



# A model of the origin of Kra-Dai tones

Laurent Sagart

► **To cite this version:**

Laurent Sagart. A model of the origin of Kra-Dai tones. Cahiers de Linguistique Asie Orientale, 2019, 48 (1), pp.1-29. 10.1163/19606028-04801004 . hal-02179092

**HAL Id: hal-02179092**

**<https://hal.archives-ouvertes.fr/hal-02179092>**

Submitted on 10 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## A model of the origin of Kra-Dai tones

*Laurent Sagart*

CNRS–EHESS–INALCO, CRLAO

*laurent.sagart@gmail.com*

### Abstract

This paper finds origins for the three Kra-Dai tones in the segmental endings of Proto-Southern Austronesian, the parent language of Kra-Dai and Malayo-Polynesian. The Kra-Dai A category originates in sonorant endings (vowels, semi-vowels, nasals, liquids) and in Proto-Austronesian \*-H<sub>2</sub>, reconstructed by Tsuchida 1976; the B category in \*-R and in \*-X, a hitherto not reconstructed ending reflected as *-h* in Amis and in the Bisayan language Aklanon; the C category, in Proto-Austronesian \*-H<sub>1</sub>, reconstructed by Tsuchida. The tonal outcomes of \*-s and \*-S are described. Kra-Dai sonorant endings in tone C are argued to come from hypothetical Austronesian prototypes in which a sonorant ending was followed by \*-s, a suffix of unknown function. Although the present model does not require Kra-Dai to be a daughter of Proto-Austronesian, the building blocks for Kra-Dai tones are shown to be in place during the Formosan phase of Austronesian phonological history.

### Keywords

Kra-Dai — Tai-Kadai — Austronesian — tones — historical phonology — tonogenesis

### Résumé

Cet article propose que les trois tons kra-dai proviennent des segments terminaux du proto-austronésien du sud, ancêtre du kra-dai et du malayo-polynésien. La catégorie tonale A du kra-dai provient des terminaisons sonantes (voyelles, semi-voyelles, nasales, liquides) et du phonème \*-H<sub>2</sub> reconstruit par Tsuchida 1976; la catégorie B provient de \*-R final et de \*-X, une final reconstruite ici pour la première fois, et reflétée comme *-h* en Aklanon; la catégorie C provient du phonème \*-H<sub>1</sub> reconstruit par Tsuchida. L'article décrit également les reflets tonaux divers des consonnes finales \*-s et \*-S du proto-austronésien du sud. En outre, l'hypothèse est proposée que les mots kra-dai se terminant par des sonantes au ton C proviennent de terminaisons sonantes suivies d'un suffixe *-s*, de fonction inconnue, en proto-austronésien du sud. Même si le présent modèle ne requiert pas que le kra-dai soit une langue issue du proto-austronésien, on souligne le fait que les consonnes finales à la base des tons kra-dai sont en place pendant la phase formosane de l'histoire phonologique austronésienne.

## Mots-clés

kra-dai — tai-kadai — austronésien — tons — phonologie historique — tonogène

### 1 Introduction

Like Old Vietnamese, Written Burmese and Proto-Hmong-Mien, the Kra-Dai languages exhibit a phonological typology strongly influenced by a form of Chinese which existed between c. 200 BCE, when Qín Shǐ Huáng's conquering armies for the first time brought Chinese into contact with Kra-Dai on a large scale, and c. 700 CE, in the beginning part of the Tang dynasty. Chinese of that period possessed a system of three true tones, not yet split into an upper and a lower series according to the voicing of the main syllable's initial consonant. The Proto-Kra-Dai tone system is structurally identical to that of Chinese of that time, as are those of Vietnamese, Burmese and Proto-Hmong-Mien: three contrasting tones on words ending in vowels and sonorants (semi-vowels, nasals, liquids) while words ending in oral stops do not show any tonal contrasts. The origin of the three-tone contrast was elucidated for Vietnamese by Haudricourt (1954b): one tone originated in words ending in [h], another in [ʔ], a third in vowels or sonorants. Haudricourt (1954a) also showed that the Chinese Departing tone, or *qùshēng* 去聲, originated in an \*-s suffix which evolved to [h] before becoming a tone, as in his scheme for Vietnamese. Pulleyblank (1962: 225), then Mei (1970) completed Haudricourt's picture for Chinese by proposing that the Chinese Rising tone, or *shǎngshēng* 上聲, goes back to words ending in a glottal stop.<sup>1</sup> The origins of the Burmese, Hmong-Mien and Kra-Dai tones have not been elucidated. Important observations were made for Kra-Dai by Ostapirat (2005) and Norquest (2013). Here I offer additional hypotheses which bring us closer to a solution. The present model is in large part a by-product of my ongoing study of the Kra-Dai language Buyang 布央 (Lángjià 郎架 dialect), described by Lǐ (1999). Among Kra-Dai languages, Buyang is especially conservative in preserving the first syllables of Austronesian words (Sagart 2004), albeit in reduced form. This helps in working out the sound correspondences between Buyang and Austronesian, and in turn increases the reliability of Buyang-Austronesian cognates. Wherever possible, Buyang forms are used in this paper, and specifics for each form are described.

