Austronesians in the Northern Waters?

Alexander Vovin*

CRLAO/EHESS, Paris, France

sashavovin@gmail.com

Abstract

The following lines are inspired by John Kupchik's seminal article 'Austronesian Lights the Way' that appears in this volume of JEAL. It demonstrated beyond any reasonable doubt for the first time that there are reliable Austronesian loanwords in Japonic that reveal quite ancient and profound contacts, because without these profound contacts the borrowing of the names of the most basic celestial bodies, such as the sun and the moon, would not be possible. In my opinion, his article opened a new and an exciting direction in the Japonic historical linguistics.

There are, however, two important differences between Kupchik's article and the present one. First, while Kupchik mostly concentrated on the Amis language from Taiwan, and to a less extent on the languages of Philippines and other Western Malayo-Polynesian, my major focus is on the Philippines languages as potential donors, and much less on other Austronesian languages of the region. Second, while Kupchik looked mostly on mysterious words in the *Omoro Sōshi*, a collection of Old Okinawan and Amami sacred and folk poems (1531–1623 AD), this article focuses more on Old Japanese in particular and Japonic in general.

Keywords

Old Japanese – Austronesian – etymology – language contact

*

This article is part of an ongoing project entitled Etymological Dictionary of Japonic Languages. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 788812).

1 Introduction

Some general remarks are in order. Overall, I agree with Lyle Campbell, who always said that “languages should speak for themselves”. This means that other humanities data, such as archeology, history, religious studies, etc., as well as genetics from the natural sciences, which is in great vogue nowadays, cannot tell us anything either in support or criticism of the purely linguistic data. Yet, sometimes these data may be interesting and worth looking at, as they can potentially point into a right direction.

Therefore, before diving into historical linguistics and philology, let me offer some preliminary extralinguistic observations.

First, the distance between Yonaguni, the last island in the Ryukyuan archipelago, and the coast of Taiwan is mere 108 km. The closest landfall in Taiwan also appears to be in the region where the Amis people live. Among the Taiwan aboriginal tribes, only Amis and Yami are true seafarers, with Amis being famous for their long boats well suited for distant ocean voyages. From this point of view, it is incomprehensible why Austronesians would sail only south, and not north, especially that in clear weather one can see Yonaguni from Taiwan and vice versa. We, of course, know nowadays that there was Shimotabaru culture in South Ryukyus between 2900–2000 BC, but the archeological evidence supporting its connection to a culture found in North-Eastern Taiwan at the same time is inconclusive, and in addition there is a hiatus in population found between 2000 BC and 900 BC (Pearson 2013: 71–81). In any case even 900 BC is too early for Japonic speakers to be present in Sakishima, as they would not be able to reach there before 1200 AD.

2 Austronesian Vocabulary in Old Japanese *Makura-Kotoba*

One of the natural sources for searching Austronesian vocabulary in insular Japonic are those permanent epithets in the Old Japanese poetry that are opaque. These permanent epithets are known in Japanese as ‘pillow words’ (*makura-kotoba* 枕詞). It must be mentioned that the majority of them are quite transparent, and can be clearly explained on the basis of Japonic itself. Then there are two other groups, a larger one based on Old Korean, and a smaller one that has Ainu origin. Finally, there is one more group that cannot be analyzed on the basis of any these three languages. I will try to show below that at least some of those, or parts thereof can have Austronesian etymologies.

One of the most spectacular cases is the permanent epithet *taratine n-a* that is used with the OJ word *papa* ‘mother’, but never with its doublet *omā ~ amo*

'id.' that is likely to be a loan from Old Korean (cf. МК *ómà* ~ *ómó* 'id.'). OJ *n-ə* in *taratine n-ə* can be easily explained as an attributive form *n-ə* of the copula *n-*, so our task boils down to determining the etymology of *taratine*. The first two syllables *tara* are strikingly reminiscent of PAN *daRa^q 'blood' (Blust and Trussel 2020).² The reflexes of PAN *daRa^q 'blood' are all over the Austronesian family, being found in Formosan languages, WMP, CMP, SHWNG, and OC, but the phonetically closest forms to OJ *tara* seem to be: Formosan: Seediq *dala*, WMP: Ilokano *dára*, Bontok *dála*, Ifugaw *dála*, Batad Ifugaw *dāla*, Pangasinan *dalá*, Tiruray *dara?*, Basap *dara?*, Bulungan *dara*, Lun Dayeh *dara?*, Tabun *dara*, Kelabit *dara?*, Iban *darah*, Maloh *dara?*, Cham *darah*, Moken *dalak*, Malay *darah*, Acehnese *darah*, Simalur *dala*, Madurese *dhara*, Buginese *dara*, CMP: Manggarai *dara* 'blood', Rembong *dara?*, Riung *dara*, Dhara/Ndao *dara*, Helong *dala*, OC: Arop *dara*, Kis *dara*, Wogeo *dara*, Manam *dara*, Takia *dala-*, Tawala *tala-*, Cheke Holo *da-dara*, Kokota *da-dara-na*.

Blust further notes that Itbayat *ka-ɾaya* means 'cosanguine', 'blood kin', and Rejang *dalea? dagiŋ* (*dalea?* 'blood') means 'kinsmen', 'blood relatives' (2020). These meanings are important for the discussion that will shortly follow below.

With such a distribution of Austronesian words for blood resembling OJ *tara*, it is practically impossible to say what group of the Austronesian languages could be a possible source. But we can considerably narrow it down by determining the place of origin of OJ *tine* in *tara tine*. It is well known that there is no contrast between OJ /e/ < PJ *ia and /ɛ/ < PJ *ay after coronals. But OJ /e/ is extremely rare, therefore our best bet is that we are dealing here with OJ *ne < PJ *nay. Thus, I reconstruct PJ *tinay. *inay is a form of address of PAN *ina 'mother' found in WMP: Tagalog *iná-y*, Bikol *iná-y*, Binukid *ina-y*, Tatana *ina-y*, Kenya *ina-y*, Kayang *ina-y*, Dusun Malang *ina-i* (Blust and Trussel 2020). The prefix *t- is honorific (Wolff 2010.2: 842) or referential/respectful (Blust and Trussel 2020) This prefix *t- in combination with the following *inay or *ina is attested in Formosan: Kavalan *tina*, Bunun *tina*, Rukai *t-ina*,³ WMP: Lun Dayeh *tina-n*, Miri *tinah*, Kiput *tinah*, Bintulu *tina*, Melanau *tina*, Siang *tina-i*, Malay *be-tina*, Banggai *tina*, Bare'e *tina*, Bungku *tina*, Moronene *tina*. The *tina* (but not *tinay*) forms are also widespread in South Halmahera – Western New Guinea and Oceanic. It seems though that *dara?* 'blood' and *ina-y* 'mother' overlap only in some Philippine languages and Malay, and more rarely in Borneo (Kalimantan). On the other hand, the form *t-ina-y* if we exclude more

1 Blust's reconstruction (2020). Wolff proposes *dayáq (2010.2: 833–834). I normally follow Blust if available, since his reconstructions appear to be more reliable.

2 One should keep in mind what in OJ there are no contrasts /t/ : /d/, /r/ : /l/ and no syllable-final or word final consonants.

3 Blust traces Formosan forms back to PAN *ta-ina (2020).

Eastern branches, is limited to Formosan and languages of Borneo in WMP, but these languages do not have *dara*? ‘blood’. This distribution, I believe, probably points to the Northern Philippines as a place of origin, which lacks honorific *t-*, but has all other features attested.

To the best of my knowledge, among all OJ texts, *tara tine* is found only in the *Man'yōshū*, the earliest and the largest Japanese poetic anthology. It is quite clear that in OJ we deal with a very old prehistoric loan, because the meaning of the permanent epithet *tara tine* was already lost in the eighth century AD. Some poems engage in a folk etymological analysis, spelling *tara ti* as 垂乳 ‘hanging breasts’ (MYS 9.1774, MYS 11.2368, MYS 11.2557, MYS 12.2991, MYS 13.3314), but such an interpretation is obviously ungrammatical, because with the vowel verb *tare-* ‘to hang down’ we would expect *tar-uru ti* ‘hang-down-ATTR breast’ or *tare ti* ‘hang.down(NML) breast’, but under no circumstances **tara ti* would be possible.

