the first person singular/plural object instead of an indefinite third person. This would change the meaning of the verb to "[I] wish [you] are not thinking about me/us" but this conflicts with the context of the poems, which takes the phrase *a^{n}ka kökörö puta yukunamö tö* 'that my heart would go to two different places' as the think being not thought about. Secondly, there is an alternative explanation: this may simply be an epenthetic consonant to break up the would-be vowel cluster *aö* (*na-ömöp)*, because OJ tended to either break up vowel clusters or remove one of the vowels (see Kupchik 2023, 105-11 for a detailed overview of EOJ vowel elision). Kupchik prefers this explanation over the CA etymology because "a singular case of epenthesis is more likely than a morphological borrowing from Ainu", and I agree with this statement. Therefore, I suspect this *y-* prefix is not a of CA origin.

4.2.6 The prefix *na-*

The prefix *na-* mentioned above is somewhat common in MYS. It is attested 11 times and is found in both WOJ and EOJ poems, which again sheds doubt on the possible Ainu origin. From these 11 times it occurs 5 times on the verb *taye* 'break up/off', two times on *ömöp/omop* 'think', and one time each on *kane* 'worry', *yak* 'burn', *nar* 'rumble', *sak* 'bloom'. Vovin glosses it as 'NEG' for 'negative', but Kupchik (2023, 293) specifies it as a 'negative imperative'. Neither of them provides an etymology, but this does not mean we can assume a CA etymology. Our reconstruction offers a few particles that could become *na*, so let us discuss them. First, the locative particle **ta*. This is problematic for two reasons: it would mean that /t/ > /n/ which is not in itself problematic, but this is not a sound change common to OJ loanwords (see Fukuzawa 2019, 9-12), and the change of locative > negative is not likely because semantically these two are very different. Thus, I reject this etymology. Second, the translative particle **ne*. This particle is also problematic for the same two reasons as **ta*, although the sound change here would be /e/ > /a/ which is also not common in OJ-CA loanwords (again, see Fukuzawa 2019, 9-12). I must reject this etymology too. Lastly, the verb suffix **-na* from the durative **-na-tara*. This does not have a phonological problem, but there is still a semantic problem because **-na* indicates completion of the action, which is also very different from the negation of an action. Thus, I must reject this etymology too on a semantic basis. To conclude, there is no evidence that *na-* is a CA borrowing.

4.2.7 Possessive prefixes

Lastly, I have not found a single example of a possessive prefix in MYS. Every instance of a possessive construction featured the possessive particle, almost always *-^{n}ka*, after the noun. We can thus be confident that OJ was not influenced by CA when it comes to possessive constructions.

4.3 Negation particle

Classical Ainu placed the negation particle **somo* before the verb, for example:

33	*somo*	*oman-an*		*somo*	*ku-oman*
	not	go-1.sg.INTR	< MA	not	1.sg-go
	"I do not go"			'I do not go'	

(Adapted from Shibatani 1990, 24)

In the MYS I have only found a single instance of a possible negation that was not in the form of a verb suffix *-a^{n}z-* or one of its variations (see Kupchik 2023, 353 for details) or the *na-* negative from section 4.2.6. This is the particle *sömö* as attested in poem 3382. Vovin analysis it as follows:

34	*na-pa*	*kopu-^{n}pa*	*sömö*
	you-TOP	long.for-COND	not
	"whether you long for [me or] not"		

(Vovin 2012, 72)

Vovin notes on the following page that this cannot be the focus particle *sö* followed by the emphatic particle *mö* as many other scholars think, and tentatively analyses it as a Japanese borrowing of the Ainu negative *somo*. I disagree with this analysis for two reasons. First, this *sömö* occurs after the verb in (34) which is the opposite of how it is used in Ainu. Second, this would mean that CA /o/ became /ö/ (which represents the vowel ə) when the Japanese borrowed it, which would be illogical because OJ also had the vowel /o/, so if OJ borrowed CA **somo* it would logically borrow it as *somo* instead of *sömö*. This makes it unlikely for *sömö* to be related to **somo*.

