



Universiteit
Leiden
The Netherlands

Vowel harmony in Ainu

Shiraishi, H.; Botma, E.D.; Ritter, N.A.; Hulst, H. van der

Citation

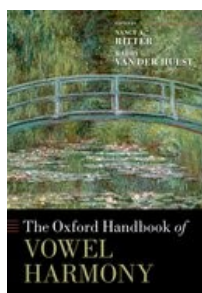
Shiraishi, H., & Botma, E. D. (2024). Vowel harmony in Ainu. In N. A. Ritter & H. van der Hulst (Eds.), *The Oxford handbook of Vowel harmony* (pp. 816-823). Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780198826804.013.64

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/4210697>

Note: To cite this publication please use the final published version (if applicable).



The Oxford Handbook of Vowel Harmony

Harry van der Hulst (ed.), Nancy A. Ritter (ed.)

<https://doi.org/10.1093/oxfordhb/9780198826804.001.0001>

Published online: 22 October 2024 **Published in print:** 10 October 2024

Online ISBN:

9780191865718

Print ISBN: 9780198826804

Search in this book

CHAPTER

64 Vowel Harmony in Ainu

Hidetoshi Shiraishi, Bert Botma

<https://doi.org/10.1093/oxfordhb/9780198826804.013.64> Pages 816–823

Published: 22 October 2024

Abstract

This chapter describes the distribution of vowels in Ainu, a linguistic isolate of Northeast Asia. Some vocalic suffixes tend to harmonize with a preceding root vowel, although the patterns are complicated by the existence of disharmonic forms. The data show that the distribution of vowels in base-suffix contexts differs from that in underived contexts, with two interesting exceptions; in both environments there is a marked preference for /o ... o/, while /o ... u/ and /u ... o/ are in general avoided. This appears to be an areal feature of Northeast Asia, since other languages in the region display similar restrictions on the distribution of /o/.

Keywords: Ainu, progressive assimilation, derived environment, areal feature, disharmonic roots

Subject: Phonetics and Phonology, Linguistics

Series: Oxford Handbooks

Collection: Oxford Handbooks Online

64.1 Introduction

Ainu is a linguistic isolate spoken by the Ainu people, an indigenous ethnic group of the islands of Honshu, Hokkaido, Sakhalin, and the Kurils in Northeast Asia. The rapid decline of the number of speakers since the end of the 19th century left only a handful of Ainu dialects available for language documentation by field linguists, on Hokkaido and Sakhalin. Their records reveal a major dialect boundary between Hokkaido and Sakhalin Ainu, and, within the Hokkaido dialects, between those spoken in the southwestern and the northeastern parts of the island.¹

This chapter describes a pattern of vowel distribution which was first described by Chiri (1952) and which has subsequently received considerable attention in the phonological literature (see, among others, Itô 1984; Mester 1986, 1988b; Ewen and van der Hulst 1988; Shibatani 1990; Vovin 1993b; Krämer 1998; Sato 2010). Part of the pattern involves a process in which a suffix vowel fully harmonizes with a preceding base vowel.

Focusing on the dialects of southwestern Hokkaido, Section 64.2 describes the pattern of vowel distribution in the relevant morphological contexts, based on Chiri (1952). Section 64.3 examines the distribution of vowels in disyllabic roots and compares this distribution with that in neighboring languages, in particular Nivkh. Section 64.4 concludes.

64.2 Distribution of vowels in derived environments

p. 817

Ainu has a five-vowel system with the canonical vowels /i, e, a, o, u/. As was first observed by Chiri (1952), these vowels are subject to co-occurrence restrictions in certain morphological contexts. Chiri investigated a total of 112 verbal and 162 nominal forms, both of which are derived by suffixation of a vowel to a CVC base. In the formation of transitive verbs, the base is an intransitive verb root that is free (1a) or bound (1b).²

- (1) a. ran-i 'lower'
 kaj-e 'break'
 mos-o 'wake up'
 jak-u 'crush'
- b. kar-i 'rotate'
 mak-a 'open'
 kom-o 'bend'
 mes-u 'tear off'

The second process involves the formation of possessive nouns. Here the vocalic suffix is added to a base which Ainu linguists call the 'conceptual form' (Kindaichi and Chiri 1936; Tamura 2000). Adding the vowel turns the conceptual form into a possessive form, with the possessor being expressed by a prefix—for example, *ku-tek-e* (1SG.POSS-hand-SFX) 'my hand'. In the absence of a prefix, the possessor signals the third-person form, as in the examples in (2).

