

**‘An examination of a possible correlation between the tone distinction of the word-initial mora of Old Japanese words and the voicing distinction of the word-initial consonant of the putative matching Austronesian words’**

**Yoshizo Itabashi**  
**University of Kyushu**

**Abstract:** This paper is an attempt to determine to what extent the rule can predict the cognateship of the putative matching pairs of Old Japanese and Austronesian words. This rule is that if the tone of a word-initial mora of an Old Japanese word is correlated with the voicing of the word-initial consonant of the putative matching word in Austronesian then it can be postulated that that word is most likely a cognate. This may eventually lead to the further postulation that there is a strong genealogical connection between Japanese and Austronesian.

**Key words:** high/low tones, voicing, word-initial consonant, Old Japanese, Austronesian

### **Abbreviations**

PAN: Proto-Austronesian

PHF: Proto-Hesperonesian (Western Austronesian + Formosan)

PF: Proto-Formosan

PMP: Proto-Malayo-Polynesian

PWMP: Proto-Western-Malayo-Polynesian

PIN: Proto-Indonesian

PPH: Proto-Philippine

PCEMP: Proto-Central-Malayo-Polynesian

POC: Proto-Oceanic

OJ: Old Japanese

PJ: Proto-Japanese

### **Introduction**

This paper is an attempt to determine if the rule below is able to make a clear distinction

between the following putative pairs of correspondences between Old Japanese and Austronesian words:

If the word-initial **high-tone** mora (e.g.[pV<sup>high</sup> -]) of an Old Japanese word corresponds to the word-initial **voiceless** consonant (e.g.[p-]) of the putative matching Austronesian word and likewise, if the word-initial **low-tone** mora (e.g.[pV<sup>low</sup> -]) of an Old Japanese word corresponds to the word-initial **voiced** consonant (e.g.[b-]) of the putative matching Austronesian word, then these pairs of tone/voicing matching endorse our postulation that they are cognates. If there is no such tone/voicing correspondence between them, then that result may eventually weaken our further postulation on the genetic relationships of Japanese and Austronesian.

This rule as a working hypothesis is made on the basis of tonogenesis, explained in the following section. The only weak point of this rule is that it can apply only to words whose word-initial consonant has a voiceless sound such as [p][t][k][s][c]. This rule, however, is to offset that lacuna to a great extent, that is, it would be able to distinguish the original voicing contrast of a number of words with a voiceless word-initial consonant.

We should bear in mind that words treated in this paper are the ones not normally borrowed partly because they are in the category of basic vocabulary and the related areas and partly because Japanese and Austronesian are geologically distant. Therefore, we can as far as to say that although the putative regular phonological correspondences between the two endorse no close genealogical connections, they definitely show a high probability of the distant genealogical relationships.

### **1. Tone and voice correspondence**

There is a linguistic phenomenon called ‘tonogenesis’, a term coined by James Matisoff (1970): tone arises as a consequence of the loss or merger of consonants, and in a non-tone language, a voiced consonant in a word-initial syllable commonly makes the following vowel in that syllable pronounced at a lower pitch than a voiceless consonant does. This is usually a minor phonetic (as opposed to phonemic) detail of voicing, and a general tendency seen in many languages. However, if consonant voicing is subsequently lost, that incidental pitch difference may be left over to carry the distinction that the voiced consonant had carried. Therefore, the voicing distinction becomes phonemic.

We find the parallel phenomenon occurred in Old Japanese. As a working hypothesis, we can interpret tone distinction in OJ resulted from the loss of the original voicing distinction in the word-initial consonant in the following way: if Pre-Proto-Japanese (PPJ; non-tone language) originally had voicing distinction such as PPJ\**pana* ‘nose’ and PPJ\**bana* ‘flower’ described below, and subsequently the word-initial consonantal voicing was lost in PJ because Proto-Japanese just like that in Old Japanese had a phonotactic rule that a word-initial consonant must be voiceless, then that tone difference may result to carry the voicing distinction. These changes are summarized as follows:

OJ *pana* (**HH**) ‘nose’ < PJ \* *pana* ‘nose’ < PPJ \* *pana* ‘nose’

OJ *pana* (**LL**) ‘flower’ < PJ \* *pana* ‘flower’ < PPJ \* *bana* ‘flower’

**H** stands for ‘high tone’ within a mora and **L** stands for ‘low tone’ within a mora and **hl** stands for ‘falling tone’ within a mora.

Thus, we can assume that the word-initial **high-tone** mora (e.g.[**pV**<sup>high</sup> -]) of an OJ/PJ word was derived from the PPJ word-initial **voiceless** consonant (e.g.[**p-**]) and likewise the word-initial **low-tone** mora (e.g.[**pV**<sup>low</sup> -]) of an OJ/PJ word was derived from the PPJ word-initial **voiced** consonant (e.g.[**b-**]).

This notion can generally apply to other Old Japanese words which contrast in tone of the word-initial consonants, there seems to be an interesting correspondence between tone and voicing in Old Japanese and Proto-Japanese.

It is noted here that the tones of these words recorded in the *Ruiju Myogisho* (RM) compiled in the 11<sup>th</sup> century are thought to be the same tones as those in OJ by most Japanese linguists such as Haruhiko Kindaichi. Namely, there was no change in tone from the early OJ period to the late OJ period.

## 2. Accent patterns of Old Japanese

Pitch accent in Old Japanese is defined as a focus on a mora within a word with some morae. That is, a pitch is a signal for the mora with the pitch to stand out within a word, and a pitch is called a tone if the pitch occurs regardless of any position of the mora in a word.

Teruhiro Hayata (1999) and other linguists such as Robert Ramsey (1990: 221) insist that Old Japanese, as well as Proto-Japanese is a tone language rather than a pitch-accented language (in this paper, however, we mainly focus on the tones of Old Japanese rather than those of Proto-Japanese). This idea has been accepted to be better able to explain the sound system of Old Japanese. An Old Japanese disyllabic noun had one of the following tones, HH, HL, LL, LH and also an Old Japanese suffix such as *nga*, *wo*, *ni* had an independent

tone. These pieces of evidence lead to their claim that Old Japanese was a tone language, which we accept in this paper. Thus we use tones instead of pitch when we refer to low or high pitch of Old Japanese or of Proto-Japanese.