4.4 Case particles

The MYS displays without fail only established EOJ case particles. There are, however, two case particles that resemble a CA case particle: the locative-genitive particle *-tu* and the locative particle *-na*. Both of which resemble the CA locative case particle **-ta*. So, could either of these two have an Ainu origin?

The Ainu origin of the locative marker *-na* presents two problems. First, in the same way to the negative prefix *na-*, this would mean that CA /t/ corresponds to OJ /n/ which is not supported by other loanwords. Second, a cognate of *-na* with the same locative function is attested in Proto-Ryukyuan (PR) (Kupchik 2023, 179), the ancestor of all Ryukyuan languages and a sister branch to Proto-Japanese (PJ) in the Japonic language family. This fact limits us to two possibilities: 1) the speakers of Proto-Japanese-Ryukyuan (PJR) borrowed **-ta* even before they split off into Proto-Japanese and Proto-Ryukyuan, or 2) CA **-ta* and EOJ *-na* are unrelated. There is simply no feasible way to determine the validity of the first possibility because of time. The speakers of PJR migrated from the Korean peninsula to the Japanese islands around 700 to 300 BCE (Vovin 2017) and we have no works written in PJR, let alone any idea of what the Ainu language looked like so long ago (remember, reconstructions of Proto-Ainu can only go back to about 1500 years ago, so any claims about Pre-Proto-Ainu from almost 3000 years ago are essentially just guesses). The problem here is

therefore that we cannot confidently confirm or deny the Ainu origin of *-na* because of a lack of information. I will nonetheless reject this possibility because at this point it is nothing more than pure speculation.

It is a similar situation for the locative-genitive marker *-tu*. It is very rare and attested in both WOJ and EOJ but only as fossilised combinations (Kupchik 2023, 181) so if it were an Ainu loan, it would have needed to be borrowed from Pre-Proto-Ainu before Proto-Japanese split off into the WOJ, EOJ and other Japanese dialects. While this is less far in the past as the previous case, it is nevertheless far enough in the past that we cannot make confident claims about it because of a lack of information about Pre-Proto-Ainu and its relations with neighbouring languages. I thus have to conclude that there are no instances of Ainu case markers in MYS.

4.5 TAM system

I have found only one mood marker that resembles a CA marker: *-asi*. This resembles the CA inchoative marker **-oasi*. It is attested in poems 3431, 3435, 3446, 3550 and 3557 respectively (glosses from Vovin 2012):

pik-asi-mö 'drag-ADJ-EXCL'
tuk-i-yör-asi-mö 'attach-INF-approach-ADJ-EXCL'
katar-i-yör-asi-mö 'talk-INF-approach-SUP-EXCL'
ita^{m}pur-asi-mo 'shake.violently-ADJ-EXCL'
nayam-asi-ke 'suffer-ADJ-ATTR'

This suffix occurs both in WOJ and EOJ poems and is always followed by another suffix in MYS. We can see that this suffix has two functions: adjectiviser and suppositional, although there is seemingly some disagreement concerning its functions[4]. Kupchik (2023, 351) does not provide an etymology for this suffix, but again, this does not validate an Ainu etymology. Most obviously, the forms of these two suffixes are not entirely the same: **-oasi* vs. *-asi*. If this were an Ainu borrowing, we would have to explain where the /o/ disappeared to, and I am not aware of a regular loss of /o/ in Japanese borrowings of Ainu words. Secondly, since *-asi* is attested in both WOJ and EOJ, we have the same problem as previous examples in this study: if it is indeed an Ainu borrowing, it must have been borrowed before WOJ and EOJ split off from PJ, in the time of Pre-Proto-Ainu, and we have not enough data to confirm this claim. Last but not least, the semantic roles of these suffixes are very different. **-oasi* emphasises the start of an event (inchoative) and *-asi* turns the verb into an adjective or indicates conjecture, and I cannot think of a simple logical way this shift in meaning could happen. I must therefore reject an CA etymology of the OJ *-asi*.