I believe that OJ *tara tine* means ‘one’s own birth mother’. It looks like it was an important distinction to make, because in all likelihood, other wives and possibly even concubines of one’s father were also referred to and/or addressed as ‘mother’. As far as I am aware, there is no textual evidence to support this claim, but there is strong linguistic evidence supporting it. Blood lines in Ancient Japan were extremely important, and the matrilineal descent was as important as the patrilineal, with the former replaced by the latter only after the Heian period (794–1192 AD). Originally, the maternal line seems to be more important than the paternal one. In the *Nihonshoki* and other WOJ and MJ texts we find the following compounds, consisting of *irə* + kinship term, which designate kinship relationship by blood:

WOJ *irə-pa* ‘one’s birth mother’, attested in the *Nihonshoki* exclusively in *kana* glosses that cannot be earlier than the Heian period: イロハ (NS 1.11⁴), (NS 1.12), (NS 1.13), etc. There is also one more attestation *irə-pa* (伊呂波) in the *man'yōgana* in WMS (2.14b), which is much more useful than the *kana* glosses, in spite of the fact that WMS is from the early Heian period, but in this case we apparently have a case of an orthographic conservatism. It is not clear what *irə* in this and other kinship terms below means. Apparently it is a different word

4 When citing the *Kojiki* or the *Nihonshoki*, Japanese scholars normally provide very vague textual addresses that include only a reference to the number of a book. This is obviously unacceptable in modern-day philology. When citing the *Nihonshoki*, I organize these addresses adding also a page number according to the modern edition in the *Newly Edited and Enlarged Series on the Japanese History* (新訂増補國史大系) by Kuroita Katsumi (黒板勝美) et al. (1974). Textual addresses for the *Kojiki* are compiled on the basis of Hisamatsu Sen'ichi (久松潜市) et al. 1958. 『古事記総索引』 *A General Index to the Kojiki*, vol. 1–2 in the 『古事記大成』 *The Kojiki Series*, vol. 7–8. They include volume number, leaf number and *verso* or *recto*, indicated as *a* or *b*.

from a homophonous OJ *irə* ‘color’, ‘appearance’. Its meaning is opaque, but it clearly refers only to one’s maternal relatives, as the combination **irə-ti* ‘*irə* + father’ simply does not exist.

WOJ *irə-ye*⁵ (伊呂兄) ‘elder brother from the same mother’ is attested in the *Kojiki* in a partially phonographic script (KJK 2.1a). Omodaka et al. do not provide this example, apparently believing that all examples come only in the logographic script as 兄 ‘elder brother’ and only from the *Nihonshoki* (1967: 107).

WOJ *irə-se* (伊呂勢) ‘brother from the same mother’ (Omodaka et al. 1967: 107), although, attested in the *man’yōgana* script (萬葉仮名) in the *Kojiki* (KJK 1.23a), and possibly in a gloss for the character 兄 in the *Nihonshoki*, although there is also an alternative gloss *irə-ye* for the same character. OJ *se* refers exclusively to an ‘elder brother’, ‘male beloved’, or ‘husband’, and never to a ‘younger brother’. Therefore, the motivation for Omodaka et al. to define this combination as ‘brother from the same mother’ rather than ‘elder brother from the same mother’ remains unclear. The examples they cite support only ‘elder brother’. Cf. also WOI *irə-ⁿdə ~ irə-tə* ‘younger brother or sister from the same mother’ below.

WOJ *irə-ⁿdə ~ irə-tə* (伊呂杼) ‘younger brother or sister from the same mother’ is attested in a completely phonographic spelling in KJK 2.13b and 2.15b. The meaning ‘younger sister’ might seem strange, but it is confirmed beyond any doubt by example (1) below:

(1) KJK 2.13a–13b

有二女。兄名蠅伊呂泥 弟名蠅伊呂杼也。

There were two women. The name of the elder sister was Fly the Elder Sister The name of the younger sister was Fly the Younger Sister.

WOJ *irə-ne* (伊呂泥) ‘elder brother or sister from the same mother’ (Omodaka et al. 1967: 108). This time Omodaka et al. seem to be right, since this combination is applicable to both sexes. The completely phonographic spelling is attested in KJK 2.13a–13b, see the example above.

WOJ *irə-mo* (伊呂妹) ‘younger sister from the same mother’ is attested only in partially phonographic spelling in KJK 1.42b, 2.28b, 2.31a, and 3.17b. In all these cases we have kana gloss 𐰇 *mo* to the right of the character 妹 ‘younger sister’.

The etymology of WOI *irə* is obscure. To the best of my knowledge, it is not attested in any other variety of Japonic except WOI. Such a limited distribution

⁵ Capitals indicate the logographic spelling.

is highly suggestive of a borrowing. But what language did this loan come from? Ainu, Koreanic and Austronesian have no word that could fill in this niche. Quite possibly, this might be a borrowing from some Jōmon period language. But certainly this hypothesis is impossible to prove or to disprove, so most likely it will remain just a speculation. There is a tiny chance that WOJ *irə-* may also be from Austronesian, cf. WMP Paku *ira?* 'blood' (Blust and Trussel 2020). But besides an obvious problem with a vocalism in the second syllable, Paku *ira?* (a reflex of PAN *daRaQ 'blood') is isolated in WMP. And finally, seeing an Austronesian behind every bush is going to be more harmful than helpful. However, notwithstanding its source language, most likely *irə* means 'blood', so *irə-pa* is probably the construction identical to *tara tine*.

There are altogether twenty-four examples of *tara tine* in the *Man'yōshū* poems⁶ not including those that appear in the Chinese prefaces, like the one to MYS 14.3550a. See exx. (2–5) below:

(2) MYS 15.3688 (WOJ)⁷

(9) 安吉佐良婆 (10) 可敝里麻左牟等 (11) 多良知祢能 (12) 波々尔麻乎之弓

(9) aki sar-a^{mba} (10) kaper-i-[i]mas-am-u tə (11) **tara tine** n-ə (12) papa-ni mawos-i-te

(9) autumn come-COND (10) return-CONV-HON-TENT-FIN QV (11) **blood mother** COP-ATTR (12) mother-DAT tell(HUM)-CONV-SUB

(12) [He] told [his] mother (11) who is [his] **mother [by] blood**: (9)

"When the autumn comes, (10) [I] will come back"

(3) MYS 15.3691 (WOJ)

(13) 多良知祢能 (14) 波々母都末良母 (15) 安佐都由尔 (16) 毛能須蘇比都知

(13) **tara tine** n-ə (14) papa mə tuma-ra mə (15) asa tuyu-ni (16) mo-nə suso pi^{ndut}-i

(13) **blood mother** COP-ATTR (14) mother FP spouse-PLUR FP (15) morning dew-LOC (16) skirt-GEN hem drench-CONV

(14) although both spouses and mother (13) who is [your] **mother [by] blood**, (16) drenching the hems of [their] skirts (15) in morning dew

6 All *Man'yōshū* poems, as well as their *kana* transcription, romanization, glossing, and translation are based on my edition and translation of this text (Vovin 2009a–).

7 In poems written in WOJ there is a very strong tendency for *tara tine* 'blood mother' to appear in *chōka*. The exceptions for *tara tine* attested in *tanka* are mostly found in books eleven and twelve that have no *chōka* at all.

(4) MYS 16.3811 (WOJ)

(21) 足千根乃 (22) 母之御事歟

(21) たらちねの₂ (22) ははの₂み₁こ₂と₂か(21) **tara tine** n-ə (22) PAPA-NƏ MI-kətə KA(21) **blood mother** COP-ATTR (22) mother-GEN HON-word IP(22) [or did I hear] the words of [my] mother (21) who is [my] **mother** [by] **blood**?

(5) MYS 20.4331 (WOJ)

(19) 多良知祢乃 (20) 波々我目可礼弓

(19) **tara tine** n-ə (20) papa-ŋGa Mɛ kare-te(19) **blood mother** COP-ATTR (20) mother-POSS eye separate.
from(CONV)-SUB(20) [he] is away from the sight of [his] mother (19) who is [his] **mother** [by] **blood**

There is also one example of *tara tine* attested in EOJ:

(6) MYS 20.4348 (EOJ)

(1) 多良知祢乃 (2) 波々乎和加例弓 (3) 麻許等和例 (4) 多非乃加里保尔
(5) 夜須久祢牟加母(1) **tara tine** n-ə (2) papa-wo wakare-te(1) **blood mother** COP-ATTR (2) mother-ACC be.separated(CONV)-SUB(2) being separated from my mother (1) who is [my] **mother** [by] **blood**.

To the best of my knowledge, *tara tine* ‘mother by blood; one’s own birth mother’ is not attested in OOK.

At the first glance, there might be, however, a fly in the ointment: in addition to twenty-four examples of *tara tine* ‘mother by blood’ there are also three examples of the permanent epithet *tara tisi* for the word *papa* ‘mother’ in the *Man’yōshū*: 多羅知斯 (MYS 5.886), 多良知子 (MYS 5.887),⁸ and 垂乳爲 (MYS 16.3791). But ultimately these three examples might also have a solution. *Tara* ‘blood’ has been already discussed above. But *t-isi* might be another

8 多良知子 (MYS 5.887) might have an alternative explanation, but is weaker than the Austronesian one given below. Although the character 子 is used in the extant Old Japanese texts only as the *ongana* sign *si*, in the Heian period *kana* it is also used with its *kungana* value *ne*. Since the Heian period *kana* is an obvious development of the *man’yōgana*, we can easily imagine that such a usage might have occurred in Old Japanese texts that are no longer extant. But this explanation certainly would not work for the examples from MYS 5.886 and MYS 16.3791.