The tones of Old Japanese monosyllabic and disyllabic nouns are classified in the following: @ (*ŋga*)<sup>1</sup>

1.1 H(H): *po* ‘mast’, *ti* ‘blood’

1.2 H(L): *ke* ‘hair’, *pa* ‘leaf’

1.3 L(L): *ta* ‘rice paddy’, *yu* ‘hot water’

A vowel of a monosyllabic word was normally lengthened twice as long in Old Japanese, thus for instance, *po* ‘mast’ was actually pronounced as [po:].

#### @@ (*ŋga*)

2.1 HH(H): *pana* ‘nose’, *kaze* ‘wind’, *usi* ‘cow’, *tori* ‘bird’

2.2 HL(L): *pasi* ‘bridge’, *isi* ‘stone’, *oto* ‘sound’

2.3 LL(L): *pana* ‘flower’, *yama* ‘mountain’, *kumo* ‘cloud’

2.4 LH(H): *umi* ‘sea’, *kata* ‘shoulder’, *iki* ‘breath’

2.5 LH(L): *kope* ‘voice’, *ame* ‘rain’, *kage* ‘shadow’

@ stands for a mora and the H or L in the parenthesis indicates the assumed tones of a suffix to those nouns. For instance, a class 2.1 noun *hana* ‘nose’ with the nominative case suffix *ŋga* is denoted as the HH(H) tonal pattern as in *hanaga*.

Most linguists and dialectologists accept the standardized hypothesis that the tonal system of Old Japanese was identical to that of the 11th century Kyoto dialect recorded in the *Ruiju Myogisho* (RM). They also agree that the tones of Proto-Japanese were similar to those of Old Japanese.

### 3. Accent patterns in Austronesian

Word accent makes a difference in meaning or in grammar in Austronesian and is phonemic. It also involves the following three details (Zorc 1993:18):

- 1) long vs. short vowels (length)
- 2) pitch or high vs. low tones (pitch accent)
- 3) loudness or amplitude (stress accent)

Zorc (1993:18) argues which of these is the most important feature of Philippine

languages, while much debate over this problem has gone on. He, then gives some Philippine examples that show the interplay of length, pitch and stress in contrastive (phonemic) word accent as follows (there are more examples to be drawn from the other languages like Butuanon, Cebuano, Bontok, Ibanag, Isneg, and others, but the examples below are enough to illustrate his point.):

	length	pitch/high vs. low	stress	meaning
Tagalog	[ʔa:so]	[ʔa(↑)so(↓)]	[ʔásó]	‘dog’
vs.	[ʔaso]	[ʔa(↓)so(↑)]	[ʔasó]	‘smoke’
Ilokano	[ba:ra]	[ba(↑)ra(↓)]	[bára]	‘hot’
vs.	[bāra]	[bǎ(↓)ra(↑)]	[bará]	‘lungs’
Aklanon	[la:la]	[la(↑)la(↓)]	[lála]	‘hurt’
vs.	[lāla]	[la(↓)la(↑)]	[lalá]	‘braid’
Bikol	[ba:ga]	[ba(↑)ga(↓)]	[bága]	‘ember’
vs.	[bāga]	[ba(↓)ga(↑)]	[bagá]	‘truly!’
Pampango	[ʔa:piʔ]	[ʔa(↑):pi(↓)ʔ]	[ʔápiʔ]	‘lime’
vs.	[ʔǎpiʔ]	[ʔa(↓)pi(↑)ʔ]	[ʔapiʔ]	‘fire’

Zorc (1993:18-19) also states that stress or pitch accent is a syntactic (=grammatical) feature in all Philippine languages, which is a means of indicating an accent group, i.e. phrase. Stress or pitch accent does not even necessarily coincide with length:

Tagalog usually [sí:no] ‘who?’ but [si:nó] ‘who?’ (impatient, angry)

Aklanon usually [ná:nu] ‘what?’ but [na:nú] ‘what?’ (irritation, duress)

This phonetic phenomenon, however, is commonplace in many languages to indicate strong emotions, so that this positional difference in stress or pitch, or even vowel length is not phonemic in this particular case. This is a minor special use of stress, pitch or length and thus normally indicates the semantic difference, as shown in many examples of the Philippine languages above.

Zorc (1993:19) insists that even length can be a feature of intonation, although it is a lot more limited: Aklanon [támbuk] ‘fat’, [katámbuk] ‘very fat’, [katámbu:k] ‘very, very fat!’. He then goes on to say that there is inherited length in the Philippine languages, which goes back to thousands of years. He reconstructed some minimal pairs for Proto-Philippine (Zorc 1993:19-20):

6 **Yoshizo Itabashi**

\*ká:yuh ‘tree, wood’ vs. \*kayú ‘you (pl.)’

\*ʔásu ‘dog’ vs. \*qasúh ‘smoke’

\*kí:taʔ ‘see’ vs. \*kitá ‘we (incl.)’

Zorc (1993:20) also contends that length may develop as the result of the loss of some consonant.

