Chapter 32

Vowel co-occurrence restriction in Ainu

Hidetoshi Shiraishi
Sapporo Gakuin University

1 Introduction

Ainu (language isolate, Japan and Russia) exhibits a co-occurrence restriction on vowels in verbal and nominal derivations. This restriction was first reported in Chiri (1952) as a type of vowel harmony. Later, it was taken up by Ito (1984) and became one of the most discussed topics in Ainu phonology to date (e.g. Mester 1986; Ewen & van der Hulst 1988; Shibatani 1990; Krämer 1998; Sato 2010).

In this paper I will give an outline of the phenomenon and review the discussions made by these authors. While Ito (1984) assumes the co-occurrence restriction as being both assimilatory and dissimilatory (disharmonic) in nature, I will point out that there is a flaw in Ito's description of the phenomenon. This flaw invalidates most of her analysis and those defended in the above literature, which depend on her description. Instead, I will propose an alternative view in which the assimilatory nature of the co-occurrence restriction plays a central role, as originally suggested by Chiri in his first report of the phenomenon in 1952.

2 Co-occurrence restriction between vowels in disyllabic derivations

The description of the co-occurrence restriction given below is based on Chiri (1952). Ainu exhibits V1–V2 co-occurrence restriction in two derivational contexts: transitive verb formation and possessive noun formation. In transitive verb formation, the base to which a V-suffix attaches is either an intransitive verb (free morpheme) or a root (bound morpheme) (the distinction between bound and free morphemes is irrelevant to the discussion here).

		Intransitive verb or root		Transitive verb	
	a.	*kay*	'to be broken'	*kay-e*	'to break'
	b.	*mos*	'be awake'	*mos-o*	'to wake up'
	c.	*yak*	'to be crushed'	*yak-u*	'to crush'
(1)	d.	*ran*	'to go down'	*ran-i*	'to lower'
	e.	*√mak*	'open'	*mak-a*	'to open'
	f.	*√kom*	'bent'	*kom-o*	'to bend'
	g.	*√kar*	'spinning'	*kar-i*	'to rotate'
	h.	*√mes*	'to come off'	*mes-u*	'to tear off'

In possessive noun formation, the V-suffix attaches to a base which is often referred to as the *conceptual form* in Ainu literature (Kindaichi & Chiri 1936; Tamura 2000). Through V-suffix affixation, the conceptual form changes to a possessive form in which the possessor is overtly expressed.

		Conceptual form		Possessive form	
	a.	*sa*	'sister'	*sa-ha*[1]	'one's sister'
	b.	*re*	'name'	*re-he*	'one's name'
	c.	*nan*	'face'	*nan-u*	'one's face'
(2)	d.	*tek*	'hand'	*tek-e*	'one's hand'
	e.	*haw*	'voice'	*haw-e*	'one's voice'
	f.	*rek*	'beard'	*rek-i*	'one's beard'
	g.	*hon*	'belly'	*hon-i*	'one's belly'
	h.	*yup*	'brother'	*yup-i*	'one's brother'

The distribution of V1 and V2 is illustrated in the tables below. Observed frequencies (= word frequency) are from Chiri (1952). Expected frequencies (in brackets) and ratios (Tables 2 and 4) are added by the current author.

Table 1: Observed and expected frequencies (transitive verb).

V1 \ V2	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	28	10	12	112

[1] The [h] before V-suffix is epenthetic.

Table 2: Observed/Expected ratios (transitive verb).

V1 \ V2	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 3: Observed and expected frequencies (possessive noun).

V1 \ V2	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 4: Observed/Expected ratios (possessive noun).