The four Kra-Dai tone categories are known as A, B, C and D. The core of these categories existed in Proto-Kra-Dai (Ostapirat 2005), probably more as

<sup>1</sup> For details see Sagart (1999).

categories characterized by different kinds of consonantal endings than as contrastive melodic contours. When borrowed into Thai (a.k.a. Bangkok Siamese) and other Kra-Dai languages of the Tai and Kam-Sui branches, the most directly exposed to Chinese influence, Middle Chinese *píngshēng* 平聲 words (ending in sonorants in Old Chinese) typically show tone A; Middle Chinese *shǎngshēng* words (ending in a glottal stop in Old Chinese) are treated by tone C; MC *qùshēng* words (ending in \*-h in late OC < early OC \*-s) by tone B. Chinese words ending in oral stops (\*-p, \*-t, \*-k) form the *rùshēng* 入聲 “Entering tone” category. That category is primarily defined segmentally, but Chinese linguistic tradition calls it a tone. Kra-Dai “tone” D corresponds to the Chinese *rùshēng*. Gedney (1986) suggests that the original phonetic values of the Thai tones A, B and C were the same as the late Old Chinese values of the Chinese *píngshēng*, *qùshēng* and *shǎngshēng* respectively under the Haudricourt-Pulleyblank-Mei solution: sonorant endings, [h]-endings, [ʔ]-endings. Subsequent attempts at solving the origin of Kra-Dai tones have assumed Gedney’s hypothesis.

Benedict (1942) claimed that the Kra-Dai languages (“Tai-Kadai” in his terminology) are a sister family to Austronesian within a larger “Austro-Thai” phylum. Based on innovations in the numeral system shared by Malayo-Polynesian and Kra-Dai, Sagart (2004) argued instead that the Kra-Dai languages are a subgroup, not a sister, of the Austronesian language family. Here I attempt an account of the Kra-Dai tone categories from the viewpoint of Austronesian consonant endings, taking these, rather than putative Proto-Austro-Thai endings, as the sources of Kra-Dai tones. Specifically, these endings are thought to have existed in an Austronesian language ancestral to Proto-Kra-Dai and Proto-Malayo-Polynesian, which I call “Proto-Southern Austronesian” (PSA). This hypothetical language was part of the Puluqish subgroup of Austronesian languages. Sagart (2008) defines Puluqish as the Austronesian subgroup cumulating all six innovative numerals \*lima ‘5’, \*enem ‘6’, \*pitu ‘7’, \*walu ‘8’, \*Siwa ‘9’ and \*puluq ‘10’. These numerals are those found in both Malayo-Polynesian and Kra-Dai. Consequently the same numerals must have been present in Proto-Southern Austronesian. Outside of Malayo-Polynesian and Kra-Dai, the Puluqish group includes three languages of southern Taiwan: Amis, Puyuma and Paiwan (Figure 1).

Two points should be borne in mind. First, while Austronesian-related words in Kra-Dai tend to belong to the basic vocabulary, they are quantitatively limited: no more than a few dozen reliable matches have been identified so far between Austronesian and any Kra-Dai language. Second, within Kra-Dai, the B and C categories are considerably smaller in terms of membership than the A and D categories. To give an idea of the proportions, Ostapirat’s (2005) list of 50 core Kra-Dai items includes 26 words with probable Austronesian cognates;

of these, twenty (77 %) are in tones A or D, one is in tone B (4 %), three in tone C (12 %), and two (8 %) alternate between tones B and C. For these reasons any discussion of the origins of the Kra-Dai B and C categories from an Austronesian point of view necessarily relies on very small numbers of examples. The sound correspondences identified in this paper account for three words in tone C and eleven in tone B. In addition, a new source of tone A is identified, with five examples. Yet, even with nineteen new tonal etymologies, this paper represents a significant advance in our understand of the formation of the Kra-Dai tone system. By comparison, Ostapirat's (2005: 123-125) proposed AN-Kra-Dai correspondences for tones B and C were illustrated by three examples in total.<sup>2</sup> Similarly Norquest (2013: 114) gave four examples of tone B corresponding to either of PAN medial \*-R- or final \*-R, and none of tone C.