Pangasinan: *ba:lo* ‘new’ (< baʔlo) vs. *baló* ‘widow’ (<PPH \*ba:lu)

Ibanag: *a:yam* ‘play’ (<ayyam) vs. *ayám* ‘animal’ (<PPH \*qa:yam)

Tausug: *í:pún* ‘slave’ (<quDípún) vs. *ipún* ‘tooth’ (<PPH \*i:pen)

Zorc (1978:71-72; 1983:4-6; 1993:20) proposes a system of classifying Austronesian languages on the basis of the role of accent involving eleven criteria (only a few criteria are shown in this paper; see Zorc 1993:20): the classification of any given language may reflect more than a single criterion. Take Tagalog for instance, it involves five criteria(out of which (a), (b) are shown here) of the following:

- (a) Phonemic length and shortness, as inherited from PPH (<PHN, PMP, PAN):  
Bisayan, Coastal and Pandan Bikol, Balangaw, Bontok, Hanunco, Ifugao, Ilokano, Isneg, Itneg, Itawit, Kalinga, kapampangan, Kankay, Malaweg, Sambal, Tagalog.
- (b) Phonemic accent (quantity or stress) as secondarily introduced, generally due to consonant loss, analogical leveling, or borrowing: Casiguran-Dumangat, Ibanag, Pangasinan, Old Javanese, Ratahan, Malagasy.
- (c) Length contrasts in the ultimate, resulting from compensation for the loss of a consonant: Tausug, Butuanon, Kamayo, Cebuano.

Zorc (1993:21) also shows evidence for the reconstruction of Austronesian accent: the PAN vowels \*a,\*i, \*u could be either long or short, but \*e could neither be long nor stressed. Thus PAN reconstructions are proposed in the following:

PAN \*bĕRáy ‘give’, \*dĕpáh ‘fathom’, \*qĕtút ‘fart’, \*Sĕmáy ‘rice’, \*Sĕdám ‘borrow’  
He continues on to state some other details of the PAN accent system.

Zorc (1993:22) concludes that the diverse accent systems are found among Austronesian languages and no single pattern is dominant. Thus, no simplistic solutions or assumptions are possible about Proto-Austronesian and it is safe to assume that PAN had a complex accentual system, which in turn yielded the diversity of accentual systems throughout the Modern Austronesian family.

As we have seen above, Austronesian languages have a complex system of word accent, but all of them have a fixed accent pattern except for the ones derived from the original fixed pattern. The latter fact is quite noticeable in any Austronesian languages and is a characteristic of Austronesian languages, while Old Japanese has a free accent pattern, as reconstructed from the accent pattern of the words recorded in the *Ruiju Myogisho*.

#### 4. Words in Old Japanese regarded as cognates with those in Austronesian

We have selected the following words on the basis of the regular phonological correspondence between Old Japanese and Austronesian. However, if the word is not listed in the RM, then we have taken it from consideration.

##### Regular Phonological Correspondences between Old Japanese and Austronesian

PAN	PJ	PF	PWMP(>PIN:PPH)	PCOMP	PAN	PJ	PF	PWMP(>PIN:PPH)	PCOMP
*p	*p	*p/*x	*p	*p	*h	*C/*O	*	*h	*h
*b	*p/*m	*p	*b	*b/*p	*m	*m	*m	*m	*m
*t	*t/*0	*t	*t/*l	*t/*n	*n	*n	*n	*n	*n
*d	*s	*l	*d	*d	*n	*n/*O	*n	*n	*n /*R
*s	*s/*t/*ts	*s/*ts	*s/*D	*s	*l/*t	*r	*l	*d/*N	*l
*z	*t	*	*d/*D	*	*w	*w	*	*w	*w
*k/*R	*k	*k	*k/*R	*k	*O	*y	*	*R	*R
*g	*k/*C	*g/*R	*g/*d/*R/*j	*R					

We need to establish the regular sound correspondence between Proto-Japanese and Proto-Austronesian because the regular phonological correspondences among Austronesian languages have been established. As shown in the table above, we are able to establish the correspondences between Japanese and Austronesian by listing all the initial consonants when an Old Japanese word is compared with the corresponding Austronesian word.

#### 5. Selection of words in Japanese and Austronesian

We have chosen the following words on the basis of correspondence in form and meaning between Japanese and Austronesian, but the only nouns with a pitch recorded in the RM are considered here simply because the ones without a pitch cannot be compared with the corresponding Austronesian words.

**Austronesian (monosyllabic) : 10 counts**

OJ <i>pa</i> 'tooth' HL	OJ <i>ki</i> <sub>1</sub> 'molar' LL	OJ <i>pi</i> <sub>2</sub> / <i>po</i> <sub>2</sub> - 'fire' LL
OJ <i>to</i> <sub>1</sub> 'outside' HH	OJ <i>te</i> <sub>2</sub> / <i>ta</i> - 'hand' LL	OJ <i>tu</i> 'saliva' LL
OJ <i>si</i> 'he/it/they' HH	OJ <i>ke</i> <sub>2</sub> / <i>ka</i> - 'food' LL	OJ <i>ki</i> <sub>2</sub> / <i>ko</i> <sub>2</sub> - 'tree' LL
OJ <i>pa</i> 'leaf' HL		

**Austronesian (disyllabic) : 23 counts**

OJ <i>kuti/kutu</i> 'mouth' HH	OJ <i>kura-si</i> 'dark' HHhl	OJ <i>pana</i> 'nose' HH
OJ <i>pana</i> 'flower' LL	OJ <i>pane</i> 'feather' HH	OJ <i>pi</i> <sub>1</sub> <i>to</i> <sub>2</sub> 'person' HL
OJ <i>posi</i> 'star' HH	OJ <i>po</i> <sub>(2)</sub> / <i>so</i> <sub>(2)</sub> 'navel' HH	OJ <i>puku</i> 'lungs' HH
OJ <i>puku-mu</i> 'include' LLhl	OJ <i>pukuro</i> 'bag' LLL	OJ <i>puku-ru</i> 'swell' HHL
OJ <i>puta/pata</i> 'two' HH	OJ <i>sisi</i> 'meat' LL	OJ <i>sita</i> 'tongue' LL
OJ <i>take</i> <sub>2</sub> 'mountain' LH	OJ <i>tana</i> 'heaven,sky' HL	OJ <i>tane/tana</i> - 'seed' LH
OJ <i>titi</i> 'breast' HH	OJ <i>to</i> <sub>2</sub> <i>mo</i> <sub>2</sub> 'companion' HL	OJ <i>tuma</i> 'husband, wife' HL
OJ <i>tume</i> <sub>2</sub> / <i>tuma</i> - 'fingernail' HH	OJ <i>tuno</i> <sub>1</sub> 'horn' LL	

**6. Correspondence between tone and voicing**

Now, the correspondences can be classified into three categories: (1) OJ voiceless (H/L) vs. Aus. voiceless, (2) OJ voiceless (H/L) vs. Aus. voiced, (3) OJ voiced (H/L) vs. Aus. voiced. The first two categories will be the primary consideration, although the third category will be briefly examined.