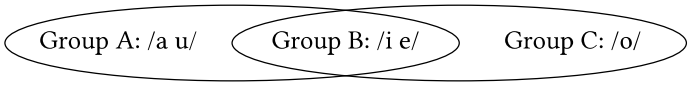
V1 \ V2	i	u	e	a	o
i	*1.56*	0.24	0.29	0.00	0.00
u	*1.26*	0.45	*1.30*	0.00	0.00
e	*0.82*	*1.40*	*1.95*	0.00	0.00
a	0.50	*2.39*	*0.52*	**4.00**	0.50
o	*1.08*	0.00	*0.93*	0.00	**4.67**

3 Review of literature

3.1 Chiri (1952)

Comparing Tables 1 and 3, Chiri (1952: 220) noticed that the distribution of V1–V2 is surprisingly similar in verbal and nominal derivations. Notably, he observed the following characteristics as being common to both nominal and verbal derivations. Firstly, /i u e/ outnumber /a o/ to a considerable degree in V2. Secondly, there is a preference for lining up identical vowels. This is especially the case with V2 /a/ and /o/, in which case V1 is nearly always /a/ and /o/, respectively. This preference is slightly weaker in V2 /e/, as non-identical V1–V2 sequences such as *u–e*, *a–e* or *o–e* are observed. In such cases, C2 is usually /y/ or /w/ (*ruy-e* 'shake', *haw-e* 'one's voice', *poy-e* 'to mix'), a regularity to which we return later. Finally, /a u/ and /o/ do not co-occur, whereas /i e/ can co-occur with any vowel. This final observation led Chiri to propose the following grouping of vowels.

(3)



In terms of vowel harmony, Group B is neutral since these vowels co-occur with any other vowel (Shibatani 1990: 14).

While (3) successfully accounts for the distribution of vowels illustrated in the tables above, it is not yet sufficient grounds to call the phenomenon vowel harmony, as Shibatani (1990) and Sato (2010) point out. Firstly, the classification in (3) cannot be characterized by a vowel feature with a phonetic basis known to trigger vowel harmony cross-linguistically, such as palatality (backness) or tongue-root position (Shibatani 1990: 15; Sato 2010: 164–165). Thus if this were a case of vowel harmony, it should have lost its transparency in the course of history. Nevertheless, such an assumption is missing in Chiri's diachronic scenario, as we will see below. Secondly, this co-occurrence restriction applies to two morphological contexts only, namely transitive verb and possessive noun formation. It is not observed in other morphological contexts such as plural suffix (*kom-pa, *kom-po*) or personal prefix (*ku-komo, *ko-komo*) formation (Shibatani 1990: 15–16; Sato 2010: 163–166). There are no alternations which make vowel sequences in these morphological concatenations harmonic. This is to say that (3) is irrelevant to most morphological operations in Ainu.

Aside from the discussion of whether or not the restriction observed is vowel harmony, Chiri's insightful observations hint at the phenomenon's historical origin. He proposed the following diachronic scenario for possessive noun formation: at a previous historical stage, the V-suffix was /i/, a third person singular affix. Subsequently, this /i/ underwent a root-control type of harmony imposed by a preceding vowel. This led to the proliferation of V1–V2 forms with identical vowels.

According to Chiri, this scenario finds the following support: firstly, there is independent evidence for the use of /i/ as a third person singular affix. For instance in the Samani dialect, it surfaces as a prefix: *i-sapa* 'his head', *i-tanehe* 'its seed', *i-an* 'it exists', *i-pon* 'it is small'. Secondly, Ainu exhibits cases of progressive assimilation, e.g.

hoku > *hoko* 'husband' (Sakhalin dialect), *erum* > *erem* 'mouse' (Iburi dialect). Here, the direction of assimilation is identical to that in root-control harmony. Finally, it conforms to a phonotactic restriction that defines **yi* and **wi* as illicit CV sequences. In Ainu, **yi* and **wi* are excluded from underived contexts, and only marginally observed across morpheme boundaries (Kindaichi 1931: 13; Tamura 2000: 23; Shiraishi forthcoming). According to Chiri, V2 /i/ underwent lowering to /e/ when preceded by /y/ or /w/, in order to avoid the surfacing of **yi* or **wi*. The lowering of /i/ to /e/ to create /ye/ and /we/ is therefore phonologically sound. This explains the high correlation between V2 /e/ and C2 /y/, /w/.