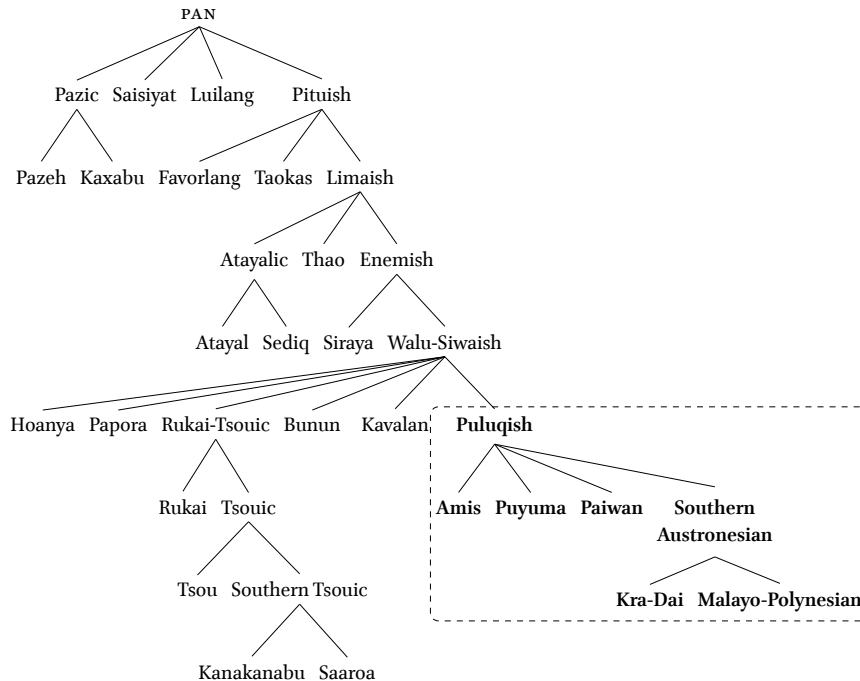


FIGURE 1: Austronesian phylogeny showing the Puluqish clade (box)

<sup>2</sup> Ostapirat (2005: 123-125) related tone B with Austronesian \*-h or medial \*-R- in Zorc's (1995) reconstruction: cited examples were 'chaff, PMP \*qepah, and 'shoulder', PAN \*qabaRa (via \*qabaR > qabah); for tone C, he correctly suggested a link to the final segment in Zorc's \*qúluH 'head'. Zorc's \*H and \*h are not contrasting phonemes, however: his \*-H words have reflexes in Taiwan while his \*-h words do not.

2 PAN vocalic and *h*-type endings2.1 Tsuchida's \*H<sub>1</sub> and \*H<sub>2</sub>

In examining the formation of Kra-Dai tone categories, reference must be made to Tsuchida's (1976) reconstruction of two PAN final consonants: \*H<sub>1</sub> and \*H<sub>2</sub>. Tsuchida's word-final \*H<sub>1</sub> is reflected as *-h* in six Formosan languages: Pazeh, Saisiyat, Atayal, Sediq, Takituduh Bunun and Amis, as well as in Aklanon, a Bisayan language of the Philippines. The Formosan language Saaroa has a glottal stop for \*H<sub>1</sub>, followed by a copy of the preceding vowel (an "echo vowel"). Tsuchida found a more restricted distribution for his \*H<sub>2</sub>: *-h* only in Takituduh Bunun and in Aklanon. Saaroa again has a glottal stop followed by an echo vowel (Table 1). He did not discover reflexes of his two phonemes in any other Austronesian languages.