Now, the first category will be considered:

The tone at the beginning of a word is always H with either H or L at the following mora. In addition to the regular phonological correspondence between them, the corresponding Austronesian consonant at the beginning of a word is always voiceless.

**Category 1: OJ voiceless (H) vs. Aus. voiceless**

- OJ *pa* (HL) 'leaf' < PJ \**pa-Ca*: PIN \**pápaq* 'plant materials';  
*palapa?* 'spiked leaf' in Tagalog;  
*papaq* 'the shape of a coconut leaf' in Javanese;  
*pálapaq* 'the shape of a coconut leaf, leaf' in Malay;  
*palapaq* 'a coconut leaf' in Ngadju-Dayak;  
*falafa* 'hemp veins' in Hova;  
*papaq* 'a leaf' in Ami



2. OJ *p̄na*(HH) 'nose' < PJ \**pana*(HH) 'nose' : PAN \**páya* 'forked projection';  
PWMP \**paŋa* 'chin (bone)'; PCEMP \*(*m*)*paŋ a* 'branch';  
*páya* 'chin' in Tagalog, Panpagan, .  
*mága* 'branch' in Samoan  
*mana* 'branch' in Hawaiian  
*mánga* 'branch' in Tongan  
*máʔa* 'branch' in Tahitian
  
3. OJ *p̄a-ne*(HH) 'feather' < PJ \**pa-ni*(HH) : PCEMP \**pan(i)* 'feather'; PAN \**pánij* 'wing'  
*paneʔ* 'wing, feather' in Bunun;  
*panid* 'wing, feather' in Yami,  
*paniʔ* 'wing' in Uma,  
*panniʔ* 'wing' in Bugis,  
*pani* 'wing' in Wolio
  
4. OJ *pi<sub>1</sub>to<sub>2</sub>* (HL) < POJ \**pi-to* < PJ \**pi-<sub>1</sub>tau*: PAN \**Tau*;  
*tau* 'person' in Puyuma;  
*tau* 'person' in Yami
  
5. OJ *p̄o<sub>(2)</sub>so<sub>(2)</sub>/peso<sub>(2)</sub>* (HH) 'navel' < PJ \**pusu*(HH) : PMP \**púsej* 'navel';  
PEMP \**pito* 'navel';  
*pūsod* 'navel' in Tagalog  
*pusud* 'navel' in Kagayanen  
*pusod* 'navel' in Murut  
*pusoh* 'navel' in Bunun,  
*pəʂəD* 'navel' in Yami,  
*pudək* 'navel' in Paiwan  
*pusat* 'navel' in Indonesian  
*poso-* 'navel' in Kilivila
  
6. OJ *p̄uta*(*puta-tu/puta-ye* HH-L)/*pat̄a* 'two' < PJ \**pata*(HH): PWMP \**páCan* 'pair';  
*pasan* 'pair' in Malay(=Indonesian), Acehnese, Sasak  
*sa-pasay* 'pair' in Toba-Batak, Sundanese, Minangkabau  
*sapasay* 'pair' in Madurese,  
*si-pasay* 'pair' in Bugis

10 **Yoshizo Itabashi**

7. OJ *titi*(HH) 'breast' < PJ \**titi*(HH): PAN \**súsu* 'breast' + \**ñúñu* 'breast' PIN \**susu* 'breast';  
*xuxu?* 'breast' in Atayal  
*tutu* 'breast' in Paiwan  
*suso* 'breast' in Tagalog
8. OJ *to<sub>2</sub>mo<sub>2</sub>*(HL) 'companion' < PJ \**tama*: PAN \**taman* 'companion'; PPH \**duma* 'companion';  
*toman* 'celibacy' in Toba-Batak;  
*təmən* 'loyalty' in Javanese  
*tuman* 'familiarity' in Javanese;  
*təman* 'friend, partner' in Malay,  
*toman-a* 'to accompany' in Fijian
9. OJ *tuma* (HL) 'husband, wife' < PJ \**tama* : the same as OJ *to<sub>2</sub>mo<sub>2</sub>* 'companion'  
PAN \**taman* 'companion'; PPH \**duma* 'companion';  
*toman* 'celibacy' in Toba-Batak;  
*təmən* 'loyalty' in Javanese  
*tuman* 'familiarity' in Javanese;  
*təman* 'friend, partner' in Malay,  
*toman-a* 'to accompany' in Fijian

As the phonotactic rule predicts, the tone at the beginning of an OJ word is always H with either H or L at the following mora and the corresponding Austronesian consonant at the beginning of a word is always voiceless, in addition to the regular phonological correspondence between them. We have nine examples of this case, though we cannot say that it is a large number.

However, it is extremely important to have such a category because those examples almost prove that there was a strong (but rather distant) genealogical connection between Japanese and Austronesian, not to mention that Japanese has prefixes which have semantic and functional similarities with those of Austronesian (Sakiyama 2010:26-27), and which have nothing in common with Altaic or any other neighboring languages or language families.

Now, we will look at the second category, which is paired with the first category. The tone at the beginning of a word is always L with either H or L at the following mora and the corresponding Austronesian consonant at the beginning of a word is always voiceless, in addition to the regular phonological correspondence between them.