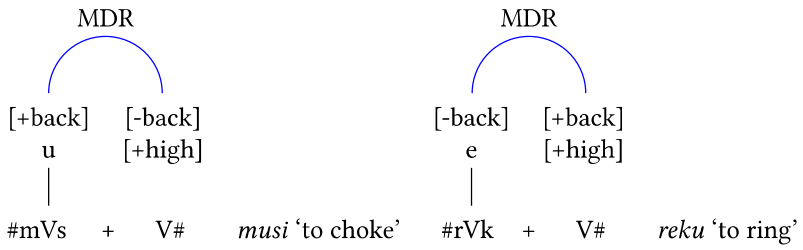
We evaluate this diachronic hypothesis in Sections 3.3 and 4 below.

3.2 Ito (1984)

Ito's primary interest lies in the pattern of disharmony, which she formalizes as a rule of dissimilation called the *Melodic Dissimilation Rule* (MDR). In Ito's analysis, MDR is operative in a subset of transitive verb formation outputs. She subcategorizes transitive verb formation into three groups: 1) a total assimilation group (*kom-o, mak-a*), 2) *a*-roots, which have /a/ as V1 and a high vowel as V2 (*kar-i, ram-u*), and 3) a disharmonic group, which has a non-low V1 and a high vowel with the opposite backness value in V1 to V2. Hence if V1 were [+back], then V2 would be [-back], and vice versa (*hum-i* 'to chop up', *pok-i* 'to lower', *pir-u* 'to wipe', *ket-u* 'to rub').

Since Ito works in the framework of autosegmental phonology, she represents the total copying (assimilation) of vowel features as the spreading of an autosegment. As for *a*-roots, Ito gives up phonological analysis and proposes to encode the backness value in the lexicon. The interesting case is 3), in which MDR is operative. MDR yields the correct output with the opposite backness value for V1 and V2 with the proviso that V2 is assigned [+high].

(4) Melodic Dissimilation Rule (MDR)
a. [+high] → [-α back] / [α back] _



Through MDR, V1–V2 sequences such as **e-i* and **o-u*, in which the vowels share [back] values, are avoided.

Ito's analysis provided grounds for discussion to other authors, who attempted to provide a formal account of disharmony. Further materials for discussion were provided by patterns of disharmony in Tzeltal (Mayan) and Ngbaka (Central African).

Whatever Ito's contribution to the discussion of disharmony, her analysis cannot be readily accepted, as her description of the phenomenon has several flaws. Firstly, Ito declares at the outset that she will illustrate her case using data from transitive verb formation only, but this choice has seemingly led to the exclusion of possessive noun formation from consideration altogether. In fact, possessive nouns contain cases of *e–i*, which is predicted not to occur by MDR. Table 3 counts fourteen such cases, which is too many to eliminate as exceptions. This fact was either overlooked or neglected by Ito. It should therefore be pointed out that MDR is successful for only a small subset of verbs (thirteen in total). Secondly, Ito argues that there is independent support for MDR in the inventory of diphthongs. According to Ito (1984: 509–510), Ainu diphthongs consist of those which obey MDR (*iw, ew, uy, oy, aw, ay*) while those disobeying MDR do not exist (*ey, ow*). However, this observation is incorrect; Ainu has *ey* in underived forms: e.g. *sey* 'bivalve' or *teyne* 'to be wet'.

There are also shortcomings in Ito's proposed analysis. For instance, it is not clear how the three subcategorized groups interact with each other. Since Ito provides three different mechanisms to account for each group, her analysis gives the impression that there are three independent mechanisms involved (total assimilation, lexical encoding and MDR) although they apply to the same morphological context. Chiri's remarkable insight that both verbal and nominal derivations provide common grounds for V1–V2 distribution is thereby lost.