Austronesian word: honorific *t-*, also discussed above + *isi* ‘flesh’, ‘meat’ (Blust and Trussel 2020: *isi₅):⁹ Formosan: Tamalakaw Puyuma *isi* ‘meat’, ‘flesh’, WMP: Kalamian Tagbanwa *isi* ‘flesh’, ‘meat’, Palawan Batak *?isi* ‘meat’, ‘flesh’, Ngaju Dayak *isi* ‘flesh’, Delang *isi* ‘meat’, Maloh *isi?* ‘meat’, ‘flesh’, Simalur *isi* ‘flesh’, Uma *ihi* ‘flesh’, Mori Atas *ihi* ‘meat’, ‘flesh’, Koroni *ihi* ‘meat’, ‘flesh’, Wawonii *ihi*, Moronene *ihi* ‘meat’, ‘flesh’, CMP: Komodo *isi* ‘meat’, Manggarai *ici* ‘flesh’, Ngadha *isi* ‘flesh without bones and fat’, Leti *isi* ‘flesh’, Luang *ihi* ‘meat’, ‘flesh’, Yamdena *isi-n* ‘flesh’, Fordata *ihi-n* ‘flesh’, Kei *ihi-n* ‘flesh of people, animals, and fruits’, Teluti *isi-colo* ‘flesh’, Paulohi *isi-ni* ‘flesh’, Saparua *isi-ni* ‘flesh’, Asilulu *isi-* ‘flesh’, Manipa *si-n* ‘flesh’, Batu Merah *isi-va* ‘flesh’, Morella *isi* ‘flesh’, Buruese *isi-n* ‘meat’, ‘flesh’, Soboyo *isi-n* ‘flesh’. Consequently, *tara t-isi n-ə* papa means ‘a mother, who is [of the same] blood and flesh’. A textual example:

(7) MYS 5.887

(1) 多良知子能 (2) 波々何目美受提

(1) *tara tisi n-ə* (2) *papa-ŋga mɛ mi-ⁿz-u-te*

(1) **blood flesh** COP-ATTR (2) mother-POSS eye see-NEG-CONV-SUB

(2) Without seeing the eyes of [my] mother, (1) who is [of the same] **blood [and] flesh**.

Finally, there is one example of *tara tune* (足常) in MYS 11.2495. I suspect that it is simply a case of an irregular development in WOJ, probably caused by progressive assimilation from the front vowel *i* to the back vowel *u* under the influence of another preceding back vowel *a*.

Another interesting permanent epithet is *nama yəmi-nə* (奈麻余美乃) with unknown meaning, which occurs only once in MYS 3.319 as a *makura-kotoba* for Kapī (甲斐) province. The *yəmi-nə* part is comparatively uncontroversial, as it is likely to be *-nə*, genitive case marker, and *yəmi* in all likelihood reflects *yəmi* in Tuku-yəmi¹⁰ ‘moon deity’. The important thing is that it cannot be Yəmi ‘Hades’, which is very popular as a folk etymology, but the second syllable vowels in *yəmi* and *Yəmi* are very different.¹¹ Sometimes Tuku-yəmi ‘moon deity’,

9 I selected only those words from Blust and Trussel’s list have meanings ‘human flesh’ and/or ‘meat’.

10 There is apparently some confusion in WOJ texts between *yomi* and *yəmi* in this word. Most importantly, there is apparently an example of raising *o* > *u*: *Tukuyomi* > *Tukuyumi* that never happens with the vowel *ə* (Omodaka et al. 1967: 466), so it is more likely that the vowel was *o* and not *ə*. But we should not also exclude a possibility of contamination.

11 Surprisingly, even some scholars support this folk etymology. E.g. Nishimiya dedicates a long discussion where he tries to circumvent this problem, although he understands

tuku-yo ‘moon-lit night’ mean ‘moon light’. Therefore, it is not unconceivable that *yami* could do the same. The first word *nama*, however, is unlikely to be Japonic: while there is *woj nama* ‘fresh’, ‘raw’, it does not fit here semantically. It is always notoriously difficult to deal with *hapax legomenoi* like this one, so our first step will be to examine the context of MYS 3.319. It is a *chōka*, but just a few lines will suffice (8–10):

(8) **Preface to the poem 3.319**

詠不盡山歌一首并短歌

A poem composed about Mt. Puⁿzi with *tanka* envoys.

Commentary

woj Mt. Puⁿzi is the same as *MdJ* Mt. Fuji, a dormant volcano and the highest mountain in Japan.

(9) **MYS 3.319**

(1) 奈麻余美乃 (2) 甲斐乃國 (3) 打縁流 (4) 駿河能國与 (5) 已知其智乃
(6) 國之三中從 (7) 出立有 (8) 不盡能高嶺者

(1) *nama yami-nə* (2) *Kapī-nə KUNI* (3) *UT-I YƏS-URU* (4) *Suru^gga-nə KUNI-tə* (5) *kəti-^ŋgəti-nə* (6) *KUNI-NƏ mi-NAKA-YU* (7) *IⁿDE-TAT-ER-U*
(8) *Puⁿzi-nə TAKA NE pa*

(1) *deep.river moon.light-GEN* (2) *Kapī-GEN province* (3) *hit-CONV approach-ATTR* (4) *Suru^gga-GEN province-COM* (5) *here-there-GEN*
(6) *province-GEN HON-middle-ABL* (7) *go.out(CONV) stand-PROG-ATTR*
(8) *Puⁿzi-GEN high peak TOP*

(5) Here and there, (2) in *Kapī* province, (1) where the **moon light** [is reflected] **in deep rivers**, (4) and in *Suru^gga* province, (3) where [waves] approach and hit [the shore], (8) the high peak of [Mt.] *Puⁿzi*, (7) which is towering high (6) over the middle of [both] provinces ...

(10) **Postscript to the poems 3.319–321**

右一首高橋連蟲麻呂之歌中出焉以類載此

The poem [3.321]¹² above appears in the poetic collection of *Takapasi-nə muraⁿzi Musimarə*. Therefore, I placed it here on the basis of similarity with other poems [in this sequence].

the difference between *kō-rui* vowel /o/ and *otsu-rui* vowel /ə/ (1984: 169–171). Omodaka is much more straightforward, just noting that the meaning of *nama yami* is unknown (1977.3: 239).

12 I did not include this or other poems in this mini-sequence here.

FIGURE 1

Mt. Fuji. Taken from an article published on August 16, 2020 on <https://hapipan.net/namayomi-kai-1/>

I glossed above *nama* as ‘deep river’, because I think that the word reflects yet another Austronesian loan in insular Japonic. Cf. the following data from Blust and Trussel 2020: PMP: **namaw* ‘sheltered water’, ‘deep place in a river’, ‘cove’, ‘harbor’, ‘lagoon’ > WMP: Casiguran Dumagat *namaw* ‘coral reef’, Karo Batak *namo* ‘deep place in a river, generally where it makes a bend and the water moves slowly’, Palauan *l̄máu* ‘small, deep spot within shallow area inside reef’, CMP: Tetun *namo-n* ‘bay’, ‘harbor’, Yamdena *name* ‘deep sea channel in the reef’, Proto-Ambon **nama(w)* ‘bay’, ‘harbor’, Buruese *nama* ‘level stretch in a stream’s course; cove, bay, harbor’, POC **namo* ‘lagoon’ > OC: Wuvulu *namo* ‘lagoon’, Kwaio *namo* ‘lake, pool, deep place in river’, Lau *namo* ‘the lagoon inside a reef, near the reef (the deep) pools towards the shore’, Arosi *namo* ‘a landlocked, shallow lagoon near the shore’, Pohnpeian *nahmw* ‘deep place within the barrier reef’, ‘lagoon’, Tongan *namo* ‘lagoon’, Niue *namo* ‘lake’, ‘pond’. One has to keep in mind that the above Austronesian words mostly reflect seashore topography. Meanwhile, Kapi is a land-locked province, so semantic shifts after the borrowing are to be expected. Kapi has its good share of lakes, ponds, and rivers, but, of course coral reefs or lagoons are not found there. To illustrate this point, I place below a photograph of Mt. Fuji with some large (unidentified) large body of water in front of it.

There are also other *makura-kotoba* in OJ, which as I suspect, might have Austronesian origin, or be partially Austronesian, but this is still the work in progress that will be presented in a future publication. One of them, is, however, dealt with in the next section of this article.

Now I turn to other insular Japonic words that might have Austronesian origin.

3 Other Austronesian Vocabulary in Insular Japonic

My attention was attracted by the Batanic (Bashiic) languages spoken on a chain of small islands, politically divided between two countries: Taiwan and Philippines. While the Northernmost island, Yami (older name: Botel Tobago, Chin. 蘭嶼 *Lán'yú*¹³) 'Orchid Island', belongs to Taiwan (it is located only 64 km from the southeastern coast of Taiwan, all other islands (Itbayat, Batan,¹⁴ and Ibatan), stretching between Northern Luzon and Yami are under Philippines' jurisdiction. In terms of speakers, Ivatan on Batan islands with 35,000 comes the first, followed by Itbayat 3,500, then Yami ca. 3,000, and finally Ibatan (Babuyan) with just 1,240. The current consensus seems to be that the Batanic languages represent a separate WMP subgroup and are not in any case connected genetically to the Philippine languages of Northern Luzon (Ross 2005) and (Blench 2015), however just a few days ago David Zorc and Robert Bust have persuaded me that the Batanic languages are just normal Philippine languages.