**Category 2: OJ voiceless (L) vs. Aus. voiceless**

1. OJ *ke<sub>2</sub>/ka*-(LL) 'food' < PJ \**ka-Ci* : PAN \**ka(n)* 'to eat', PWMP \**kan/kai* 'to eat', PCEMP \**kani* 'to eat';  
*kain* 'food' in Tagalog;  
*da-han* 'boiling rice' in Toba-Batak;  
*k-un-an* 'to eat' in Ngadju-Dayak;  
*kan-a* 'to eat' in Fijian;  
*kai* 'to eat' in Tongan, Futuna
2. OJ *ki<sub>2</sub>/ko<sub>2</sub>*-(LL) 'tree' < PJ \**kaCu-Ci*: PAN \**káquy* 'tree';  
*kaqoy* in Tagalog;  
*kayo* in Ilokano,  
*kaquy* in Atayal
3. OJ *pi<sub>2</sub>/po<sub>2</sub>*-(LL) 'fire' < PJ \**paCu -Ci*: PAN \**xapuy* 'fire'; PF \**qapuy*;  
*puzu* in Tso;  
*puniq* in Sedek; PWMP  
 (\**api* in Ainu)
4. OJ *si* (LL) 'he/it/they' < PJ \**si*: PCEMP \**si* 'they';  
*si* 'they(animate)' in Biak, Sio  
*si* 'we' in Trukese
5. OJ *sisi* (LL) 'meat' < PJ \**sisi*: PAN \**Sesi* 'meat' PIN \**isi* 'content';  
*isi* 'content' in Javanese, Malay  
*isi* 'meat' in Ngadju-Dayak  
*siə* 'meat' in Acehnese  
*isi-n* 'meat' in Buru  
*?asiasi?* 'meat' in Yami
6. OJ *take<sub>2</sub>*(LH) 'mountain' < PJ \**tsaka-Ci* (HL): PAN \**sá:kay* 'to ascend' ; PMP \**sákay* 'to ascend',  
*zake* 'upward' in Fijian  
*qake* 'upward' in Tongan

12 **Yoshizo Itabashi**

7. OJ *tane* (LH) /*tana*- 'seed' < PJ \**tana-Ci*: PAN(PMP?) \**tánem* 'seed'  
*ta'nəm* in Kagayanen,  
*tanom* in Bangingi Sama,  
*-tanom* in Murut,  
*Tamən* in Madurese,  
*natan* in Iraputu,  
*tano* in Manam,  
*tanu* in Tahitian
8. OJ *te<sub>2</sub>*(LL)/*ta*- 'hand' < PJ \**ta-Ci* (LL): PAN \**tánan* 'hand';  
*tanan* 'hand' in Javanese, Malay(=Indonesian), Ngadju-Dayak;  
*tanana* 'hand' in Hova;  
*tana* 'hand' in Fijian
9. OJ *tuno<sub>1</sub>* (LL) 'horn' < PJ \**tsunu* (LL): PWMP \**tsuRe* 'horn', PHF \**tsúRey* 'horn'  
( R=[Y] )  
*súju* in Tso, Javanese  
*taju<sub>kiy</sub>* in Atayal,  
*tə<sub>kuj</sub>* in Paiwan,  
*súujay* in Tagalog, Aklanon  
*tánduk* in Palawan, Indonesian, Sundanese, Balinese  
*tónu?* in Uma  
*tánru?* in Bugis, Konjo  
*sunar* 'light' in Javanese, Madurese

These nine OJ examples would expect the word-initial voiced consonant of the Austronesian words above, but they, instead, have the voiceless consonant in Austronesian. They appear to be exceptions of the phonotactic rule. Or these examples seem to disprove the OJ phonotactic rule above, but as long as the rule is working in OJ which may extend to Austronesian, there might be some explanation for this category. At this point, we cannot say more on this. This is our problem to solve in the near future.

Again, although we have the problem on this category, the rule still holds true for the first category, which is more important for this type of diachronic linguistic problem.

Consider the following category. The tone at the beginning of a word is always L

with either H or L at the following mora and the corresponding Austronesian consonant at the beginning of a word is always voiceless, in addition to the regular phonological correspondence between them.

**Category 3: OJ voiceless(L) vs. Aus. voiced**

1. OJ *ki* (LL) 'molar' < PJ \**ki*/\**gi*: PIN \**gigi* 'molar';  
*gigi* 'the tip of a pick' in Toba-Batak;  
*gigi* 'teeth' in Malay;  
*hihi* 'the gum' in Hova
  
2. OJ *pa* (Lhl?) 'tooth' < PJ \**pa-Ca*: PIN \**baRang* 'mortar tooth';  
*bagang* 'mortar tooth' in Tagalog;  
*bagʔang* 'mortar tooth' in Bikol, Hiligaynon
  
3. OJ *pana* (LL) 'flower' < PJ \**pana* 'flower': PA \**bunah* 'flower';  
*bəyəlay* in Rukai  
*buyəŋ* in Acehnese,  
*buya* in Indonesian, Toba-Batak, Da'a,  
*buyə* in Balinese,  
*buna-k* in Roti,  
*fuya* in Samoan
  
4. OJ *puku* (LL?) 'lungs' (same root as OJ *puku-ro* 'a bag', OJ *puku-mu*(LL-hl) 'to include', OJ *puku-ru*(HH-L) 'to swell', J *Fugu* 'blow fish', J *Fuguri* 'scrotum', J *Fukura-Fagi* 'calves (of leg)' < PJ \**puku-* 'something bulged' ; PJ *puku-ro* (LL-L) 'bag': PAN \**buku* 'swell, a joint'; PEMP \**puku* 'protrusion, mound';  
*buku* 'hump, lump' in Malay,  
*buku* 'swell' in Hova,  
*buko* 'flower bud' in Fijian  
*buko* 'something bulged' in Fijian  
*mbuku* 'humpy' in Fijian,  
*pu'u* 'hill, mound, pimples, bumps, heart, stomach, fist' in Hawaiian;  
(*puhoa* 'lungs' in Middle Korean)