- (2) rek-i 'beard'
 hon-i 'belly'
 jup-i 'brother'
 tek-e 'hand'
 haw-e 'voice'
 nan-u 'face'
 sar-a 'tail'
 kok-o 'son-in-law'

The forms in (1) and (2) show that the suffix variants include each of /i, e, a, o, u/. However, Chiri observes that in these processes certain combinations of base and suffix vowels are avoided. The results from his corpus of CVC roots are summarized in the following tables, which present the observed and expected frequencies of vowels in V1 (base) and V2 (suffix) position in transitive verb formation (Tables 64.1 and 64.2) and possessive noun formation (Tables 64.3 and 64.4). Expected frequencies are indicated between brackets.

Table 64.1 Observed and expected frequencies of vowels in transitive verbs

	V2 →					
V1↓	i	u	e	a	o	
i	9 (4.0)	3 (4.7)	3 (4.0)	0 (1.4)	0 (1.7)	16
u	5 (4.8)	7 (5.6)	6 (4.8)	1 (1.7)	0 (2.0)	19
e	0 (4.5)	9 (5.3)	9 (4.5)	0 (1.6)	0 (1.9)	18
a	9 (10.0)	14 (11.8)	8 (10.0)	9 (3.6)	0 (4.3)	40
o	4 (4.8)	0 (5.6)	3 (4.8)	0 (1.7)	12 (2.0)	19
	28	33	29	10	12	112

Table 64.2 Observed/expected ratios: Transitive verbs

	V2 →				
V1 ↓	i	u	e	a	o
i	2.25	0.64	0.75	0.00	0.00
u	1.04	1.25	1.25	0.59	0.00
e	0.00	1.70	2.00	0.00	0.00
a	0.90	1.19	0.80	2.50	0.00
o	0.83	0.00	0.63	0.00	6.00

Table 64.3 Observed and expected frequencies of vowels in possessive nouns

V2 →						
V1 ↓	i	u	e	a	o	
i	22 (14.1)	1 (4.1)	1 (3.4)	0 (1.2)	0 (1.2)	24
u	28 (22.3)	3 (6.6)	7 (5.4)	0 (1.9)	0 (1.9)	38
e	14 (17.0)	7 (5.0)	8 (4.1)	0 (1.4)	0 (1.4)	29
a	12 (24.0)	17 (7.1)	3 (5.8)	8 (2.0)	1 (2.0)	41
o	19 (17.6)	0 (5.2)	4 (4.3)	0 (1.5)	7 (1.5)	30
	95	28	23	8	8	162

Table 64.4 Observed/expected ratios: Possessive nouns

	V2 →				
V1 ↓	i	u	e	a	o
i	1.56	0.24	0.29	0.00	0.00
u	1.26	0.45	1.30	0.00	0.00
e	0.82	1.40	1.95	0.00	0.00
a	0.50	2.39	0.52	4.00	0.50
o	1.08	0.00	0.93	0.00	4.67

The observed frequency is the number of times a vowel combination occurs in the corpus. Consider, for example, the combination of /o/ in V1 and /o/ in V2 in Table 64.1, whose observed frequency is 12. The expected frequency is calculated by multiplying the frequency of /o/ in V1 (=19) by the frequency of /o/ in V2 (=12), and by dividing this number by the total number of disyllabic forms (19x12/112). This yields an expected frequency of 2.0; see Table 64.1. The observed frequency of /o ... o/ combinations is therefore six times as high as the expected frequency (12/2.0 = 6); see Table 64.2. ↴

p. 818

Chiri himself makes the following observations about the distribution of vowels in base-suffix contexts:

p. 819

- (3) a. /i, u, e/ are more frequent in V2 than /a, o/.
- b. There is a preference for identical vowels.
- c. /i, e/ occur with any vowel, whereas /a, u/ do not occur with /o/.