The Batanic languages are studied comparatively well, with the first publications produced by Japanese scholars, starting with a ground-breaking study of the Yami language by Asai (1936), and also keeping tempo with more recent work by Tsuchida et al. (1987, 1989), Yamada (2002, 2014), and Yamada and Zayas (1997). From 1960s Western scholars also pitched in: Reid (1966, 1971), Maree (2007), and Maree and Tomas (2012), as well as above-mentioned Ross (2005) and Blench (2015), closely followed by Taiwanese researchers: Li (2000), Zhāng Xún-huì (2000), Yang (2002), and Rau et al. (2012). Finally, there is welcoming sign that Filipino scholars are also getting involved in the study of the Batanic languages: Hidalgo (1998) in addition to the cooperation by Ernesto Constantino in Tsuchida et al. (1989), and C. N. Zayas with Yamada (1997).

Let me start with two Batanic words. The Yami language is naturally saturated with loanwords from Japanese, as the period of Japanese colonial rule in Taiwan was long enough (1895–1945 AD). Therefore, Japanese loans have penetrated even the Yami basic vocabulary, cf. for example, *inu* 'dog' < Jpn. *inu* 'id.' and *saru* 'monkey' < Jpn. *saru* 'id.' But not all Yami-Japanese look-alikes can be classified in this way. For example, Yami *takei* 'mountain' (Zhāng 2000: 132) cannot be a loan from Jpn. *take* (< *woj take*) 'peak', 'mountain', or Ryukyuan *taki* 'id.', which more frequently appears as *u-taki* 'sacred grove on a mountain', 'shrine on a mountain'. There are several facts that strongly speak against the directionality of borrowing from Japanese to Yami in this case. First, Jpn. *take* 'peak', 'mountain' is not a colloquial word in the modern Japanese language.

13 Note that this reading is irregular: we should expect *Lán xù*.

14 Actually, several islands.

It appears mostly either in place names as *take* ~ *-dake*, or in poetry. Therefore, it is highly unlikely that soldiers from the Japanese occupational force on Orchid Island would use this word, and not the much more habitual *yama* ‘mountain’ in reference to local mountains. Second, Yami *takai* ‘mountain’¹⁵ looks suspiciously more like PJ **takay* rather than its descendant forms. Third, there are apparent cognates at least in the Itbayat and Ivatan languages: Itbayat *takey* ‘field’, ‘farm’, ‘forest’, ‘wilderness’ ‘island space other than *hili* (town where people live)’ (Yamada 2002: 252); Ivatan *takey* ‘field’ (GIEID). Proto-Batanic **takai*¹⁶ was apparently borrowed into insular PJ as **takay* due to Arisaka’s constraint on the combination of /a/ and /ə/ within the same morpheme. Note also that in WOJ *take* < PJ **takay* functions as a single morpheme and does not have a compounding form *taka-* resulting from a *-y loss in preconsonantal position. Cf. WOJ *take* < PJ **takay* ‘bamboo’, for which both a free form *take* and a compounding form *taka-* are well attested. We can firmly establish the Batanic languages as a source of borrowing into insular Japonic, because as far as I can tell, *takay* ‘mountain’, ‘wilderness’ is not attested anywhere else in the Austronesian language family.

WOJ *asi* ‘foot’, ‘leg’ is an interesting word, appearing only as a free form, the compounding form as well as the form found before the OJ plural suffix *-na* is just *a-*, for example *a^mbumi* ‘stirrup’ < **a-nə pum-i* ‘foot-GEN step-NML’, *ato* ‘footprint’, ‘track’ < **a-to* ‘foot-track’ or ‘foot-place’, *ayupi* ‘leg cord’ < **a-yup-i* ‘leg-tie-NML’, *a-na suwe* ‘feet end’, etc. Omodaka et al. (1967: 1). The weird behavior of WOJ *asi* ~ *a-* ‘foot’, ‘leg’ is also paralleled by WOJ *unaⁿzi* ‘nape of the neck’ (a free form) and *una-* ‘id.’ (a compounding form). On the other hand, EOJ *a* ‘foot’ appears as a free form in two examples: MYS 14.3387 and MYS 14.3533, but the form *asi* is not attested in the extant EOJ texts. The EOJ examples are below:

(11) MYS 14.3387¹⁷

(1) 安能於登世受 (2) 由可牟古馬母我

(1) *a-nə otə se-nⁿz-u* (2) *yuk-am-u koma mənⁿga*

(1) *foot-GEN sound do-NEG-CONV* (2) *go-TENT-ATTR stallion*

(2) [I] want a stallion, which would go (1) without making any noise with its **hoofs**.

15 Rau and Dong define Yami *takei* as ‘mountain’, ‘field’ in their online dictionary (2009).

16 In both Yami and Itbayat orthography /ə/ is rendered as *e*.

17 This poem is from Simotupusa province, corresponding to the northern part of modern Chiba prefecture.

The tricky part of this example is that EOJ *a* ‘foot’ is found before the genitive case marker *-nə*, which always requires a free form of a noun before, cf. EOJ *ke-nə* ‘tree-GEN’ (MYS 20.4375), unlike WOJ where both free and compounding forms can be found before *-nə*: *kī-nə* ‘tree-GEN’ (*kī* is a free form) and *kə-nə* ‘tree-GEN’ (*kə-* is a compounding form).

(12) MYS 14.3533

(4) 安奈由牟古麻能 (5) 乎之家口母奈思

(4) *a* *nayum-u koma-nə* (5) *wosi-keku mə na-si*

(4) **foot** suffer-ATTR stallion-GEN (5) be.regrettable-ATTR/NML FP exist. not-FIN

(5) [I] have no sorry feelings (4) for [my] stallion that will hurt [its] **legs**.

Here we clearly have a free form of *a* ‘foot’, ‘leg’. Thus, we can reconstruct PJN **a*, but not PJN **asi*, which is clearly innovative, being confined to WOJ, MJ and all modern Japanese dialects that are descendants of MJ (except Hachijō, which is at its core is a descendant of EOJ).

In Ryukyuan the situation is very different as compared to the one in Japanese. While *ʔasi* is attested in Shuri (Kokuritsu kokugo kenkyūjō 1998: 618)¹⁸ and *asi* in Tonosiro (Uchima and Arakaki 2000: 356),¹⁹ given the fact that it is a form with *-si*, which is innovative in Japanese, these two attestations must be viewed as Japanese loans in Ryukyuan. In addition, two other words for ‘foot’, ‘leg’ are attested in Ryukyuan. The first one appears on the periphery as *hagi* in Yuwan and Koniya dialects on Amami island in the North, and as *pagi* in Higashi Nakasone, and *pagi* in Yonaha on Miyako island, and finally as *pay* in Tonosiro and Hateruma dialects of Yaeyama island group in the South of the Ryukyuan archipelago (Uchima and Arakaki 2000: 356). This word is cognate to WOJ *paʔgi* ‘shin’, MJ *fagi* ‘id.’, and other Japanese forms, such as MdJ *-hagi* in *fukura-hagi* ‘calf of the leg’, etc. Both WOJ *paʔgi* ‘shin’ and PR **panki* reflect PJ **panki* ‘foot’, ‘leg’, which underwent a semantic shift in Japanese. There are three independent pieces of evidence that point to PR **panki* as an original Japonic word. First, it is preserved on periphery of Ryukyuan, and innovations typically happen in the center, not on the periphery. Second, the semantic shift

18 Shuri attestation is weird, because it appears only in the Japanese to Okinawan index, but not in the main part of the dictionary.

19 Aleksandra Jarosz also informed me that there are other attestations of this word in Miyako, and that it looks like an old loan due to the compound *asida* ‘geta’ and the shift of *i* > *ɨ*. John kupchik further found *anduu* ‘heel’ in Ryukyuan, probably also related to Tōhoku *aguto* ‘id.’ The main problems, is however that all these words include opaque ‘suffixes’ which makes the establishment of *a-* somewhat not 100% reliable.

'leg' > 'shin' is easy to imagine, while 'shin' > 'leg' is not. Third, this semantic shift in Japanese was undoubtedly triggered by the appearance of OJ *a* 'foot', 'leg'. Finally, there is Shuri *hwisja* (Kokuritsu kokugo kenkyūjo 1998: 618), Oku *pjsa*, Benoki *pjsa*, *Φjsa*, Kumejima *çjsa*, etc. 'foot', 'leg' (Uchima and Arakaki 2000: 356), found in the center of the Ryukyuan language area. This word is apparently cognate to WOJ *piⁿza* 'knee' < PJ **pinsa* 'id.' This time we are dealing with a clear innovation in Ryukyuan: it occurred only in the center, and is a derivation from the word for 'knee'.

The important fact is that PJ **panki* 'foot', 'leg' has reflexes in both Japanese and Ryukyuan, while OJ *a* 'id.' is confined to Japanese. The limited geographical distribution of *a* 'foot', 'leg' makes it a perfect candidate for a loanword.