14 **Yoshizo Itabashi**

5. OJ *sita* (LL) 'tongue' < PJ *\*tsita* : PIN *\*dilaq/d'ilaq* 'tongue':  
*dilaʔ* 'tongue' in Tagalog, Aklanon  
*dilaʔ* 'tongue' in Palawan, Molbog, Kagayanen, Murut  
*lidah* 'tongue' in Malay, Indonesian  
*lidamə* 'tongue' in Rukai,  
*lidam* 'tongue' in Puyuma
6. OJ *tu* (LL) 'saliva' < PJ *\*tu*: PAN *\*ludaq* 'saliva' *\*luvaq* 'to spit':  
*luraq* 'saliva' in Tagalog  
*luwaq* 'to spit' in Tagalog;  
*ludaq* 'saliva' in Malay  
*luaq* 'to spit' in Malay;  
*rura* 'saliva' in Hova  
*lua* 'to spit' in Hova;  
*lua* 'to vomit' in Fijian, Tongan

As the phonotactic rule predicts, the tone at the beginning of a word is always L with either an hl or a L at the following mora and the corresponding Austronesian consonant at the beginning of a word is always voiced, other than the regular phonological correspondence between them. We have only six examples of this category. However, we find it extremely important to have such a case because those examples may lead to our postulation that there is a strong genetic relationship between Japanese and Austronesian.

Now, we will consider the fourth category, paired with the third category. The tone on the word-initial position is always H with an H on the second mora. In addition to the regular phonological correspondence between them, the corresponding consonant at the beginning of the Austronesian word is always voiced.

**Category 4: OJ voiceless(H) vs. Aus. voiced**

1. OJ *kura-si* (HHhl) 'dark' < PJ *\*kura-*: PIN *\*gəlap* 'darkness';  
*golap* 'darkness' in Toba-Batak;  
*gə lap* 'dark, darkness' in Malay
2. OJ *kuti* (HH)/*kutu* 'mouth' < PJ *\*kutu-Ci* (HH): PAN *\*ɲusu* 'mouth',  
*ɲoso* 'mouth' in Yami,  
*ɲuzu* 'mouth' in Roviciana,  
*ɲusu-* 'mouth' in Eastern Fijian,

- ɲuhu* 'mouth' in Western Fijian,  
*nucu* 'mouth' in Rotuman,  
*ɲutu* 'mouth' in Tongan, Samoan, Mele-Fila;  
*gutu* 'lips' in Mele-Fila,  
*ʔutu* 'lips' in Tahitian,  
*ɲutu* 'lips' in Rapanui
3. OJ *posi* (HH) 'star' < PJ \**pəsi* (HH) : PAN \**bi-(n)tuqen* 'star'; PWMP \**bintan* 'star';  
 PCEMP \**pituqu* 'star',  
*bituwīn* in Tagalog,  
*vicuk-an* in Paiwan,  
*bintan* in Acehnese, Indonesian, Madurese;  
 (*pyol* 'star' in Middle Korean)
4. OJ *tana* (HH) 'heaven, sky' < PJ \**tana* (HH): PWMP \**lanit*  
*lāyit* 'heaven' in Tagalog, Isnag, Kalinga Linimos,  
*layit* 'heaven' in Kagayanen, Palawan, Molbog, Sarangani Blaan  
*laji* 'heaven' in Burut
5. OJ *to<sub>1</sub>*(HH) 'outside' < PJ \**tu*: PIN \**luwal* 'outside';  
*luwal* in Tagalog;  
*ruar* in Toba-Batak;  
*luwar* in Javanese;  
*luar* in Malay;  
*ruar* in Ngadju-Dayak
6. OJ *tume<sub>2</sub>* (HH) /*tuma-* 'fingernail' < PJ \**tuma-Ci* : PIN \**Du(m)put* 'to pick up  
 with fingers';  
*dampot* 'to pick up with fingers' in Tagalog;  
*damput* 'to pick up with fingers' in Kapanpagan;  
*Du(m)put* 'to pick up with fingers' in Javanese;  
 (*top* 'fingernails' in Middle Korean)

Unlike the phonotactic rule predicts, the tone at the beginning of an word is always H, instead of L (the third category, paired with the fourth), with either an H or a L at the following mora and the corresponding consonant at the beginning of an Austronesian word is always voiced.

Now the fifth category will be considered. The tone at the beginning of an OJ word with a **C<sup>vd</sup>**- or a **V**- is always L with either an H or a L at the following mora and the corresponding consonant or vowel at the beginning of an Austronesian word is always voiced.

**Category 5: OJ voiced(L) vs. Aus. Voiced**

1. OJ **a (LL)** 'I' < PJ **\*a** : PAN **\*aku** 'I'  
*a* 'I' in Barim, Kagayanen, Sio  
*au* 'I' in Toba-Batak, Roti, Tongan, Rapanui  
*ya* 'I' in Sawai
  
2. OJ **me<sub>2</sub>/ma- (LL)** 'eye' < PJ **\*ma-Ci**: PAN **\*mata** 'eye';  
*maca* in Paiwan, Rukai,  
*mata* in Yami, Puyuma, Ami, Bunun Palawan, Tagalog, Bikol, Bisaya, Pangasinan, Manam,  
Mbula, Wolio, Tongan, Samoan, Tahitian,  
*maa* in Mekeo, Lau,  
*maa(-na)* in Kwaio
  
3. OJ **me<sub>1</sub> (LH)** 'female' < PJ **\*bei** < PPJ **\*baya** : PAN **\*baya** 'female';  
*ababay* 'female' in Rukai  
*va-vay-an* 'female' in Paiwan  
*baya* 'mother' in Tagalog;  
*bai* 'female' in Kagayanen  
*bai* 'grandmother' in Pangasinan;  
*bayi* 'baby' in Javanese
  
4. OJ **mimi (LL)** 'ear' < PJ **\*<sup>m</sup>bi<sup>m</sup>bi (LL)** : PAN **\*bi+biR** 'lips';  
*bibian* 'lips' in Rukai,  
*bibig* 'lips' in Isnag,  
*bibi* 'lips' in Achenese,  
*bibir* 'lips' in Indonesian, Madurese,  
*BiBi* 'lips' in Uma,