(3a) reveals a striking asymmetry in the vocalism of the suffix. Barring two exceptions (a single form with /u ... a/ and a single form with /a ... o/), /a/ and /o/ occur in V2 only when there is an identical vowel in V1. More generally, the data show that 41.1 percent of derived transitives and 29.6 percent of derived possessives have identical vowels. According to Chiri, the preference for identical vowels is the result of a diachronic change in which the suffix vowel harmonized to the base vowel.

Chiri advances two arguments in support of the view that identical vowels result from harmony. The first is that Ainu displays other cases of (historical) progressive assimilation. Compare the following forms from Sakhalin Ainu and Iburi Ainu with those from the more conservative Saru dialect (Iburi and Saru are dialects of southwestern Hokkaido). The forms in (4) and (5) are all underived and include loans from Japanese.

(4) Saru	Sakhalin	
tomari	tomara	'seaport' (Japanese <i>tomari</i>)
menoko	meneko	'woman' (Japanese <i>menoko</i>)
hoku	hoko	'husband'
tokuj	tokoj	'intimate friend' (Japanese <i>tokui</i>)
kosonte	kosonto	'wadded silk clothing' (Japanese <i>kosode</i>)

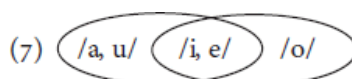
(5) Saru	Iburi	
erum	erem	'mouse'
ontaro	ontoro	tub' (Japanese <i>ōdaru</i>)
ohetu	ohotu	'to spill out'

Second, there are some forms which, according to Chiri, originate from the same morpheme but have two distinct root vowels. Both vowels condition an identical vowel in a following transitive suffix, as in (6). ↴

p. 820

- (6) pus-u 'to push out'
- pos-o 'to penetrate'

The observation in (3c) is a statement about the combinability of vowels, which can be pictorially represented as in (7).



Chiri's claim, then, is that there are three sets of vowels, /a, u/, /i, e/, and /o/, of which /i, e/ can combine with the other two sets, while /a, u/ and /o/ are mutually exclusive. However, the data show that this is not sufficiently precise. One problem is that the distribution of /i/ and /e/ is asymmetric; /o ... i/ and /o ... e/ are attested, but /i ... o/ and /e ... o/ are not. In addition, there is, as we have seen, a restriction on /a/ in V2, which almost always occurs together with an /a/ in V1.

Starting with Itô (1984), the interpretation of the restrictions in the theoretical literature has been based on the idea that disharmonic forms (i.e., forms with non-identical vowels) have a high suffix vowel which is subject to backness dissimilation, so that "when the root vowel is non-low, the suffix is [...] opposite in backness to the root vowel" (Mester 1988b: 122; for a similar characterization, see Itô 1984: 508). This accounts for the absence of /e ... i/ and /o ... u/ in transitive verbs. Note, however, that /e ... i/ does occur in possessive nouns. (This point is not mentioned by Itô and Mester, who focus only on the transitive suffix.) There are no fewer than 14 such forms in Chiri's corpus (see Table 64.3).

Chiri attributes the surface variation in the transitive and possessive suffix to phonological conditioning by the base. He suggests the following diachronic scenario for the possessive suffix. In an earlier stage of Ainu, the suffix was *-i and marked the third-person singular (e.g., *hon-i* 'his/her/its belly'). The suffixed form later became a base to which another personal prefix could be attached (e.g., *ku-honi* 'my belly').³ Other variants of the suffix derive from this *-i. For example, Kindaichi and Chiri (1936: 24) point to the existence of nouns which allow both -i and -u, as in (8).