4.6 Applicatives

Ainu uses various prefixes to indicate applicative meanings (see section 3.3.3), these prefixes being **e-*, **o-*, and **ko-*. There are two similar prefixes in MYS: *i-* and *o-*, as seen in poems 3409, 3473, 3518, and 3540:

i-tu^{n}k-i 'I-follow-INF' (3409)
i-katar-u '*i*-hang-ATTR' (3518)
i-yuk-i '*i*-go-INF' (3540)

[4] Vovin glosses only the third example as SUP, but Kupchik (2023, 351-2) glosses only the last example as ADJ.

There are two theories as to the function of *i-*. First, Vovin and Kupchik gloss it as a direct-locative marker. Second, as offered by Yanagida & Whitman (2009, 117-9), it might be an active marker that reinforces the agency of the subject and the event as a whole. The option this study is interested in is the first one, the direct-locative marker, because it might be related to the CA locative applicative **e-*. This would mean that CA /e/ became /i/ when it was borrowed into EOJ. Fukuzawa (2019, 10) does provide a few instances where Japanese /i/ corresponds to Ainu /e/ or vice versa in loanwords (Ainu *otcike*, *kani*, *emo* – Japanese *osiki*, *kane*, *imo*), but these instances are few and far between. Furthermore, there is a morphosyntactic problem. Ainu uses the applicative to promote adjunct or oblique noun to the role of direct object, which can then be incorporated into the verb complex. If this particle *i-* is indeed an Ainu borrowing, we would expect it to fulfil this same function. However, poem 3540 disproves this:

5	*tek-ari*	*kar-pe*		*tek-e-kar-pe*
	hand-INST	make-NML	->	hand-APPL-make-NML
	'things made by hand'			'things made by hand'

Example from Ainu applicative construction (Repeat from section 3.3.4)

35	*te^{n}ko-ni*	*i-yuk-i*
	maiden-DAT	*i*-go-INF
	"I went to the maiden"	

Example from MYS (3540)

In (35) we see that the direction of the verb is marked with the non-compressed dative particle *-ni*, which means that *te^{n}ko* 'maiden' is 1) not the direct object of the sentence, and 2) not incorporated into the verb complex. These two factors demonstrate that the prefix *i-* is not an applicative marker and thus is unlikely to have CA **e-* as its origin.

Originally I added *o-yun-i* '*o*-sleep-NML' (poem 3473) to this list as a Japanese borrowing of the Ainu locative applicative **o-* because Vovin (2012, 160) offered this etymology too. Kupchik (2023, 139) has offered a more convincing explanation that is in line with EOJ grammar and phonology. I thus conclude that there is no evidence that CA influenced EOJ when it comes to applicative constructions.

4.7 Other potential Ainu influences

I have found two instances of possible CA influences in MYS that are worth discussing. First, the EOJ word *si^{n}ta* 'time; when', as found in poems 3363, 3461, 3478, 3515, 3520, and 3533. This word functions as a temporal conjunction and turns a clause into a temporal clause as in poem 3363:

36	*mat-u*	*si^{n}ta*
	wait-ATTR	time
	"when I wait"	

(Adapted from Vovin 2012, 52)

It is considered an Ainu loanword by both Vovin ad Kupchik. Vovin (2012, 53) offers three reasons why he thinks this:

1. *Si^{n}ta* is not attested in any WOJ texts;
2. It always occurs in the same grammatical position, which differs in semantic and morphosyntactic ways from a similar temporal noun *töki* 'time; when';

3. It can, based on established EOJ phonology, be traced back to the Ainu word *hi-ta* 'time-LOC', which fulfils the same function as *si^{n}ta*.