- vivi* 'lips' in Da'a,  
*bibig* 'mouth' in Tagalog;  
*bibir* 'lip, rim' in Malay
5. OJ *na* (LL) 'I, you' < PJ \**na*: PEMP \**na* 'he/she/it';  
*na* 'I' in Siassi;  
*na* 'he/she/it' in Bare'e
6. OJ *nani* (LH) 'what' < PJ \**n-ani*: PIN \**anu* 'what, something';  
*ano* 'what' in Tagalog  
*anuh* 'what' in Aklanon  
*anu* 'what' in Malay, Palawan, Kagayanon  
*onu* 'what' in Molbog  
*nanu-an* 'what' in Atayal  
*a-nəma* 'what' in Paiwan
7. OJ *ne*<sub>2</sub> (HL) 'sound' < PJ \**na-Ci*: PAN \*(*m*)*N-tajiT* 'to cry';  
*taji* 'to sound, sound(n)' in Nggela;  
*kani* 'sound, to ring, to cry' in Hawaiian
8. OJ *no*<sub>2m</sub>- (Lhl) 'to drink' < PJ \**naCum*-: PMP \**inum*- 'to drink';  
*inom* in Yami, Tagalog,  
*inum* in Kalinga Limos, Madurese,  
*minum* in Palawan, Toba-Batak, Indonesian,  
*num* in Buang,  
*num*- in Adzera,  
*-nom* in Yabem
9. OJ *wi*<sub>2</sub> (LL) 'water' < POJ \**wei* < PJ \**wayi*; PAN \**way* v; PWMP \**wai*;  
PCEMP \**wai*(R);  
*aek* in Toba-Batak;  
*we-da* 'hot water' in Javanese;  
*air/ayer* 'water' in Malay;

18 **Yoshizo Itabashi**

*wai* 'water' in Maori, Fijian;

*vai* 'water' in Futuna, Samoan, Tongan;

*wi* 'water' in Lewo

10. OJ *wo*<sub>1</sub> (LL) 'tail' < PJ \**wu-Ca* : PAN \**uRat* 'blood vessel, tendon';

*urat/ot-ot* 'blood vessel, tendon, muscle' in Javanese

*urat* 'blood vessel' in Malay;

*ugat* 'blood vessel' in Tagalog

11. OJ *yo*<sub>1</sub> (LL) 'night' < POJ \**yo<sub>1</sub>-pi* < PJ \**yupi* 'dusk': PAN \**Rabi?iH* 'night'

PWMP \**Rabi* 'night'; PCEMP \**Rabi* 'night';

*Gabi* in Palawan,

*gabi* in Tagalog, Isnag,

*gobi* in Molbog

The tone at the beginning of an OJ word with a **C<sup>vd</sup>**- or a **V-** is always L with either an H or a L at the following mora and the corresponding consonant or vowel at the beginning of an Austronesian word is always voiced. This category is beyond the scope of the rule, but this correspondence between the two can be understood when compared with the rule for the third category: (the third category) the LOW tone of an OJ word-initial VOICELESS consonant corresponds to the Austronesian word-initial VOICED consonant, whereas (the fifth category) the LOW tone of an OJ word-initial VOICED consonant corresponds to the Austronesian word-initial VOICED consonant (the fifth category).

The fifth category is more explicit than the third category in that the fifth category is based more on the natural tendency of the tonal association that a VOICED consonant or vowel is generally associated with a LOW tone rather than a high, and so a LOW tone tends to correspond to the VOICED consonant or vowel of a COGNATE word when both go back to the common source. This notion is supported by the high number of examples in this category (73%: 11 counts out of 15, i.e. the added number of examples of the fifth with the sixth category: both categories have OJ VOICED (H/L) vs. Aus. VOICED in common) in addition to the regular correspondence between them.

Now, we will consider the last category which is paired with the fifth category: the tone at the beginning of an OJ word with a **C<sup>vd</sup>**- or a **V-** is always H with an H at the following mora and the corresponding consonant or vowel at the beginning of the PAN

word is always voiced.

**Category 6: OJ   voiced(H) vs. Aus. Voiced**

1. OJ *mi*<sub>2</sub> (HH) 'body' < PJ \**mi*: the same as *mi* 'fruit' PAN \**bu'aq* 'fruit';  
*buai /bui* 'fruit' in Atayal;  
*boa/boway* 'fruit' in Paiwan;  
*bu*a 'fruit' in Toba-Batak;  
*bu*aq 'fruit' in Malay;  
*vua* 'fruit' in Fijian;  
*fua* 'fruit' in Tongan, Futuna, Samoan
  
2. OJ *mi*/*mu*- (HH) 'fruit' < PJ \**mu-Ci*: PAN \**bu'aq* 'fruit';  
*buai /bui* 'fruit' in Atayal;  
*boa/boway* 'fruit' in Paiwan;  
*bu*a 'fruit' in Toba-Batak;  
*bu*aq 'fruit' in Malay;  
*vua* 'fruit' in Fijian;  
*fua* 'fruit' in Tongan, Futuna, Samoan
  
3. OJ *mi*<sup>n</sup>*do*<sub>2</sub>*ri* (HHL) 'green' < PJ \**mi*<sup>n</sup>*daCuri* 'green': PIN \**qi**Dau* 'green';  
*ijo* in Acehnese, Javanese,  
*qijau* in Indonesian,  
*qejo* in Sundanese,  
*qijo* in Wolio,  
*qilao* in Tagalog
  
4. OJ *wo* (HH) 'male' < PJ \**u-Ca* : PAN \**uRa* 'young man';  
*wo* 'person' in Javanese;  
*ora* 'person' in Malay;  
*uran* 'child' in Toba-Batak;  
*ule-* 'brothers and sisters' in Urawa;

This category, paired with the fifth category, has only four examples. This category is beyond the scope of the rule, unlike the second or fourth category. Thus, we do not pursue it further.

### Conclusion

Almost all the putative matching nouns found in both Japanese and Austronesian have been considered according to the phonotactic rules mostly based on the tonogenesis promulgated by James Matisoff.