(8)	am-i ~ am-u	'fingernails'
	kam-i ~ kam-u	'flesh'
	jam-i ~ jam-u	'chestnut'
	ham-i ~ ham-u	'leaf'
	mim-i ~ mim-u	'fish meat'
	nan-i ~ nan-u	'face'
	por-i ~ por-u	'cave'
	rar-i ~ rar-u	'eyebrows'

p. 821

It is unclear what could have motivated the emergence of the -u variant. It may be relevant that most of the roots in (8) contain /a/. Itô (1984) speculates that the variation between -i and -u in the transitive suffix reflects an earlier contrast between a front-low and a back-low vowel, which conditioned front and back realizations of the suffix, respectively; following the merger of the two low vowels, the variation between -i and -u would then have become ↯ lexicalized. Perhaps a similar development occurred in the possessive suffix. If this were the case, the choice between -i and -u would be synchronically unpredictable after /a/. This might then explain why, in some forms, speakers apparently vary in their choice of suffix vowel.

Chiri (1952: 108, 116) observes that the -e variant of the suffix frequently occurs with a preceding *j* or *w*, as in the transitive verbs in (9a) and the possessive nouns in (9b).

(9)	a.	kaj-e	'to break'
		raj-e	'to kill'
		tuj-e	'cut'
	b.	haw-e	'voice'
		maw-e	'wind'
		ruj-e	'whetstone'
		kew-e	'body'

Chiri attributes the frequent occurrence of /e/ with *j* and *w* to a constraint prohibiting sequences of *j*, *w*, and *i* (Ainu lacks underived forms with **ji* and **wi*).

The data considered above show that the form of the possessive (and transitive) suffix is not fully predictable after CVC roots. The situation is different for CV roots.⁴ The data in (10) show that after such roots the suffix vowel is identical to the base vowel if the latter is non-high. In these forms, the consonant preceding the suffix ([h], [j], or [w]) is epenthetic.

(10)

ke-he	‘oil’
re-he	‘name’
ko-ho	‘flour’
po-ho	‘child’
ka-ha	‘thread’
sa-ha	‘elder sister’

When the base vowel in a CV root is high, the suffix vowel is consistently realized as *-e*, as in the Hokkaido Ainu examples in (11). (Cognate forms in Sakhalin Ainu have epenthetic [h].)

(11)

ni-je	‘tree’
pi-je	‘seed’
ru-we, ru-je	‘trace’

Note that here, as in (9), we find *-e* occurring with a preceding glide, which in the forms in (11) is epenthetic.

64.3 Distribution of vowels in underived environments

The contexts in which we find harmony in Ainu are severely restricted; they are limited to the two morphological environments described above. As is pointed out by Shibatani (1990) ↴ and Sato (2010), other morphological processes do not show any trace of vowel harmony—for example, plural suffixation (*kom-pa* ‘bend.PL’, **kom-po*) and person-marking prefixation (*ku-komo* ‘bend.1SG’, **ko-komo*). Shibatani and Sato note that the same appears to be the case for underived forms. Many disyllabic roots violate the restrictions described above—for example, *tokap* ‘day’, *hapo* ‘mother’, *hoku* ‘husband’, *husko* ‘old’.

In order to get a better insight into the distribution of vowels in underived forms, we extracted 565 disyllabic roots from the Ainu–Japanese dictionary of the Saru dialect (Tamura 1996) and tabulated the observed and expected (O/E) frequencies of vowels in V1 and V2 position. In Table 64.5, expected frequencies are indicated between brackets.

Table 64.5 Observed and expected frequencies of vowels in disyllabic roots

	V2 →					
V1 ↓	i	u	e	a	o	
i	20 (17.1)	15 (19.2)	27 (22.5)	47 (48.7)	10 (11.6)	119
u	18 (18.1)	20 (20.3)	21 (23.7)	61 (51.5)	6 (12.3)	126
e	9 (12.2)	18 (13.7)	20 (16.1)	30 (34.8)	8 (8.3)	85
a	20 (18.2)	26 (20.5)	20 (24.1)	51 (51.9)	10 (12.4)	127
o	14 (15.5)	12 (17.4)	19 (20.5)	42 (44.2)	21 (10.5)	108
	81	91	107	231	55	565