Kupchik (2023, 138) agrees with Vovin's arguments, and thus the Ainu origin of *si^{n}ta* is highly probable. There is, however, a small problem with the etymology of *hi-ta*, because this is the MA form and MA /h/ is the result of various older sounds in CA. Vovin justifies *hi* by explaining that we expect Ainu /hi/ to be pronounced as /çi/ in the same way that Modern Japanese does. This /çi/ would then have been interpreted by speakers of EOJ as /si/, which was the closest approximation possible in EOJ. The problem is the fact that MA /h/ is the result of CA /h^{j}/ and /ʔ/ depending on the phonetic environment of the word, which is why I offer both **h^{j}i* and **ʔi* as reconstructions of the *hi* in *-hitara* in section 3.3.2. Out of those two reconstructions, /h^{j}/ seems the most likely candidate for *hi* 'time' for two reasons. First, the glottal stop /ʔ/ being interpreted as /s/ before /i/ seems highly unlikely to me, because it is not a very strong and noticeable sound. Secondly, Alonso de la Fuente (2022, 163-4) explains that PA *ʔ was largely an epenthetic consonant that would be inserted at the start of vowel-initial words and in vowel clusters and thus was not a true phoneme. I prefer /h^{j}/ because it would have been phonemic, and being already palatalised, /h^{j}/ allows us to skip the assumed middle step /çi/ in /hi/ > /çi/ > /si/ and go straight from /h^{j}i/ to /si/. Following this, we get EOJ *si^{n}ta* < CA *h^{j}i-ta*.[5]

The second instance is the emphatic particle *si* as seen in poem 3400 for example:

37 *kimi* *si* *pum-i-te-^{m}pa*
lord *si* step-INF-PERF-COND
"if my lord would have stepped"
(Vovin 2012, 89-90)

Both Vovin and Kupchik gloss it as an emphatic particle, but are not able to determine more nuance nor can they provide an etymology. Ainu may shed light on this. Being *si* like the previous case, we can already say that the original Ainu word could have been **h^{j}i*. Alternatively, it could have been **si* or **s^{j}i*, because both /si/ and /s^{j}i/ would have been interpreted as /si/ by the Japanese. This leaves two possibilities for MA words: *hi* (< *h^{j}i*) and *shi* (< *si, s^{j}i*). Interestingly, Batchelor (1905) provides two words in MA that could fit this *si* particle: *hi* 'yes; so' and *shi* 'true; great; very' (page 148 and 397 respectfully). Both of these words fit the phonetic requirements and semantically they could work too. MA *hi* 'yes; so' would act as an affirmative particle which could be re-interpreted as a particle that adds emphasis, and MA *shi* 'true; great; very' would probably exaggerate or enlarge the noun, which would then be re-interpreted as emphasis. *Shi* has a problem that *hi* does not, however: *shi* is an adjective, which is placed *before* the noun it modifies (see section 3.1) which is not the case in the examples from MYS. *Hi* on the other hand is an adverb, which means it is placed right before the verb and thus after the noun, which is the case for almost every instance (see (37) for example). The strongest counter-argument is probably that *si* is attested in both WOJ and EOJ, which makes this theory improbable for reasons I have explained before. I will tentatively support OJ *si* < Ainu *hi*, assuming that MA *hi* 'yes; so' < CA **h^{j}i* and that it was borrowed from Ainu either before WOJ and EOJ split off from PJ or after they split off and WOJ borrowed it from EOJ after EOJ borrowed it from CA.

[5] CA /t/ > EOJ /^{n}t/ can be explained too: Ainu voiceless stops are often "phonetically realized as voiced or half-voiced" (Vovin 2012, 53), which would have been interpreted in OJ as a prenasalised stop.

5 Discussion: Reasons for absence of grammatical exchange and concluding thoughts

In the previous section I have analysed my findings from the MYS based on the main features of CA grammar, which were:

Grammatical feature:	**Result:**
Noun incorporation:	no convincing CA influences on EOJ;
Prefixes:	no convincing CA influences on EOJ;
Negation:	no convincing CA influences on EOJ;
Case particles:	no convincing CA influences on EOJ;
TAM-system:	no convincing CA influences on EOJ;
Applicatives:	no convincing CA influences on EOJ.