3.3 Sato (2010)

Sato (2010) faithfully follows Chiri's description of the phenomenon and focuses on the prevailing pattern of identical V1–V2.[2] Like Chiri, Sato (2010: 172) assumes it to occur as a consequence of root-control harmony, thus V1 affecting V2, and proposes to add the following step to Chiri's diachronic scenario. In a similar fashion to Chiri's reconstruction of possessive noun formation with a common V-suffix /i/, Sato proposes to reconstruct transitive verb formation with a common V-suffix /ə/. This /ə/ underwent root-control harmony (= total assimilation) (*kom-ə* > *komo*, *yas-ə* > *yasa*, *yup-ə* > *yupu*). Deviations from total assimilation were /ə/ > /e/ after /y/ (and possibly /w/) (*noy-ə* > *noye*, **noyo*), and V1 /ə/ > /e/ and V2 /ə/ > /u/ after C2 /s/ (*məs-ə* > *mesə* > *mesu*).[3]

Sato's hypothesis reinforces Chiri's diachronic scenario by assuming a common vowel suffix /ə/ for transitive verb formation. He thereby succeeds in providing a common phonological context, root-control harmony, for both nominal and verbal derivations.

Nevertheless, Sato's hypothesis is not as convincing as Chiri's as he fails to provide a source of the reconstructed /ə/. Unlike Chiri's postulation of /i/ in possessive noun formation, /ə/ is not associated to any known morphological entity nor to any phonological process which might have produced such a vowel. It is this latter point to which we turn in the next section.

[2] I would like to thank Yasushige Takahashi (Hokkaido University) for bringing this work to my attention.

[3] Sato does not provide explanations of these irregular processes.

4 The nature of root-control harmony

An interesting case of V1–V2 distribution is found in Nivkh (language isolate, Russia), a geographic neighbor of Ainu. Although the vowel inventory is not entirely identical to that of Ainu (Nivkh has a high central vowel /ɨ/), vowel distribution in Nivkh disyllabic roots exhibits some common characteristics. Table 5 illustrates the observed and expected frequencies of V1–V2 in Nivkh disyllabic roots, and Table 6 the observed/expected ratios (Shiraishi and Botma 2015, data from Pukhta 2002).

Table 5: Observed and expected frequencies (Nivkh).

V1 \ V2	i	ɨ	u	e	a	o	
i	12 (11.1)	14 (3.7)	5 (4.5)	0 (0.8)	1 (8.4)	0 (3.5)	32
ɨ	24 (17.0)	15 (5.7)	9 (6.9)	1 (1.1)	0 (12.8)	0 (5.4)	49
u	21 (14.6)	4 (4.9)	13 (5.9)	0 (1.0)	4 (11.0)	0 (4.6)	42
e	10 (10.4)	0 (3.5)	1 (4.2)	1 (0.7)	17 (7.8)	1 (3.3)	30
a	21 (28.1)	1 (9.5)	8 (11.4)	3 (1.9)	42 (21.1)	6 (8.9)	81
o	13 (19.8)	0 (6.7)	5 (8.0)	2 (1.4)	12 (14.9)	25 (6.3)	57
	101	34	41	7	76	32	291

Table 6: Observed and expected frequencies (Nivkh).

V1 \ V2	i	ɨ	u	e	a	o	
i	*1.09*	**3.50**	*1.00*	0.00	0.13	0.00	32
ɨ	*1.41*	**2.50**	*1.29*	*1.00*	0.00	0.00	49
u	*1.40*	*0.80*	**2.17**	0.00	0.36	0.00	42
e	*1.00*	0.00	0.25	*1.00*	**2.13**	0.33	30
a	*0.75*	0.11	*0.73*	*1.50*	**2.00**	*0.67*	81
o	*0.65*	0.00	*0.63*	**2.00**	*0.80*	**4.17**	57
	101	34	41	7	76	32	291

Like Ainu, the high O/E ratio in the diagonal cells from upper left to lower right (Table 6) shows the tendency to line up identical vowels. In Ainu, 41.1% of the verbs have identical V1–V2. In nouns it is 29.6%. In Nivkh, it is 37.1%.

In addition, there are the following asymmetries in the distribution of vowels. Firstly, a high vowel–/a/ sequence hardly ever occurs (**i–a, *u–a, *ɨ–a*). Secondly, V2 /o/ exhibits a strong tendency to line up with V1 /o/. In contrast, there seems to be no such restriction on V2 /i u/. Interestingly, these restrictions concern mainly V2 and resemble Chiri's observation on Ainu seen earlier.