I think that Batanic is again the best candidate for the source of this loanword. I was able to find Yami *ai* and Itbayat *ayi*, the latter, according to Blust and Trussel 2020 being the case of a secondary resyllabification in order to confirm to PAN canonical disyllabic structure. Ultimately, Yami *ai* and Itbayat *ayi* go back to PAN **qaqay* ~ **waqay* 'foot', 'leg'. The situation with *ai* 'foot', 'leg' is more complicated than with *takai* 'mountain' discussed above, as it is attested outside Batanic as well. I cite below all cases of *ai* in Austronesian according to Blust and Trussel 2020. WMP: Bolinaw Sambal *a:yi* 'foot', 'leg', Maguindanao *ai* 'foot', 'leg', Bintulu *ai* 'foot', Banggai *ai* 'foot', 'leg', CMP: Tetun *ai-n* 'leg', 'foot', Ujir *ai* 'leg', 'foot', Kamarian *ai* 'leg', 'foot', Wahai *ai* 'foot', Hitu *ai* 'leg', 'foot', Batu Merah *ai-va* 'foot', Morella *ai-ka* 'foot', SHWNG: Pom *ai* 'foot'. Still, Batanic remains the likeliest source due to its proximity to Ryukyus and Japan. Blust and Trussel posit the following process of resyllabification: **ay* > *ai* > *ayi*. Our task is to determine at which of these stages the borrowing took place. Certainly not at the monosyllabic **ay* stage, because **ay* would contract to **ε* in pre-Old Japanese, resulting in OJ *e*.²⁰ It is the same process that we have already seen above with PJ **takay* > WOJ *take* 'mountain', 'peak'. If OJ *a* 'foot', 'leg' were from **ay*, we would expect WOJ *e*, not *a* for its free form. And not at the *ayi* stage, because it would stay intact in OJ.²¹ This leaves us with *ai* 'stage', which is further supported by OJ phonotactics that prohibits vowel clusters vv. If certain phonological or morphonological processes lead to formation of such a cluster, one of the vowels has to go, and it is normally higher and narrower vowel that does. In the case of Batanic *ai*, /i/ is certainly higher and narrower than /a/, so in the process of borrowing *ai* > *a*, because *ai* could not be borrowed as such. As much as I generally dislike etymologies

20 The contrast between PJ **ε* and **e* is lost in OJ in the initial position, after coronals, *w* and *y*.

21 See Vovin (2020.1: 48–54) on the contrast /i/:/yi/ in WOJ.

for monophonemic words, I cannot find anything faulty with OJ *a* ‘foot’, ‘leg’ derived from Batanic *ai*.

The last etymology that I am going to discuss here concerns not the Batanic languages, but the Amis language on Taiwan. As demonstrated by Kupchik (2021), the names of several celestial bodies in Ryukyuan and to a lesser degree in Japanese, come from Austronesian and in particular from Amis. This time, however, we are going to descend from heaven to the depths of the ocean.

There are two words in WOJ for ‘whale’: *kuⁿdira* and *isa*. Both words have obscure etymologies, or to put it frankly, no etymologies at all. But there are differences. WOJ *kuⁿdira* is likely to be a non-poetic word: except in the poems KK 9 and its variant NK 7, *kuⁿdira* does not appear even once in whole text of the *Man’yōshū*, but there are plenty of attestations of *isa* in the *Man’yōshū*, while *kuⁿdira* seems to be confined to documents, historical and mythological texts, and dictionaries. Furthermore, *kuⁿdira* undoubtedly enjoys PJ status, as it is everywhere in Japonic: WOJ, MJ, Ryukyuan, MdJ, and modern Japanese dialects. Meanwhile, *isa* is attested only in WOJ, although there is also a piece of evidence that it was attested as well in the dialect of Iki islands: 俗云鯨爲伊佐 ‘In the local dialect ‘whale’ is called *isa*’ (IFIB 528²²). My position has always been that among the two doublets, the one with more limited distribution is a perfect candidate for a loanword.

Indeed, there is Amis *?iso* and Sakizaya *?isu* ‘whale’. The Amis orthography with *o*, normally stands for [o] or [ʊ], therefore Wolff believes that it is the same word as PAN **qisu* ‘shark’ (2010.2: 965). I beg to differ, since sharks are fish, and whales are mammals. Admittedly, people of old might not have such a solid foundation in zoology, as we have today: there are cases when whales are called fish, e.g. WOJ *isa-na* ‘whale-fish’, or Russian folklore *рыба-кит* ‘fish-whale’. Apparently, whatever swims in water is ‘fish’. But still such a maritime people as Amis would know the difference between sharks and whales: it would be highly unlikely if both were perceived simply as ‘big fish’. Interestingly enough, Blust and Trussel (2020) do not include Amis *?iso* ‘whale’ under their PAN **qiSu* ‘shark’, the only representative of Formosan languages cited is Paiwan *qisu* ‘shark’. There is no dearth of numerous WMP and CMP forms provided in Blust and Trussel 2020, but they are not relevant to our purposes here, as none of them could be a source for WOJ *isa*.

Thus, at the initial glance, Wolff’s Austronesian etymology for Amis *?iso* ‘whale’ has a semantic problem, while WOJ etymology *isa* ‘whale’ has a phonetic problem of vowel correspondences Amis /u/ : WOJ /a/ in the second syllable. I will attempt to demonstrate now that this problem with different vocalism

22 Pages are cited according to Akimoto Kichirō’s edition of Fudoki.

in Amis and WOJ is actually a pseudo-problem. First, the phonetic sequence [isu] within the same morpheme in WOJ is extremely rare. To the best of my knowledge, it is found only in five cases: 1) *Isu-nə kami* ‘Top of the rock’, a name of a famous shrine, where *isu* is certainly a result of *o > u raising, and *Iso-nə kami*, an original form that is attested wider and better than *Isu-nə kami*; 2) in three WOJ words: *isusuki*, *isusuku*, and *isurəkəp-* with unclear meaning; 3) in *isu kupasi*, a *makura-kotoba* for *kuⁿdira* ‘whale’, which is also considered to have an unknown meaning (Omodaka et al. 1967: 75–76). Actually, this *makura-kotoba* is extremely important for our purposes here. I think that it represents an Austronesian-Japonic hybrid, where *isu* is Austronesian, being a loan from Amis *ʔiso* ‘whale’. The remaining part is Japonic. Certainly, it is unlikely that we are dealing here with WOJ *kupasi* ‘to be beautiful’, ‘to be perfect’, ‘to be superb’: ‘whale (*kuⁿdira*) that is beautiful [like] whale (*isu*)’ would make little, if any sense at all. The primary function of whales in Japan from times immemorial to the present day, has always been a food source. Therefore, I think that WOJ *kupasi* in this case consists of *kup-* ‘to eat’ and adjectivizer *-asi*, invariably found after verbal stems. If I am right, then *isu kup-asi* may not seem as syntactically modifying *kuⁿdira*, because the attributive form of *-asi* is *-asi-ki*. However, since we are dealing with poetry, *-ki* could have been also omitted to confirm to 5–7 poetic meter. Also, attributive in *-ki* in WOJ is a relatively recent innovation, and as Kupchik indicated to me, it is unlikely to predate this *makura-kotoba*. Then *isu kup-asi kuⁿdira* can be understood as ‘a whale, which is an edible whale’. It is also interesting that all attestations of *isa* in the *Man’yōshū* are in the compound *isa-na* ‘whale-fish’,²³ and all of them are found invariably in the context referring to whale hunting like in the example from MYS 17.3893 below. Therefore, it seems that we have both *isu* and *isa* ‘whale’ in WOJ. I suspect that the second syllable vowel *a* in *isa* is due to the regressive assimilation that took place in the compound *isa-na* < **isu-na*, as well as to overall structural pressure to eliminate the sequence [isu]. Relevant WOJ textual examples (13–15):

(13) KK 9

伊須久波斯久治良佐夜流

isu kup-asi kuⁿdira sayar-u

whale eat-ADJ whale be.caught-FIN

A whale, an edible whale, got caught.

23 There is another possibility for analyzing *isa-na* as ‘whale-PLUR’. See the commentary to MYS 17.3893 (Vovin 2016: 16–17). The possibility of the connection with Amis *ʔiso* ‘whale’ was also suggested for the first time in the same commentary.

(14) NK 7

伊殊區波辭區旆羅佐夜離

isu kup-asi kuⁿdira sayar-i

whale eat-ADJ whale be.caught-CONV

A whale, an edible whale, got caught, and ...

(15) MYS 17.3893

(3) 伊佐魚取 (4) 比治奇乃奈太乎 (5) 今日見都流香母

(3) isa-NA TƏR-U (4) Piⁿdiki-nə naⁿda-wo (5) KEPU MI-t-uru kamə(3) whale-fish catch-ATTR (4) Piⁿdiki-GEN open.sea-ACC (5) today

see(INF)-PERF-ATTR EP

(5) [I] wonder whether [I] will see today (4) the open sea of Piⁿdiki

(3) where [they] catch whales.

4 Austronesian Affixes in Insular Japonic

These are *a priori* difficult to find, because Austronesian uses mostly prefixes and infixes, while in Japonic suffixes rule supreme, except in oj, which has some prefixes as well.

Thus, for example, there is an oj locative prefix *sa-*, which marks not the separate nouns, but NP. This prefix is not recognized in the traditional Japanese grammar, but see its full description in Vovin (2020.1: 101–105, 108–109). Some textual examples (16–20):

(16) NK 4

佐祢耐據茂阿黨播怒介茂

sa-ne-ⁿ-dəkə mo atap-an-u kamo

PREFIX-sleep(NML)-GEN-bed FP give-NEG-ATTR EP

[my wife] will not give me a place to sleep!