TABLE 1: PAN vocalic endings and Tsuchida's \*H endings

	Vocalic endings	*-H <sub>1</sub>	*-H <sub>2</sub>
Pazeh <sup>LT</sup>	-∅	-h	-∅
Saisiyat <sup>L04</sup>	-ʔ	-h	-ʔ
C'uli' Atayal <sup>L04</sup>	-ʔ	-h	-ʔ
Truku Sediq <sup>L04</sup>	-∅	-h	-∅
Saaroa <sup>Ts</sup>	-∅	-ʔ	-ʔ
Takituduh Bunun <sup>Ts</sup>	-ʔ	-h	-h
Takituduh Bunun <sup>L88</sup>	-ʔ	-∅	-∅
Amis <sup>Na</sup>	-∅	-h	-∅
Aklanon <sup>SZN</sup>	-∅	-h	-h

Earlier discussions of Kra-Dai tone formation in relation to Austronesian did not make use of Tsuchida's reconstructions. Ostapirat's (2005) discussion of sound correspondence between Kra-Dai tones and Austronesian relied on Zorc's (1995) Austronesian reconstruction, which does not distinguish between \*H<sub>1</sub> and \*H<sub>2</sub>. As a result, Ostapirat's observations on tones B and C are limited. Had he relied on Blust's or on Wolff's instead, he would not have been more successful. Kra-Dai tonogenesis cannot be understood outside of Tsuchida's distinction, as will be shown below. In §4.2, I argue that one more consonantal ending, \*X, hitherto not reconstructed, existed in PAN, and is involved in Kra-Dai tonogenesis.

Forty-three years after Tsuchida's dissertation, only minor points need to be added to his statement of the correspondences. First, in Li's (1988) study of

Bunun dialects, PAN words with \*-H<sub>1</sub> and \*-H<sub>2</sub> have lost *-h* in Takituduh Bunun, but are still distinct from words with original vocalic endings, which end in *-ʔ*. Although Li's phonetic account of Takituduh Bunun differs from Tsuchida's, the difference is superficial: the same sets of words are distinguished. Second, other Bisayan languages, like Tausug, essentially show the same distinction as Aklanon. Third, several Philippine languages such as Itbayaten and a number of languages spoken by Negrito populations, show final *-h* in some, but not all, of Tsuchida's words with PAN \*H<sub>1</sub> and \*H<sub>2</sub>.

The apparent randomness in intra-Philippine correspondences for Tsuchida's two \*H phonemes motivated Blust's rejection of these phonemes until he changed his stance in a recent paper (Blust 2018) on the basis of revised material from Aklanon, provided to him by R. D. Zorc. Although he now silently accepts Tsuchida's two \*H phonemes, he continues to be puzzled by the irregularity of their reflexes in the Philippines, which seem to require *at least* twelve distinct correspondences: he speaks of a "major comparative puzzle" (p. 243), finds that "there is no plausible basis for treating the mystery aspirates as either innovations or retentions" (p. 244), and feels that this poses a serious problem to comparative linguistics.

In order to clarify the nature of Blust's "mystery aspirates" and to show that they do not pose a threat to \*H<sub>1</sub> and \*H<sub>2</sub>, I compare in Table 2 the distribution of *-h* and of vocalic endings in the Philippine language Aklanon and in the Formosan language Takituduh Bunun. The top panel in Table 2 tabulates all thirteen Aklanon words ending in *-h* in the 210-word basic vocabulary database of Greenhill et al. (2003)<sup>3</sup> and their cognates in Takituduh Bunun (data from Li 1988, Li 2004, Tsuchida 1976). Outside of two instances of vacillation in Li's data ('ash' and 'two', marked with an exclamation mark in Table 1), final consonants correspond as *h* :: ∅ :: ∅ :: *h*. There are no exceptions. The bottom panel lists all seven Aklanon words ending in a vowel in Greenhill et al. (2003) which have cognates in Takituduh Bunun. Endings in this set correspond as ∅ :: ʔ :: ʔ :: ʔ. There is a single exception, a personal pronoun. These correspondences are quite regular.

Why, then, are the correspondences involving final *-h* and vowel endings not regular among Philippine languages? Each of the languages in Blust's sample retains a different set of *h*-words, so that direct comparison between them returns an inordinately high number of correspondences; it is noteworthy that the number of correspondences increases each time a new language is added. It seems very probable that this is a situation where languages retaining *-h*

<sup>3</sup> The Aklanon data in Greenhill et al. (2003) are a composite of sources, not homogeneous as to phonetic notation.

have borrowed extensively from locally more influential languages, their close relatives, which have lost *-h*. Donor and borrower being closely related and in an intimate contact relationship, loans are both numerous (including basic vocabulary) and difficult to detect: this gives the impression that irregular sound change has interfered with phonological transmission. It is significant that some of the smaller Philippine languages, in particular those spoken by Negrito populations, are among the languages investigated by Blust.