Table 64.6 Observed/expected ratios: Disyllabic roots

	V2 →				
V1 ↓	i	u	e	a	o
i	1.17	0.78	1.20	0.97	0.86
u	0.99	0.99	0.89	1.18	0.49
e	0.74	1.31	1.24	0.86	0.96
a	1.10	1.27	0.83	0.98	0.81
o	0.90	0.69	0.93	0.95	2.00

The data show that vowels in underived forms, unlike those in base-suffix contexts, are quite evenly distributed. The O/E ratio in most of the cells in Table 64.6 is around 1.00, so that the observed frequency closely matches the expected frequency. There are, however, two vowel combinations which show a similar asymmetry to that found in base-suffix forms: /o ... o/ and /u ... o/. In disyllabic roots, /o ... o/ also has the highest O/E ratio (2.00). /u ... o/, on the other hand, has the lowest O/E ratio (0.49), followed by the reverse order, /o ... u/ (0.69). Recall that both these combinations are completely absent in derived transitive verbs and possessive nouns. The distribution of /o ... o/, /u ... o/ and /o ... u/ therefore suggests a shared property which underlies vowel combinations in both derived and non-derived forms.

p. 823

64.4 A brief areal perspective

Ainu is spoken in a region where tongue-root (TR) harmony is an areal feature, which has presumably spread through language contact (Janhunen 1981a; Comrie 1997; Ko et al. 2014). Among the geographic neighbors of Ainu, the Tungusic languages are characterized by root-controlled TR harmony (Chapter 61, this volume). Nivkh, another geographical neighbor, has remnants of what may have been TR harmony in certain prefix-base domains, and it displays synchronic height-based restrictions in disyllabic roots (Chapter 62, this volume). Based on a corpus of 291 Nivkh roots of this type, Shiraishi and Botma (2017) find that 37.1 percent have identical vowels, and that the combination /o ... o/ is particularly frequent. The data considered in Section 64.3 show much the same picture for Ainu. Interestingly, the preference for /o ... o/ is also found in languages with labial harmony, such as Mongolic (Chapter 60, this volume) and Tungusic (Chapter 61, this volume).

Another similarity between Ainu and Nivkh is that both languages have restrictions on V2. In Nivkh roots, low vowels (/e, a, o/) in V2 position cannot be preceded by a high vowel (e.g., */i ... a/, */u ... o/). This is rather similar to Chiri's observation in (3c), with the main difference being that /e/ is marginal in Nivkh (in particular in V2), but not in Ainu. The cause of these restrictions seems different, however. In Nivkh, the V2 position forms the weak syllable of a trochaic foot, which leads Shiraishi and Botma (2017) to analyze the restrictions as the result of vowel reduction. Such an analysis cannot readily be applied to Ainu, which has a pitch accent system and lacks significant vowel reduction (Shiraishi 2022b).

64.5 Conclusion

Ainu displays an intricate distribution of vowels in certain specific morphological contexts. Transitive verbs and possessive nouns with CVC roots are marked by a vocalic suffix which is fully harmonic with the preceding root vowel in some forms but disharmonic in others. The pattern of vowel restrictions in disharmonic forms resists a straightforward analysis, both synchronically and diachronically. The distribution of vowels in base-suffix contexts differs from that in underived contexts, with two interesting exceptions; in both environments there is a marked preference for /o ... o/, and a dispreference for /o ... u/ and /u ... o/. This appears to be an areal feature, since other languages in Northeast Asia display similar restrictions on the distribution of /o/.

Acknowledgments

We are grateful to the editors and to three anonymous reviewers for useful comments and suggestions.

Notes

- 1 For a sketch of the sociolinguistic situation of Ainu, see Shiraishi (2022a).
- 2 Transitive verbs are also derived by suffixation of *-ke*, *-re*, *-te*, and *-ka*. According to Chiri (1952), the vocalic suffix in (1) represents an older process.
- 3 According to Chiri (1952: 118), the synchronic reflex in the Samani dialect (northeastern Hokkaido) is the prefix *i-*, e.g. *i-sapa* 'his/her/its head', *i-pon* 'he/she/it is small'.
- 4 Recall that the data in Tables 64.1–64.4 are limited to base-suffix combinations with CVC roots.