(17) MYS 5.800

多爾具久能佐和多流伎波美

tani-ⁿ-guku-nə sa-watar-u kipam-i

valley-GEN-frog-GEN PREFIX-cross-ATTR go to the end-NML

[until] the end [of the Earth], where a toad goes

(18) MYS 5.804

遠等咩良何佐那周伊多斗乎意斯比良伎
 wotəme-ra-ŋga sa-n-as-u ita to-wo os-i-pirak-i
 maiden-PLUR-POSS **PREF-sleep-HON-ATTR board door-ACC push-**
CONV open-CONV
 pushing open the **wooden doors** [of the chambers] **where** maidens
sleep

(19) MYS 15.3735

左奴流欲能伊米爾毛伊母我美延射良奈久尔
 sa-n-uru yo-nə imɛ-ni mo imə-ŋga mi-ye-ⁿz-ar-an-aku n-i
PREF-sleep-ATTR night-GEN dream-LOC FP beloved-POSS see-PASS-
 NEG(CONV)-exist-NEG-NML DV-CONV
in the dreams of the nights when [I] sleep, [my] beloved always
appears

The examples above are all from WOJ, but prefix *sa-* is also attested in EOJ:

(20) MYS 14.3489

左祢度波良布母
 sa-ne-ⁿ-do parap-umə
PREF-sleep-GEN-place clean-EXCL
 [I] am cleaning the **place where [we] sleep [in]!**

As noted in Vovin (2020.1: 109): “There is a dative-locative case marker *-sa* in all Tōhoku dialects. The proper locative function seems to be confined to Tōhoku (those dialects that have it do not have Central Japanese *-ni*). In directive and dative functions case marker *-sa* is found also in Northern and Southern Kyūshū and in Northern Kantō”. In peripheral modern Japanese dialects (Tōhoku and Kyūshū), where it probably became a suffix under the pressure of the SOV word order.

More than fifteen years ago, in the first edition of my WOJ grammar I have suggested that this *sa-* might have an Austronesian origin (Vovin 2005: 90–91), repeated in Vovin (2020.1: 109). At this time, I mentioned the existence of the locative *sa-* in Tagalog, Atayal, and Rukai, noting also that in Atayal *sa-* marks NP like in OJ:

- (21) Atayal (Asai and Ogawa 1935: 34)
 sa-ko maki? BaBao na rahejal
 PREF-I exist top GEN earth
 on the earth, **where** I am
- (22) Tagalog (Schachter and Otnes 1972: 384)
- a. b-um-alik siya **sa Maynila**
 re-AF-turn he LOC Manila
 He returned **to Manila**.
- (Schachter and Otnes 1972: 384)
- b. um-akyat siya **sa puno**
 AF-climb he LOC tree
 He climbed **the/a tree**.
- (Schachter and Otnes 1972: 77)
- c. um-upo **sa silya** ang ale
 AF-sit LOC chair TOP lady
 The lady sat **on the chair**.
- d. (Schachter and Otnes 1972: 76)
 mag-bigay ng pera **sa pulubi** ang tatay
 AF-present ADJC money DAT beggar TOP father
 Father gave **the beggar** some money.

However, the locative *sa* is not limited to Tagalog. I was able to find it in other Philippine languages as well, Hiligaynon (23) and Cebuano (24):

- (23) Hiligaynon (Motus 1971: 70)
 di'in **sa Texas**
 where LOC Texas
 Where **in Texas**?
- (24) Cebuano (Borello 2018: Leksyon 5, p. 3)²⁴
 duol **sa post office**
 close LOC post office
 Close **to post office**.

²⁴ This book has no pagination.

The same holds true in Bikol (*sa* locative class marker occurring before general nouns, see Mintz 1971: 368) and Maranao (*san* ‘there’, ‘in’, ‘at’, see McKaughan and Macaraya 1967: 389). No textual examples for any of them.

This list is probably far from conclusive, but accessible resources on the Philippine languages that I am able to reach are quite limited during the COVID-19 pandemic years. But even on the basis of the short list of attestations above, one can see that the locative *sa* is found in the Northern, Central and Southern Philippines. This is important, since both Atayal and Rukai are the languages spoken in the interior of Taiwan. Consequently, their speakers are not seafarers, and the Austronesian source of OJ *sa-* is unlikely to be in Taiwan. It is a different story with Philippines, which seem to be the likeliest source of OJ *sa-*.

In addition, OJ directive-locative focus verbal prefix *i-* might also have had an Austronesian origin. The locative prefix *i-* attested in the Amis language is, on the contrary, a nominal prefix. This fact makes this etymology less credible, but not impossible. Furthermore, given the situation that the Amis language as demonstrated by Kupchik (2021), is the most prodigious source of Austronesian loans in the insular Japonic, I think that it is even more or less likely. See Old Japanese examples (25–28) below:²⁵

(25) KK 13

伊勢能宇美能意斐志爾波比母登富呂布志多陀美能伊波比母登富理宇
知弓志夜麻牟

Ise-nə umi-nə opi-[i]si-ni pap-i-mətəpər-əp-u sitaⁿdami-nə

i-pap-i-mətəpər-i ut-i-te si yam-am-u

Ise-GEN sea-GEN grow(CONV)-stone-LOC crawl-CONV-go.around-ITER-
ATTR seashell-COMP **DLF-crawl-CONV-go.around-CONV** hit-CONV-SUB
EP stop-TENT-FIN

like the shellfish that are constantly crawling around on the growing
rocks of the Ise sea, [we] **will crawl around** [them] **there**, smite and
stop [them].

(26) KK 86

意富岐美袁斯麻爾波夫良婆布那阿麻理伊賀弊理許牟叙

opə kimi-wo sima-ni pa^mbur-a^mba puna-amar-i i-ⁿgaper-i-kə-m-u ⁿzə

great lord-ACC island-LOC exile-COND boat-exceed-CONV

DLF-return-CONV-come-TENT-ATTR FP

25 All OJ examples are cited from Vovin (2020.2: 505–512).

if [they] exile [my] great lord to an island, there are many boats, and
[he] **would return here.**

(27) **MYS 5.804**

遠等咩良何佐那周伊多斗乎意斯比良伎伊多度利与利提
wotəme-ra-³ga sa-n-as-u ita-to-wo os-i-pirak-i i-taⁿdor-i-yər-i-te
maiden-PLUR-POSS PREF-sleep-HON-ATTR board-door-ACC push-CONV
open-CONV **DLF-pursue-CONV-approach-CONV-SUB**
[gentlemen] push open wooden doors where maidens sleep, and [they]
pursue [maidens] there ...

(28) **MYS 14.3518** (an EOJ example)

伊波能倍尔伊可賀流久毛
ipa-nə [u]pɛ-ni i-kakar-u kumo
rock-GEN top-LOC **DLF-hang-ATTR** cloud
clouds, **hanging** over the rocks

The directive-locative focus prefix *i-* is also attested in Old Ryukyuan:

(29) **OS 12.658**

としが三年いきよてとしが四年いきよて
tosi-ga SAN-NEN i-kiyo-te tosi-ga YO-NEN i-kiyo-te
year-POSS three-year **DLF-invite-SUB** year-POSS four-year
DLF-invite-SUB
inviting [the goddess] **here** for three years, **inviting** [the goddess] **here**
for four years

Now let us turn to the examples in the Amis language:²⁶

(30) Amis (Cài and Zēng 1997: 307)

- a. adihay ko kapaysinan no 'Amis **i-tiya ho**
be.many SM taboo GEN 'Amis **LOC-long.time ago**
There were many taboos for Amis **in the times now past.**

26 The following examples are all from (Cài and Zēng 1997). However, since Càì and Zēng's glosses are sometimes incomplete and also include omissions, I used the Amis dictionary by Virginia Fey (1986) to fill in the necessary gaps.

(Cài and Zēng 1997: 306)

- b. **i-tiraay** ingata no lalan no tosiya ko niyaro'
LOC-there be.close GEN road GEN car SM village
 The village **there** is close to the motor-way.

(Cài and Zēng 1997: 306)

- c. hinam-han nira **i-taliyok** no talo'an ira ko mapeliay a narakatan no
 tamdaw
 look.around-DEB he **LOC-perimeter** GEN small.hut exist SM be.left
 CP trace GEN person
 When he was prompted to look **around** the small hut, there were
 traces left by a person.

5 Conclusion: When and Where?

I will try to give very approximate and tentative answers to two questions here: “When these Austronesian loanwords entered insular Japonic?” and “Where did it happen?”

The first question is easier than the second. I do not see any traces of these loanwords in peninsular Japonic, therefore the upper temporal boundary must not be earlier than 300 BC, the rough date when the Yayoi culture that could be defined as Proto-Japonic first appeared in Japan. Although it has become trendy recently to push back the beginnings of the Yayoi culture to 700 BC or so, I am not convinced, because the evidence presented is not very reliable. As for lower temporary boundary, 400 AD seems to be quite likely, because while our knowledge of Ancient Japan increases exponentially after this date, leaving no likely room for significant population movement or language contacts on a large scale (and judging by the nature of these loans, the language contact was indeed on a very large scale, since we are dealing not with just a peripheral, but with a basic vocabulary). However, before 400 AD we know precious little about the ethnolinguistic history of Japan.