TABLE 2: *-h* endings and vowel endings in Aklanon and Takituduh Bunun

	Aklanon	Takituduh Bunun		
		Li (1988)	Li (2004b)	Tsuchida (1976)
'hand'	<i>ʔalímah</i>	—	<i>ʔima</i>	—
'left'	<i>waéah</i>	<i>ta-vili</i>	<i>tavili</i>	<i>tana-vilih</i>
'mother'	<i>ʔinah</i>	<i>t-ina</i>	<i>tina</i>	<i>t-inah</i>
'father'	<i>ʔamáh</i>	<i>t-ama</i>	<i>tama</i>	<i>t-amah</i>
'to work'	<i>ʔumáh</i>	—	—	<i>qumah</i> 'farm'
'to dig'	<i>kalih</i>	<i>ma-kai</i>	<i>makai</i>	<i>makaih</i>
'louse'	<i>kú:tuh</i>	<i>kutu</i>	<i>kutu</i>	<i>kutuh</i>
'stone'	<i>batoh</i>	—	—	<i>batuh</i>
'ash'	<i>ʔabúh</i>	<i>qabu!</i>	<i>qabu?!</i>	<i>qabuh</i>
'new'	<i>bágʔuh</i>	<i>baqlu</i>	<i>baqlu</i>	<i>baqluh</i>
'one'	<i>ʔisayáh</i>	—	<i>tasʔa</i>	—
'two'	<i>daywáh</i>	<i>dusa?!</i>	<i>dusa!</i>	—
'three'	<i>tatlúh</i>	<i>tau</i>	<i>tau</i>	<i>tauh</i>
'breast'	<i>sóso</i>	<i>cucu?</i>	<i>cucu?</i>	—
'eye'	<i>matá</i>	<i>mata?</i>	<i>mata?</i>	<i>mata?</i>
'to pound, beat'	<i>bayo</i>	—	<i>mabađu?</i>	—
'you'	<i>kamo</i>	—	<i>ʔamu</i>	—
'five'	<i>lima</i>	<i>hima?</i>	<i>hima?</i>	—
'seven'	<i>pitu</i>	<i>pitu?</i>	<i>pitu?</i>	<i>pitu?</i>
'eight'	<i>waeo</i>	<i>vau?</i>	<i>vau?</i>	<i>vau?</i>

## 2.2 Blust's reinterpretation of Tsuchida's \*H<sub>2</sub>

Having finally accepted the validity of Tsuchida's two \*H phonemes, Blust (2018) proposes that the reflex of PAN \*-H<sub>2</sub> in Seediq, an Atayalic language of Taiwan, is *-x*, rather than Tsuchida's *-ʔ*. If Blust was right about this, Seediq would be the only Austronesian language having different reflexes for PAN vocalic end-



ings, \*H<sub>1</sub> and \*H<sub>2</sub>. This leads him to the claim that PAN \*-H<sub>1</sub> was \*-h and \*-H<sub>2</sub> was \*-x. He does not give consideration to Tsuchida's Seediq examples of \*-H<sub>2</sub>: *ini?* 'this', *təlu?* 'three', *k<əm>ali?* 'to dig'.<sup>4</sup>

The view of the correspondences of Proto-Atayalic fricative endings across Seediq dialects embodied in Blust's paper is not accurate. According to Li's (2004a) study, Proto-Atayalic had two fricative endings: \*h and \*x. Final \*h reflects Tsuchida's \*-H<sub>1</sub>, while \*-x has no known PAN source, being found only in words not represented outside of Atayalic (Li 2004a: 652). Some Seediq dialects maintain the distinction, usually as -h vs. -x, e.g. Paran, Toda, Hecuo, Truku Seediq (Li 2004b), all reproduced in Greenhill et al. (2003); others merge the two, e.g. Ferrell's (1969) Sakura Seediq and Pecoraro's (1977) Taroko-French dictionary. In the latter work, the outcome of the merger is noted as -x, with an orthographic free variant -ç after rounded vowels (e.g. *koyoç* alternating with *koyux* 'rain'). Examples of Taroko -x (or -ç) from PAN \*-H<sub>1</sub> include *imax* 'to drink' < \*imaH<sub>1</sub> 'id.', *sapox* 'ritual cleansing' < \*CapuH<sub>1</sub>, 'to sweep', *odox* 'to roast on embers' < \*CuNuH<sub>1</sub> 'to roast, burn',<sup>5</sup> *dalex* 'almost, near to' < \*dzaNiH<sub>1</sub> 'near', *onoç* 'female breast' < \*nunuH<sub>1</sub> 'id'.