We are now able to answer the question raised as the topic ‘Is tone distinction of the word-initial mora of an Old Japanese word correlated with voicing distinction of the word-initial consonant of a putative Austronesian cognate?’; as a result of its application of the phonotactic rule to the H/L tone of an OJ word-initial voiceless/voiced noun, the putative corresponding Austronesian word-initial consonant is either voiceless or voiced:

#### 1) Category 1

As the phonotactic rule predicts, the word-initial HIGH tone with the word-initial VOICELESS consonant of an OJ word corresponds to the word-initial VOICELESS consonant of an Austronesian word, which is based on the regular phonological correspondence between them.

$$\begin{array}{ccc} \text{OJ} & & \text{Austronesian} \\ \text{C}^{\text{vls}}\text{V-}(\underline{\text{H-}}) & = & \text{C}^{\text{vls}}\text{V-} \end{array}$$

#### 2) Category 3

As the phonotactic rule predicts, the word-initial LOW tone with the word-initial VOICELESS consonant of an OJ word corresponds to the word-initial VOICED consonant of an Austronesian word, which is based on the regular phonological correspondence between them.

$$\begin{array}{ccc} \text{OJ} & & \text{Austronesian} \\ \text{C}^{\text{vls}}\text{V-}(\text{L-}) & = & \text{C}^{\text{vd}}\text{V-} \end{array}$$

#### 3) Category 5

This category is beyond the prediction of the phonotactic rule. But this correspondence between the two can be understood clearly when this category is compared with the third category; the third category states that the LOW tone of an OJ word-initial VOICELESS consonant corresponds to the Austronesian word-initial VOICED consonant, whereas the fifth category says that the LOW tone of an OJ word-initial VOICED consonant corresponds to the Austronesian word-initial VOICED consonant. In other words, the tone at the beginning of an OJ word with a Cvd- or a V- is always L with either an H or a L at the following mora and the corresponding consonant or vowel at the beginning of an Austronesian word is always voiced.

$$\begin{array}{ccc} \text{OJ} & & \text{Austronesian} \\ \text{C}^{\text{vd}}\text{V- (L-)} & = & \text{C}^{\text{vd}}\text{V-} \end{array}$$

The first two categories (category 1 and 3) with the fifth category match the predictions of the rule, which indicates that the rule functions as a classifier into the two separate conditions of the rule shown in the categories 1 and 3.

But those nouns in these categories seem to function as cognates, which is more important than anything else in comparative linguistics because those ones, even a small number of counts may be able to clarify the genealogical relationships between Japanese and Austronesian.

#### 4) Other categories (2,4,6)

These categories are not predicted by the rule, although the number of these examples (6 counts) is the same as those of categories 1 and 3. The number of the examples in the sixth category is only four counts, when compared with the number of the fifth category (11 counts; 72%) being predicted as a tonal tendency explained above. These categories need to be studied further in the near future.

Finally, it is worth noting that Japanese and Austronesian have a common variety of prefixes, which are treated as morphology and syntax as opposed to phonology, and which are normally not borrowed, as well as the regular phonological correspondences and much common basic vocabulary.

## References

- Edmondson, Jerold A. and Kenneth J. Gregerson (eds.) (1993) *Tonality of Austronesian languages*, Honolulu: University of Hawaii Press.
- Hayata, Teruhiro. (1999) *Onchoo no taiporojii (Linguistic typology of tones)*, Tokyo: Taishukan Shoten.
- Haudricourt, Andr -Georges. (1954) *De l'origine des tons du vietnamien*, Journal Asiatique 242: 69-82.
- Kitahara, Yasuo, et al. (2000) *Nihon Kokugo daijiten (Japanese dictionary)*, Tokyo: Shogakkan.
- Matisoff, J. (1970) Glottal dissimilation and the Lahu high-rising tone: A tonogenetic case-study. *Journal of the American Oriental Society* 90 (1): 13-44.
- Matsumori, Akiko. (2008) On the reconstruction of the proto-accentual system of Japanese, Frellesvig, Bjarke & John Whitman (eds.) *Proto-Japanese Issues and Prospects*, Amsterdam:

- John Benjamins Publishing Company. 103-124.
- Ramsey, Robert. (1990) *Nihongo oyobi Choosengo ni okeru akusento no kigen* (The origin of pitch accent in Japanese and Korean), Osamu Sakiyama (ed.) *Nihongo no Keisei (The Formation of the Japanese Language)* 218-228.
- Sakiyama, Osamu. (2010) Nihongo no keiseikatei to gengosesshoku (The formation of Japanese and language contact), Katsumi Shibuya (ed.) *Gengosesshoku no Sekai (The Realm of Language Contact)* Special Issue, Vol. 29, No. 14, 18-31. Meijishoin.
- Shimabukuro, Moriyo. (2008) A reconstruction of proto-Japanese accent for disyllabic nouns, Frellesvig, Bjarke & John Whitman (eds.) *Proto-Japanese Issues and Prospects*, Amsterdam: John Benjamins Publishing Company. 125-139.
- Tadmor, Uri, Haspelmath, Martin, and Taylor, Bradley. (2010) Borrowability and the notion of basic vocabulary, *Diachronica, Special issue: Quantitative Approaches to Linguistic Diversity* Vol.27, No.2, 226-246.
- Tryon, Darrell (ed.) (1995) *Comparative Austronesian dictionary, Part 1 ~ Part 4*, Trends in Linguistics Documentation 10, New York: Mouton de Gruyter.
- Zorc, David R. (1978) Proto-Philippine word accent: innovation or Proto-Hesperonesian retention? *Papers of the Second International Conference on Austronesian Linguistics. Fascicle I*. Pacific Linguistics C-61. Canberra: The Australian National University, 67-119.
- Zorc, David R. (1983) Proto-Austronesian revisited, *Philippine Journal of Linguistics* 14:1-24.
- Zorc, David R. (1993) II: Overview of Austronesian and Philippine accent patterns, Edmondson, Jerold A. and Kenneth J. Gregerson (eds.) *Tonality of Austronesian Languages*, Honolulu: University of Hawaii Press, 17-24.