The only convincing CA grammar that was found in MYS was the temporal conjunction *si^{n}ta*, and there is one somewhat convincing etymology for the emphatic particle *si*. Looking back at Table 4, we can now try to determine the intensity of contact between the Ainu and Japanese in the 8th century.

Most likely	Non-basic vocabulary items
Less likely	Relatively superficial phonological features; simple structural influences; simple function words
Rather unlikely	Syntactic features; more basic vocabulary items; derivational and affixes; deeper (morpho)phonological features
Most unlikely	Inflectional morphology; fundamental changes to syntax

Repeat of Table 4 (see section 1.3).

In section 1.2, we already established that linguistic features from the 'most likely' (to be borrowed) category have already been confirmed in Japanese, so where should the CA features from this study be placed? Applicatives, prefixes, and case particles would be in the 'rather unlikely' category, because they rely on affixation and syntax. Noun incorporation and the TAM-system probably belong to the 'most unlikely' category, because these rely on multiple systems (noun incorporation also influences the transitivity of the verb, as explained in section 3.3.4) and are part of the basic construction of sentences. Negation probably belongs to the 'less likely' category, although this might also be a 'rather unlikely' feature depending on how integrated it is in the verb complex. Negation in Ainu is less integrated, so I will place it in the 'less likely' category. The temporal conjunction *si^{n}ta* is an example of a simple structural influence and a simple function word, so it is in the 'less likely' category:

Less likely	Negation; *si^{n}ta*
Rather unlikely	Applicatives; prefixes; case particles
Most unlikely	Noun incorporation; TAM-system

Table 12: the "likeliness scale" for the borrowing of Ainu features.

This study has found that all of these features have not been found in EOJ, which indicates that influence of CA grammar was limited to 'most likely' and 'less likely' features, although only one example of 'less likely' features is attested. This absence of 'unlikely' features in EOJ indicates that language contact between the Ainu and Japanese from the Japanese perspective was on the low-intensity side rather than the high-intensity side. I can think of two explanations for this low-intense language contact on the Japanese side. The first explanation is the nature of the sources. All documentation of OJ is in poetic written form, which differs from spoken language because poetic language is typically planned and structured, while spoken language is spontaneous, less formal, and less structured. This means that the poems in MYS do not fully and accurately represent all of EOJ, and a colloquial spoken version of EOJ might have been more influenced by CA grammar than written EOJ was. The second explanation is the power imbalance between the Ainu and Japanese. As explained in section 1.2, the Yamato state gradually conquered the Tohoku region of the Honshu island and attempted to assimilate the Emishi people. This placed the Japanese culture above the native cultures in importance, and thus pressured the Ainu from Tohoku to adopt Japanese customs and the Japanese language. As a result of these policies, there was far less incentive for Japanese people to learn Ainu than for the Ainu to learn Japanese, and this made it easier for Japanese features to be integrated into Ainu than vice versa (notice the length of the Japanese borrowings into Ainu compared to Ainu borrowings into Japanese in Fukuzawa (2019)).

This study has attempted to determine to what degree Classical Ainu has influenced Old Japanese, which consisted of two parts. The first part of this study consisted of the reconstructing of Classical Ainu grammar by comparing features of different Modern Ainu dialects and applying the reconstructed phonology of Proto-Ainu. After a basic grammar had been reconstructed, the poems of book 14 of the *Man'yōshū*, which was mostly written in Eastern Old Japanese, were analysed in order to determine whether features of Classical Ainu grammar could be found in this text. This study concluded that no features of the reconstructed Ainu grammar are present in the *Man'yōshū* 14, but one feature of Ainu grammar was attested. Including the already well-documented Ainu loanwords in Japanese, this means that 'most likely' and 'less likely' features of Classical Ainu are present in Old Japanese which indicates that, as predicted in section 1.3, contact between the Ainu and Japanese from the Japanese perspective was indeed "*slightly more intense"* – it was not so intense.