The second question is much more difficult to answer. However, at the present point we already have several pieces of the puzzle at our disposal. First, we know that the initial landfall of the Yayoi culture in Japan was in Northern Kyūshū, from where it steadily and rapidly spread to the rest of Kyūshū, Shikoku, as well as to the most of Honshū, excluding its Northern part, which was dominated by Ainu at least until the late ninth or early tenth century AD. Ainu-Japanese bilingualism was still present in the Kantō region of Japan well into the early tenth century AD (Vovin 2009b, 2022). Second, we know

that Proto-Ryukyuan speakers did not start to migrate to Ryukyuan islands earlier than the late ninth century or early tenth century AD. There are different opinions as to where Proto-Ryukyuan was spoken: in the North-Eastern Kyūshū, North-Eastern Kyūshū, or Southern Kyūshū (Serafim 2003), but we can say for sure that the Austronesian loanwords detected by Kupchik (2021) in Old Okinawan come back to the time when Proto-Ryukyuan was still spoken on Kyūshū. Third, although we know very little about the Hayato (WOJ Payatə, 隼人) and almost nothing about Kumasō (WOJ Kumasə, 熊襲) tribal languages in Kyūshū, their customs, headgear, clothes, shields, etc. look strikingly Austronesian (Ōbayashi 1975: 24–33). A treatment of the bits and pieces of the Hayato and the Kumasō languages would require a separate article, so here I will limit myself to just one note. Almost half a century ago it was suggested by Murayama that WOJ *-sə* in Kumasə and *-tə* in Payatə reflect PAN *Cau ‘person’, ‘human being’ (1975: 258–260). Apparently, there is a phonetic difficulty: we would rather expect WOJ *-to* and *-so* rather than WOJ *-tə* and *-sə* as a reflex of PAN *Cau. But there is a spelling Kumasəo (球磨磨於) in Fudoki (CFIB 503), (HFIB 520, 521), (BF 358, 368), (HF 380), or almost identical 球磨磨於 (HF 404). This spelling apparently violates the OJ phonotactics prohibiting vowel clusters, and, consequently could only be an attempt to represent a foreign word. It still does not explain the vowel [ə], but, nevertheless brings us one step closer to PAN *Cau. The Extra-Formosan MP languages all uniformly have initial consonant *t-*, while the Formosan languages show here a variety of reflexes: Papora *sō*, ‘person’, ‘human being’, Pazeh *saw* ‘person’, ‘human being’, Thao *caw* ‘person’, ‘human being’, Hoanya *sau* ‘person’, Tsou *cou* ‘person’, ‘human being’, Puyuma *Tau* ‘to be human’ (Blust and Trussel 2020). For WOJ *-sə* Formosan origin will work better, but for *-tə* either Formosan or MP will do, because OJ has never had any affricates [c] or [č]. A possible solution to [ə] vowel may be offered by Tagalog that has *táʔo* ‘person’, ‘human being’ with penultimate stress, but *taʔó* ‘mortal’ with a stress on the last syllable. It is quite possible that WOJ *səo* (贈於, 贈於) represents an attempt to transcribe *səʔo*. Furthermore, unfortunately our knowledge about OJ accentual system is speculative at best, but if *səʔo* had an accent pattern LH, it is not inconceivable that the vowel [ə] is reduction of an original *a. Therefore, we can posit a hypothetical *saʔu, which will be quite in agreement with Austronesian forms.

All three independent pieces of evidence discussed above point in the same direction: the quite intensive language contact between Austronesian and insular Japonic took place on Kyūshū sometimes between 300 BC and 400 AD. There will be many years ahead of the hard work before we get more or less coherent general picture, but it looks like that starting with Kupchik (2021) a new subfield of Japonic historical linguistics is born.

To put an icing on the cake, it is quite possible that Austronesians might have travelled to much more distant lands in the North than Japan. Long ago I have noticed PE *iqaluγ ‘fish’ (no Aleut cognates). In this word *-luγ looks like a postbase, but it cannot be *-luγ* ‘to be bad’, and needs to be explained. The remaining part *iqa- suspiciously looks like PEMP *ikan ‘fish’ < PMP *hikan < PAN *Sikan. Aleut *tana-x̣* ‘land’, ‘earth’ (traditionally compared with Eskimo *nuna* ‘id.’, but this etymology has too many problems) no less suspiciously looks like Maranao, Manobo, Visayan, etc. *tana?*, Malay, Old Javanese, Javanese, Balinese, etc. *tanah* < PMP *taneq ‘land’, ‘earth’, ‘soil’ (Blust and Trussel 2020). Certainly, these might be just chance look-alikes, but, in my opinion, this possibility is worth investigating a little bit further, in spite of the fact that poor Austronesians would be chilled to their bones in these frigid Arctic waters.

Acknowledgments

Partially funded by the ERC Advanced Grant 788812 AN ETYMOLOGICAL DICTIONARY OF THE JAPONIC LANGUAGES (PI Alexander Vovin). I use the occasion to express my gratitude to my friend and colleague Osada Toshiki (長田 俊樹), who provided me with valuable Japanese materials on the Batanic languages, which are almost impossible to find outside Japan during the COVID-19 pandemic. No lesser amount of my gratitude goes to John Kupchik, my former student from the University of Hawai’i and now an independent and excellent scholar, and one of most valuable colleagues, and to Aleksandra Jarosz, no less valuable or brilliant colleague and friend, with the primary specialization in the Miyako language of the Ryukyuan. Both Aleksandra and John read an earlier draft of this paper and provided many useful comments that helped to improve my article tremendously. Needless to say, I am alone is responsible for any shortcoming or mistakes in this article.

Abbreviations

Languages and Sources

Chn.	Chinese
CMP	Central Malayo-Polynesian
EOJ	Eastern Old Japanese
GIEID	Glosbe Ivatan-English and English-Ivatan Dictionaries. n. p., n. d.
Jpn.	Japanese
MdJ	Modern Japanese

MJ	Middle Japanese
MP	Malayo-Polynesian
OC	Oceanic
OJ	Old Japanese (both Western and Eastern)
OOK	Old Okinawan
PAN	Proto-Austronesian
PEMP	Proto-Eastern-Malayo-Polynesian
PJ	Proto-Japonic
PJN	Proto-Japanese
PMP	Proto-Malayo-Polynesian
POC	Proto-Oceanic
PR	Proto-Ryukyuan
SHWNG	South Halmahera and Western New Guinea
WMP	Western Malayo-Polynesian
WOJ	Western Old Japanese

Linguistic Terms

ABL	Ablative
ADJ	Adjectivizer
AF	Actor focus
ATTR	Attributive
COM	Comitative
COND	Conditional converb
CONV	Converb
COP	Copula
CP	Connecting particle
DAT	Dative
DEB	Debitive
DP	Desiderative particle
FIN	Final
FP	Focus particle
GEN	Genitive
H	High pitch
HON	Honorific
IP	Interrogative particle
L	Low pitch
LOC	Locative
NEG	Negative
NML	Nominalizer

NP	Noun phrase
PLUR	Plural
POSS	Possessive
PROG	Progressive
QV	Quotational verb
SM	Subject marker
SUB	Subordinative converb
TENT	Tentative
TOP	Topic

References

Primary Sources

BF *Bungo Fudoki* (豊後風土記), *Bungo Fudoki*, 713 AD
CFIB *Chikuzen Fudoki Itsibun* (筑前風土記逸文), *Fragments of the Chikuzen Fudoki*, between 713 and 737 AD
HF *Hizen Fudoki* (肥前風土記), *Hizen Fudoki*, 713 AD
HFIB *Higo Fudoki Itsibun* (肥後風土記逸文), *Fragments of the Higo Fudoki*, between 713 and 737 AD
IFIB *Iki Fudoki Itsibun* (壱岐風土記逸文), *Fragments of the Iki Fudoki*, between 713 and 737 AD
KJK *Kojiki* (古事記), *Records of Ancient Matters*, 712 AD
KK *Kojiki kayō* (), *Songs of the Kojiki*, 712 AD
MYS *Man'yōshū* (萬葉集), *A Collection of Ten Thousand Leaves [of Words]*, between 759 and 785 AD
NK *Nihonshoki kayō* (日本書紀歌謠), *Songs of the Nihonshoki*, 720 AD
NS *Nihonshoki* (日本書紀), *Annals of Japan*, 720 AD
OS *Omoro sōshi* (おもろさうし), *Omoro Anthology*, 1531–1623 AD
WMS *Wamyō Ruijushō* (倭名類聚抄), *Classified Japanese Words*, 931–938 AD

Secondary Sources

Akimoto, Kichirō 秋本吉郎, ed. 1958. 『風土記』。「日本古典文學大系」2。[Gazetteers. Series of the Japanese Classical Literature, vol. 2.] Tokyo: Iwanami.
Asai, Erin. 1936. *A study of the Yami language: An Indonesian language spoken on Botel Tobago Island*. Leiden: Ginsberg.
Asai, Erin 淺井恵倫 and Ogawa Naoyoshi 小川尚義. 1935. 『原語による台湾高砂族伝説集』。[A Collection of Legends of Taiwan Hill Tribes Based on the Native Languages.] Taihoku: Taihoku Imperial University.