Blust's proposal that Seediq -x reflects Tsuchida's \*-H<sub>2</sub> relies on two Seediq forms in total, of which he only gives the endings, and no source: 'head' and 'stone', presented as ending in -x (Blust 2018: 233).

'Head' very probably comes from Pecoraro's Taroko-French dictionary, which has *qolox* 'skull'. Blust infers the reconstruction PAN \*quluH<sub>2</sub>. However, as we saw, Taroko -x is the regular reflex of PAN \*-H<sub>1</sub>; moreover Saisiyat *ta-?ələh* both demands PAN \*-H<sub>1</sub> and excludes \*-H<sub>2</sub>. This word evidently reconstructs as \*quluH<sub>1</sub>.<sup>6</sup> This removes Blust's first example of Seediq -x reflecting \*-H<sub>2</sub>.

Blust's second example is Tsuchida's PAN \*batuH<sub>2</sub> 'stone'. Tsuchida (1976: 134) indeed cited the Seediq form as *bətu-nu-x*, which apparently means a stem *bətux* with pseudo-infix <nu>, where -x reflects the stem's final consonant: that is the interpretation followed by Blust. However this is not Tsuchida's interpretation: he finds this -nu- "inexplicable" (Tsuchida 1976: 189), evidently because it would imply that \*-H<sub>2</sub> evolves to Seediq -x, when all the evidence he has assembled shows that \*-H<sub>2</sub> evolved to Seediq -?. The difficulty is now solved by Li's (2004c: 1052) argument that -nux is an Atayalic male register pseudo-suffix. Li cites another example: Mayrinax Atayal *yipun* 'tooth' (< Proto-Atayalic \*gipun) in the female register, but *yip?mux* (< \*gip-nux < \*gip(un)-nux) in the male reg-

4 Unlike the variety of Seediq in Table 1, the variety of Seediq reported on by Tsuchida has a glottal stop for PAN vocalic endings and for \*-H<sub>2</sub>.

5 With \*N > d as in *adup* 'hunt', *idas* 'moon', and irregular loss of the initial consonant.

6 As Blust recognized in an e-mail message to me dated Nov. 27, 2018, and entitled "Oops". PAN \*quluH<sub>1</sub> is the reconstruction given in Sagart (2005).

ister. Consequently, the final consonant in Seediq *bətunux* is part of a suffix, not of the base. PAN *\*batuH<sub>2</sub>* regularly evolved to *\*batu* in the Atayalic branch; the male register pseudo-suffix *\*-nux* was attached to it before PAT. The regular reflex of *\*-H<sub>2</sub>* in Atayalic is zero, as all of Tsuchida's examples, including 'stone', indicate, although Blust ignores them.

This removes the second, and last, of Blust's two examples of PAN *\*-H<sub>2</sub>* being treated as *-x* in Seediq. One can safely conclude that Seediq does not support a reinterpretation of Tsuchida's correspondences for *\*H<sub>2</sub>*, and that Seediq provides no evidence for assigning the phonetic value [x] to *\*H<sub>2</sub>*.

It so happens that the present paper proposes the same phonetic value for *\*H<sub>2</sub>* as Blust; but on completely different grounds (§6). In the following sections, I examine the correspondences between Austronesian endings and Kra-Dai tones.

### 3 Tone C

#### 3.1 *Tone C from \*H<sub>1</sub>*

Austronesian words ending in *\*H<sub>1</sub>* are a source of the Kra-Dai tone C (Table 3).