To sum up, both Nivkh and Ainu exhibit 1) a preference for lining up identical

vowels (the strongest being /o/–/o/), and 2) restriction on V2, namely, a high (front) vowel (*i, e*) is common whereas back non-high vowel (*o*) is not, unless preceded by an identical vowel. According to Shiraishi & Botma (2015), these two characteristics stem from a single phonological process in Nivkh: stress-dependent height harmony. The stress pattern of Nivkh is trochaic in polysyllabic roots – V1 stressed and V2 unstressed. Stress is realized as high pitch and increased duration. Shiraishi and Botma claim that this prosodic asymmetry affects the distribution of vowels. Being unstressed, V2 is a prosodically weak licensor and therefore able to host only two types of vowels:

1. Vowels which are supported sufficiently by the prosodically strong V1 by means of total or partial assimilation. Total assimilation is the strongest in /o/–/o/.[4] Partial assimilation comprises a case of height harmony, in which a non-high V2 is allowed only with a non-high V1 (*a–e, o–e, e–a, o–a*), but not with a high V1. High–non-high sequences (**i–a, *u–a, *ɨ–a*) are excluded since a high V1 fails to license a non-high V2, having no height feature in common.

2. The intrinsically short vowels /i u/. Being *small*, these vowels do not interfere with the prosodic weakness of V2 and conform to the canonical trochaic stress pattern in Nivkh.[5] This is a case of unstressed vowel reduction, in which weak prosodic positions are "more susceptible to co-articulatory effects from neighboring strong vowels" (Barnes 2006: 193). It is this mechanism that Shiraishi and Botma assume to underlie the asymmetric distribution of vowels in Nivkh.

The question is whether a similar mechanism of stress-dependent harmony can be assumed in the asymmetric distribution of vowels in Ainu. One critical difference from Nivkh is that Ainu has iambic as a dominant stress pattern in underived contexts. In the derived disyllabic forms discussed above, stress depends on morphological boundaries: CVC-V is iambic (*ka'y-e*) while CV-CV is trochaic (*'sa-ha, 're-he*). The vast majority of cases are iambic as there are only 25 CV-CV forms. Thus if we were to assume a stress-dependent root-control type of harmony, we are forced to assume that Ainu originally had a trochaic stress pattern which later became iambic at some point in its history. Is such an assumption justifiable?

Possible support for this idea comes from a correlation between trochaic stress patterns and morphological complexity. Ainu is a typical iamb language; in simplex forms, stress falls on the second syllable unless the first is heavy (CVC): *sa'pa* 'head', *'sinrit* 'ancestors'. The interesting case is the morphologically complex forms. According to Sato (2015; forthcoming), the stress pattern in morphologically complex forms is largely dependent on morphological boundaries. When the first element of a complex is CV, trochee is dominant. Sato counts 58 trochees and 6 iambs in Tamura's Ainu dictionary (1996).

[4] This is possibly a remnant of labial harmony prevailing in the area, as Tungusic and Mongolic languages exhibit a similar pattern (Li 1996).

[5] Like /i u/, /i/ has a short duration. However, it is not small, as it is not a focalized vowel (Harris 2005). See Shiraishi & Botma (2015) for further discussion.

(5) a. 're-kor
name-have
'to have a name'

b. 'e-re
eat-CAU
'let someone eat'

c. ni-'esisuye
stick-swing
'to swing a stick'

When the first element is CVC, the first segment in the second element plays a decisive role. If it is C, the complex is exclusively trochaic (*'tek-moymoke* 'to move hands'). If it is V, iambic is dominant but trochees are also observed. Sato counts 109 iambs and 29 trochees in Tamura's dictionary.

(6) a. sik-'erayke
eye-kill
'to glare at'

b. am-'us-pe
claw-attach-thing
'crab'

c. 'cip-o
boat-get_on
'to get on a boat'

d. 'mim-us
meat-stick
'to be fat'

Thus while iamb is dominant in simplex forms, trochee is dominant in complex forms. Sato (forthcoming) infers from this fact that trochee is the basic and the archaic stress pattern in complex forms. Iambs (6a, b) are innovative, possibly developed under the influence of simplex forms.