Sources

Adachi, Noboru; Kakuda, Tsuneo; Takahashi, Ryohei; Kanzawa-Kiriyama, Hideaki; Shinoda, Ken-ichi. 2017. “Ethnic derivation of the Ainu inferred from ancient mitochondrial DNA data”. *American Journal of Physical Anthropology* 165(1): 139-148.

Batchelor, John. 1905. *An Ainu-English-Japanese dictionary*. Tokyo: Tokyo Methodist Pub. House.

Alonso de la Fuente, José Andrés. 2022. “The Ainu language through time” in *Handbook of the Ainu Language*, edited by Anna Bugaeva, 147-183. Berlin: De Gruyter Mouton.

Blake, Barry. 1992. “The Case Hierarchy”. *La Trobe Working Papers in Linguistics*, 5: 1-6.

Bugaeva, Anna. 2022. “Ainu: A head-marking language of the Pacific Rim” in *Handbook of the Ainu Language*, edited by Anna Bugaeva, 23-55. Berlin: De Gruyter Mouton.

Dixon, R.M.W. 1994. *Ergativity.* Cambridge: Cambridge University Press.

Fitzhugh, William W.; Dubreuil, Chisato O. 1999. *Ainu: Spirit of a Northern People*. Arctic Studies Center, National Museum of Natural History, Smithsonian Institution in association with University of Washington Press. ISBN 978-0-96734-290-0. OCLC 42801973.

Frellesvig, Bjarke. 2010. *A History of the Japanese Language*. 1st ed. Cambridge: Cambridge University Press.

Fukazawa, Mika. 2019. “Ainu Language and Ainu Speakers.” In *Routledge Handbook of Japanese Sociolinguistics*, vol 1, p.3-24.

Hattori, Shirō. 1964. *An Ainu dialect dictionary : with Ainu, Japanese and English indexes*. Tokyo: Iwanami Shoten Publishers.

Hokkaidō Government. 2017. "平成２９年北海道アイヌ生活実態調査報告書" [2017 Hokkaido Ainu Life Survey Report]. https://www.pref.hokkaido.lg.jp/fs/5/5/7/2/6/5/7/_/2%E5%B9%B3%E6%88%9029%E5%B9%B4%20%E5%8C%97%E6%B5%B7%E9%81%93%E3%82%A2%E3%82%A4%E3%83%8C%E7%94%9F%E6%B4%BB%E5%AE%9F%E6%85%8B%E8%AA%BF%E6%9F%BB%E5%A0%B1%E5%91%8A%E6%9B%B8.pdf (accessed 11 June 2024)

Kupchik, John. 2023. *Azuma Old Japanese : A Comparative Grammar and Reconstruction*. First edition. Berlin: Walter de Gruyter.

Onze Taal [our language]. 2023. “Een ‘boom’ aan Engelse leenwoorden” [a ‘boom’ of English loanwords]. https://onzetaal.nl/tijdschrift/01-2023/artikel/een-boom-aan-engelse-leenwoorden (accessed 11 June 2024)

Shibatani, Masayoshi. 1990. “The Ainu Language” in *The Languages of Japan*. Cambridge: Cambridge University Press.

Tezuka, Kaoru. 2009. “Interaction between the Ainu of Hokkaido and Honshu Japanese during the Past 1,000 Years: Use Pattern Analysis of Wooden Artefacts from the Ishikari Lowland: Interactions between Hunter-Gatherers and Farmers: from Prehistory to Present”. *Senri Ethnological Studies*, 73, 179–201. https://doi.org/10.15021/00002591.

Thomason, Sarah Grey. 2001. *Language Contact.* Edinburgh: Edinburgh University Press.

Vovin, Alexander. 1993. *A Reconstruction of Proto-Ainu*, Vol. 4. Leiden: Brill.

Vovin, Alexander. 2012. *Man'yōshū, book 14. A New English Translation Containing the Original Text, Kana Transliteration, Romanization, Glossing and Glossary*. Leiden and Boston: Brill.

Vovin, Alexander. 2020. *A Discriptive and Comparative Grammar of Western Old Japanese*. Boston: BRILL.