TABLE 3: *Kra-Dai tone C from AN \*H<sub>1</sub>*

	PAN <sup>S</sup>	Buyang <sup>LJF</sup>	PHI <sup>O</sup>	PT <sup>P</sup>
'to come, go'	<i>*uwaH<sub>1</sub></i>	<i>va C</i>	—	—
'head'	<i>*quluH<sub>1</sub></i>	<i>qa ðu C</i>	<i>*uRəu C</i>	<i>*kraw C</i>
'shoot, outgrowth, flower'	<i>*buŋaH<sub>1</sub></i>	<i>ma ŋa C</i>	—	—

The following evidence support the reconstructions in Table 3:

- *\*uwaH<sub>1</sub>* 'to come, to go': Atayal *uah* 'come', Mantauran Rukai *oa* 'go', Babuza *m-oa* 'come', Puyuma *ua* 'go'; the Buyang tone is ambiguous for B and C but Kra comparison shows the etymological tone is C (Ostapirat 2000: 98, 125, 182); Buyang has *v* for PSA *\*-w-* and *a* for PSA *\*-a-*, as in *va B* 'nine' < PSA *\*siwa*; lack of a minor syllable reflecting PAN *\*u-* is regular with AN words beginning in a vowel;
- *\*quluH<sub>1</sub>* 'head': Paiwan *qulu*, Seediq (Taroko) *qolox* 'skull', Saisiyat *ta-ʔælæh*, Aklanon *úeoh*; the tone in Proto-Lakkja *\*kleu A* 'head' (Theraphan 1992) is unexplained; the Buyang minor syllable *qa* reflects an Austronesian first syllable beginning in a velar or uvular stop as in 'louse', Buyang *qa tu A* < PSA *\*kuCuH<sub>2</sub>*; Buyang *ðu* for PSA *\*-lu* as in 'eight', Buyang

*ma ðu*, PSA \*walu;

- \*buŋaH<sub>1</sub> ‘shoot, outgrowth, flower’: Saisiyat *poŋæh* ‘flower’ (p- unexplained; expect b-), Saaroa *vujavuja* ‘ear of foxtail’ (own fieldwork, 2014), Kanakanabu *buyabuŋa* ‘flower’, Aklanon *bú:ŋah* ‘fruit’; in Amis, Thao, Puyuma, Atayal and Seediq, a word reconstructible as \*buŋa, without final \*H<sub>1</sub>, specifically refers to the sweet potato; the sweet potato is a south American domesticate introduced to the Austronesian world in prehistoric times across the Pacific; sweet potatoes are grown from cut and rooted shoots from the tuber: a word meaning ‘shoot’ is an appropriate name for such a plant; lack of \*H<sub>1</sub> in e.g. Amis *foŋa* ‘sweet potato’ is presumably the result of secondary spread from an Austronesian language having lost PAN \*H<sub>1</sub>; in Buyang, the minor syllable *ma* reflects an AN first syllable beginning in a labial consonant, as in *ma ta A* ‘eye’ < PAN \*maCa, ‘eight’, Buyang *ma ðu A* < PSA \*walu, ‘kill’, Buyang *ma te A* < PSA \*patay; Buyang *ŋ-* for AN \*ŋ- as in ‘deaf, Buyang *ŋan B* or *ŋan C*, PMP root \*-ŋəl.

## 4 Tone B

### 4.1 From PAN \*-R

It is becoming increasingly clear that Proto-Austronesian word-final \*-R, probably a voiced uvular fricative or approximant [ʁ], is one source of the Proto-Kra-Dai tone B. This was first pointed out in Norquest (2013), where the first two comparisons in Table 4 were noted.

TABLE 4: Proto-Kra-Dai Tone B from AN \*-R

	PAN <sup>B</sup>	PMP <sup>BS</sup>	PHI <sup>N</sup>	Buyang <sup>LJF</sup>	PT <sup>P</sup>
‘natural spring’	—	*tebuR	—	—	*bo: B
‘snake’	*SulaR	*hulaR	*lja:ɸ = B	—	—
‘dry, to dry’	—	*kaRkaR	*k <sup>h</sup> u:ɸ = B	<i>qha B</i>	*χauɸ B
‘saliva’	—	*qi(d)zuR	—	<i>qa tu B</i>	—
‘sprinkle, spurt’	—	*-buR	*p <sup>h</sup> uɸ = B	—	—

Unaware of Norquest’s earlier observations, and independently from him, I noticed the match between tone B and final \*-R in ‘dry’ and ‘saliva’, while working on the Kra language Buyang in September-early October 2017. I became aware of Norquest’s proposal on an email discussion group started by

Cecil H. Brown, through Norquest's message of October 22, 2017 to which he had attached his 2013 paper. I added the comparison for 'sprinkle, spurt' later.