Typically, an iamb in a complex form undergoes resyllabification from CVC.V to CV.CV (*si.'ke.ray.ke, a.'mus.pe*), thereby conforming to the canonical iambic stress pattern in simplex forms (Sato forthcoming). If the iamb-trochee asymmetry were associated with morphological complexity, then the hypothetical historical shift of stress from V1 to V2 that we are assuming above could be associated with a shift in morphological complexity – the shift signals a change from morphologically complex to simplex in the mind of language users. It is this shift that could have led to a mismatch between the stress pattern and the distribution of vowels that we are witnessing now.

References

Barnes, Jonathan. 2006. *Strength and weakness at the interface: positional neutralization in phonetics and phonology.* Berlin/New York: Mouton de Gruyter.

Chiri, Mashiho. 1952. Ainugo ni okeru boinchowa [Vowel harmony in Ainu]. *Bulletin of the Hokkaido University Faculty of Arts* 1. 101–118.

Ewen, Colin & Harry van der Hulst. 1988. [high], [low] and [back] or [I], [A] and [U]? In Peter Coopmans & Aafke Hulk (eds.), *Linguistics in the Netherlands 1988*, 49–57. Dordrecht: Foris.

Harris, John. 2005. Vowel reduction as information loss. In Philip Carr, Jacques Durand & Colin Ewen (eds.), *Headhood, elements, specification and contrastivity*, 119–132. Amsterdam: John Benjamins.

Ito, Junko. 1984. Melodic dissimilation in Ainu. *Linguistic Inquiry* 15 (3). 505–513.

Kindaichi, Kyosuke. 1931. Ainu yūkara gohō tekiyō [An outline grammar of the Ainu epic poetry]. In *Ainu jojishi yukara no kenkyu*, 1–233. Tokyo: Tōyō Bunko.

Kindaichi, Kyosuke & Mashiho Chiri. 1936. Ainu gohō gaisetu [An outline of Ainu grammar]. In *Chiri mashiho chosakushū [Writings of Chiri Mashiho]*, vol. 4, 3–197. Tokyo: Heibonsha.

Krämer, Martin. 1998. A correspondence approach to vowel harmony and disharmony. *Working papers Theorie des Lexicons* 107.

Li, Bing. 1996. *Tungusic vowel harmony: description and analysis.* The Hague: University of Amsterdam PhD thesis.

Mester, Armin. 1986. *Studies in tier structure.* Amherst: University of Massachusetts PhD thesis.

Pukhta, Maria. 2002. *Nivkh-Russian conversation and daily-life thesaurus* (Endangered languages of the Pacific Rim A2-012). Osaka Gakuin University.

Sato, Tomomi. 2010. Chiri hakase to Ainugogaku [Doctor Chiri and Ainuology]. In Hironobu Kosaka (ed.), *Chiri Mashiho*, 159–173. Sapporo: Crews.

Sato, Tomomi. 2015. Ainugo no goseigo no akusento kisoku to sono reigai nit suite [Compound accent rules and exceptions in Ainu]. In Anna Bugaeva & Iku Nagasaki (eds.), *Ainugo kenkyu no shomondai*, 1–13. Sapporo: Hokkaido shuppan kikaku center.

Sato, Tomomi. Forthcoming. Ainugo no goseigo ni okeru reigaiteki akusento to sono rekishiteki kaishaku [Exceptions to compound accent rule and its historical implications in Ainu]. In Yeonju Lee (ed.), *Akusento ronshu.* Sapporo: Hokkaido University.

Shibatani, Masayoshi. 1990. *The languages of Japan.* Cambridge: Cambridge University Press.

Shiraishi, Hidetoshi. Forthcoming. Phonetics and phonology. In Anna Bugaeva (ed.), *The handbook of the Ainu language.* Berlin: Mouton.

Shiraishi, Hidetoshi & Bert Botma. 2015. *Stress-dependent harmony in Nivkh.* Presentation at the 23rd Manchester Phonology Meeting.

Tamura, Suzuko. 1996. *Ainu dictionary (Saru dialect).* Tokyo: Sofukan.

Tamura, Suzuko. 2000. *The Ainu language.* Tokyo: Sanseido.