Vovin, Alexander. 2017. "Origins of the Japanese Language." *Oxford Research Encyclopedia of Linguistics*. Accessed 22 Jun. 2024. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-277.

Yanagida, Yuko 柳田 優子. 2006. "古代語における動詞連体形節内の格付与について [About case assignment in adnominal clauses in Old Japanese]." *Scientific Approaches to Language* 5: 181-94. http://id.nii.ac.jp/1092/00000146/

Yanagida, Yuko, John Whitman. 2009. "Alignment and Word Order in Old Japanese." *Journal of East Asian linguistics* 18, 2: 101-44. DOI 10.1007/s10831-009-9043-2

日本経済新聞 [Japanese Economics Newspaper]. 2018. "北海道のアイヌ、10 年余で 4 割減 実態反映せず" [Ainu in Hokkaido down by 40% in 10 years, not reflecting the real situation]. https://www.nikkei.com/article/DGXMZO34630960X20C18A8CR8000/ (accessed 11 June 2024)

Appendix A: Abbreviations

This study uses the following abbreviations:

ADJ	Adjectiviser
APPL	Applicative
ATTR	Attributive
CA	Classical Ainu, Common Ainu (same definition in this research)
COND	Conditional
COP	Copula
DAT	Dative
DBT	Dubitative
DV	Defective verb (see Kupchik 2023, 287 for more information)
EOJ	Eastern Old Japanese
EXCL	Exclamation
GER	Gerund
INF	Infinitive
INS	Instrumental
INT	Intensive
INTR	Intransitive
LOC	Locative
MA	Modern Ainu
MJ	Middle Japanese
MYS	Man'yōshū book 14 (unless stated to be another book)
NEG	Negative
NJ	Modern Japanese (from New Japanese to prevent confusion with Middle Japanese)
NML	Nominaliser
NOM	Nominative
OJ	Old Japanese
PA	Proto-Ainu
PERF	Perfective
POSS	Possessive
POT	Potential
PROG	Progressive
SCA	Southern Classical Ainu
SUB	Subordinate
SUP	Suppositional
TAM	Tense-aspect-mood
TENT	Tentative
TOP	Topic
TRN	Transitive
V	Vowel
WOJ	Western Old Japanese
1	1st person
2	2nd person
3	3rd person
sg.	singular
pl.	plural

Appendix B: Additional information

This appendix provides context and additional information that does not quite fit in the main text.

1) I recommend reading the works of Chiri Yukie (知里 幸恵), who became famous for her translations of *yukar* which are published in her *Ainu Shinyōshū* (アイヌ神謡集).

2) The theory of case hierarchy was put forward by Barry Blake (1992) and states that there is a strong tendency for noun cases to follow this hierarchy from most common to most uncommon: NOM/ABS < ACC/ERG < GEN < DAT < LOC < ABL <INS/COM < others. So if a random language has the locative case, we would expect it to have the dative, genitive, accusative/ergative, and nominative/absolutive cases as well.

3) An active-stative language is a language in which nouns are treated a certain way based on their agency, which means that the transitive subject and the active intransitive subject are treated the same (through the same case marking for example) and the inactive intransitive subject and transitive object are treated the same way, as opposed to English for example which treats both active and inactive intransitive subjects like transitive subjects and only treats the transitive object differently. See Dixon (1993, 71) for more information.

4) Latin example: *celsa sedet Aeolus arce sceptra tenens* 'Aeolus sits high in the castle, holding the sceptre' (Aeneid 1, line 58-9, my translation) where the main verb *sede-t* 'sits' is marked with the third person singular marker *-t* while the subordinate verb *tenens* 'holding' does not have person marking. Modern Ainu example: *cikor hapo unhawekoyki* 'our mother scolded us' (Bugaeva 2022, 29), in which the main verb *hawekoyki* 'to scold' takes the first person plural object marker *un-* and the subordinate verb *kor* 'to have' takes the first person plural subject marker *ci-*.