- Blench, Roger. 2015. *A New Look at Bashiic*. Draft version. https://www.academia.edu/14178249/A_new_look_at_Bashiic_a_divergent_subgroup_of_Malayo_polynesian.
- Blust, Robert and Stephen Trussel. 2020 (last revision). *The Austronesian Comparative Dictionary*. <https://www.trussel2.com/acd/>.
- Borello, Alfonso. 2018. *My Little Book of Cebuano Visayan*. Vol. 1. n. p., n. p.
- Cài, Zhōng-hán 蔡中涵 and Zēng, Sī-qí 曾思奇. 1997. 《阿美族母語語法結構分析》 [An Analysis of Syntactic Constructions in the Grammar of the Amis Tribal Mother Tongue.] Taipei: Taiwan Aboriginal Foundation.
- Fey, Virginia (Fāng Mǐn-yīng 方敏英). 《阿美語辭典》 / *Amis Dictionary*. Taipei: The Bible Society.
- GIEID. Glosbe English-Ivatan Dictionary. n. p., n. d. <https://en.glosbe.com/en/ivv>.
- Hidalgo, Cesar A. 1998. *Ivatan-Filipino-English dictionary: the cultural dictionary of Batanes*. Manila: Academics Foundation.
- Hisamatsu, Sen'ichi 久松潜一 et al. 1958. 『古事記総索引』. [A General Index to the Kojiki], vols. 1–2. In: 『古事記大成』 *The Kojiki Series*, vol. 7–8. Tokyo: Heibonsha.
- Kokuritsu kokugo kenkyūjo (国立国語研究所), eds. 1998. 『沖縄語辞典』. [A Dictionary of the Okinawan Language.] Tokyo: Ōkurashō insatsu kyoku.
- Kupchik, John. 2021. Austronesian Lights the Way: on the Origin of the Words for SUN in Old Ryukyuan. *International Journal of Eurasian Linguistics* 3 (2). 237–257.
- Kuroita, Katsumi 黒板勝美 et al. 1974. 『日本書紀. 前編–後編』. [The Nihonshoki], vols. 1–2. 『新訂増補國史大系』. [The Newly Edited and Enlarged Series on the Japanese History.] Tokyo: Yoshikawa Kōbunkan.
- Li, Paul Jen-Kuei. 2000. Subgrouping of the Batanic Languages. n.p., n.p.²⁷
- Maree, Judith Y. M. and Orlando R. Tomas. 2012. *Ibatan to English dictionary: with English, Filipino, Ilokano, Ivatan indices*. Manila: SIL Philippines.
- Maree, Rundell D. 2007. *Ibatan: A Grammatical Sketch of the Language Babuyan Claro Island*. Manila: SIL Philippines.
- McKaughan, Howard P. and Batua A. Macaraya. 1967. *Maranao Dictionary*. Honolulu: University of Hawai'i Press.
- Mintz, Malcolm W. 1971. *Bikol Dictionary*. Honolulu: University of Hawai'i Press.
- Motus, Cecile L. 1971. *Hiligaynon Lessons*. Honolulu: University of Hawai'i Press.
- Murayama, Shichirō 村山七郎. 1975. 「隼人の言語」. [The Language of Hayato.] In: Ōbayashi, Taryō 大林太良 (ed.). 『隼人』. 日本古代文化の探究. [Hayato. The Search for the Ancient Culture of Japan Series.] Tokyo: Shakai Shisōsha. 249–263.
- Nishimiya, Kazutami 西宮一民. 1984. 『萬葉集全注』. 卷第三. [The Man'yōshū Fully Annotated, vol. 3.] Tokyo: Yūhikaku.

27 I was able to obtain a very strange .pdf file which includes only the first two pages.

- Ōbayashi, Taryō 大林太良. 1975. 「民俗学から見た隼人」. [Hayato from the Point of View of Ethnography.] In: Ōbayashi, Taryō 大林太良 (ed.). 『隼人』. 日本古代文化の探究. [Hayato. The Search for the Ancient Culture of Japan Series.] Tokyo: Shakai Shisōsha. 2–62.
- Omodaka, Hisataka (澤瀉久孝). 1977.3. 『萬葉集注釋』. 卷第三. [The Man'yōshū with a Commentary, vol. 3.] Tokyo: Chūō kōronsha.
- Omodaka, Hisataka 澤瀉久孝 et al. 1967. 『時代別国語大辞典』. [A Large Dictionary of Japanese by Periods – Old Japanese.] Tokyo: Sanseidō.
- Pearson, Richard. 2013. *Ancient Ryukyu*. Honolulu: University of Hawai'i Press.
- Rau, Victoria, Dong, Maa-Neu. 2009. *Yami (Tao) Dictionary Project*. http://yamibow.cs.pu.edu.tw/index_en.htm.
- Rau, Victoria, Dong, Maa-Neu and A. H. -H. Chang. 2012. *Yami (Tao) dictionary*. Taipei: National Taiwan University Press.
- Reid, Lawrence A. 1966. *An Ivatan syntax*. Honolulu: Hawai'i University Press (Oceanic Linguistics Special Publication #2).
- Reid, Lawrence A. 1971. *Philippine minor languages: Word lists and phonologies*. Honolulu: Hawai'i University Press (Oceanic Linguistics special publication #8).
- Ross, Malcolm. 2005. The Batanic Languages in Relation to the Early History of the Malayo-Polynesian Subgroup of Austronesian. *Journal of Austronesian Studies* 1 (2). 1–23.
- Serafim, Leon A. 2003. When and from Where did the Japonic Language Enter the Ryukyus? – A Critical Comparison of Language, Archaeology, and History. In: Vovin, Alexander and Osada Toshiki 長田俊樹 (eds.). 2003. 『日本語系統論の現在』/ *Perspectives on the Origins of the Japanese Language*. Kyoto: International Research Center for Japanese Studies. 463–475.
- Tsuchida, Shigeru, Yamada, Yukihiro and Moriguchi, Tsunekazu. 1987. *Lists of selected words of Batanic languages*. Tokyo: Department of Linguistics, University of Tokyo.
- Tsuchida, Shigeru, Constantino, Ernesto, Yamada, Yukihiro and Moriguchi, Tsunekazu. 1989. *Batanic Languages: Lists of Sentences for Grammatical Features*. Tokyo: Department of Linguistics, University of Tokyo.
- Uchima, Chokujin 内間直仁 and Arakaki, Kumiko 新垣公弥子. 2000. 『沖縄北部・南部方言の記述的研究』. [A descriptive Study of Northern and Southern Ryukyuan Dialects.] Tokyo: Kazama shobō.
- Vovin, Alexander. 2005. *A Descriptive and Comparative Grammar of Western Old Japanese*. Part 1: Phonology, Script, Lexicon and Nominals. First edition. Folkestone: Global Oriental.
- Vovin, Alexander. 2009a–. *Man'yōshū*. Folkestone: Global Oriental (books 5, 14, 15, 20), and Leiden and Boston: Brill (1, 2, 16, 17, 18, 19 so far).
- Vovin, Alexander. 2009b. 『万葉集と風土記に見られる不思議な言葉と上代日本列島におけるアイヌ語の分布』. [Strange Words in the Man'yōshū and the Fudoki]

- and the Distribution of the Ainu Language in the Japanese Islands in Prehistory.] Kyoto: The International Research Center for the Japanese Studies.
- Vovin, Alexander. 2016. *Man'yōshū*, book 17: *A New English Translation Containing the Original Text, Kana Transliteration, Romanization, Glossing and Commentary*. Leiden and Boston: Brill.
- Vovin, Alexander. 2020. *A Descriptive and Comparative Grammar of Western Old Japanese*. Volumes 1 and 2. *Handbuch der Orientalistik* v: 16.1–2. Leiden and Boston: Brill. Second edition.
- Vovin, Alexander. 2022 (forthcoming). Ainu Elements in Early Japonic. In: Bugaeva, Anna (ed.). *Handbook of Ainu linguistics*. Berlin: De Gruyter.
- Wolff, John U. 2010. *Proto-Austronesian Phonology with Glossary*, 2 vols. Cornell: Cornell University Press.
- Yamada, Yukihiro 山田幸宏. 2002. *Itbayat-English Dictionary*. Endangered Languages of the Pacific Rim Publications Series A3–006. Suita: Osaka Gakuin University.
- Yamada, Yukihiro 山田幸宏. 2014. *A Grammar of the Itbayat Language of the Philippines*. Himeji: Himeji Dokkyo University.
- Yamada, Yukihiro 山田幸宏 and C. N. Zayas. 1997. *A Bibliography of the Bashiic languages and cultures*. CSSP Publications, UP Diliman, Office for Research and Development, Initiative for the Study of Ethnolinguistic Groups in the Philippines.
- Yang, D. 2002. *Subgrouping and Reconstruction of Batanic Languages*. Unpublished MA thesis. Taichung: Providence University.
- Zhāng Xún-huì 張旬慧. 2000. 《雅美語參考語法》。[A Reference Grammar of the Yami Language.] Taipei and Hong Kong: Yuanliu chuban gongsi (Austronesian Languages of Taiwan 13).