Supporting evidence for the PAN and PMP reconstructions in Table 4 can be found in Blust & Trussell (2010) (where 'saliva' is written as \*qizuR). Concerning 'snake', PHL \*l- for PAN \*l is not regular; one would expect PHL \*-R-, as in 'head'. Possibly there has been dissimilation between the medial and final consonants. In the Buyang word for 'saliva', the minor syllable *qa* regularly reflects the AN first syllable \*qi-, according to the rule outlined above under 'head'; *t*- reflects \*d(z)-, as in *qa tam C2* 'blade (of knife)', cognate with PMP<sup>B</sup> \*ka-ta(d)zem 'sharpness';<sup>7</sup> the Buyang tone is 11 = B2, the lower variant of tone B, indicating the initial was once voiced. Buyang *u* for \*u as in 'head' (Table 3), 'eight' (Table 8). PMP<sup>S</sup> \*kaRkaR 'to dry' is supported by Ilianen Manobo, Western Bukidnon Manobo *kagkag* 'to dry', Casiguran Dumagat *kag'kag* 'to dry'.<sup>8</sup> Buyang aspirated initials go back to word-internal clusters of two voiceless obstruents such as *Xk* < \*-Rk- in \*kaRkaR, compare Buyang *khut D* 'to dig' < P<sup>W</sup>M<sup>P</sup> \*kutkut 'to dig'. Although Buyang retains the Austronesian velar/uvular distinction at the initial of Austronesian final syllables, the velar in \*kaRkaR shows place assimilation to \*R, a uvular, resulting in Buyang *qh*.

Not all Austronesian words ending in \*-R give tone B in Kra-Dai. Ostapirat (2005: 122 and Table 7.18) cited PAN \*SimaR 'grease, oil, fat' as illustrating the evolution of Proto-Austronesian final \*-R to *-n* in tone A.<sup>9</sup> The Proto-Tai word 'to hear' \*ɲin A (Pittayaporn 2009) shows the same development (see Table 5). Its Austronesian prototype is either PMP<sup>B</sup> \*dəŋəR 'to hear' or a word ending in root \*-ɲiR 'to hear'.<sup>10</sup> In Northern Tai, Saek *ɲia A* 'to hear' confirms that this etymon did not end in plain \*n in PKD. The high-series tone in *ɲia A1* 'to hear' in Bouyei, also a Northern Tai language, points to a word-internal cluster with voiceless first consonant, resulting from penultimate vowel syncope in a form based on, e.g. \*teɲiR (cf. Amis *ma-teɲil* 'to hear').

The very limited evidential basis at our disposal suggests that the factor conditioning evolution of \*-R to tone B or to tone A was the nature, nasal or non-nasal, of the preceding consonant. Proto-Kra-Dai would have had two condi-

7 Where *-am* reflects AN \*-em, as in 'six', PSA \*enem, Buyang *nam A* (Table 8).

8 Philippine data cited from Reid (1971).

9 Ostapirat (2005: 122) also cited a PAN \*pudeR 'navel' as cognate with Proto-Tai *duuu*, Kam-Sui ?*daa*, Hlai *reu*, Laha *dau*, a tone-A set. However the evidence for a PAN \*pudeR 'navel' is obscure. Tsuchida reconstructs the Proto-Austronesian word for 'navel' as \*pujək, Wolff as \*pugek, and Blust as \*puja or \*pujek. Ostapirat's set for 'navel' appears to reflect \*puja. The probable Buyang (Lángjià) cognate ?*duə* 'navel', a tone-A word, regularly reflects PAN \*puja (Sagart, in progress).

10 Amis *ma-teɲil* 'to hear' < \*teɲiR, Proto-Siraya<sup>A</sup> \*x<m>iliɲix < \*-ɲiR 'to hear', Proto-Sangiric<sup>Sn</sup> \*diɲiR 'id.'

tioned variants, nasalized and nonnasalized, of PAN \*-R. Only the nonnasalized variant led to tone B. However, this will need to be confirmed when more examples become available.

TABLE 5: *Proto-Kra-Dai tone A from PAN \*-R*

	PAN	PSA	PKS <sup>Td</sup>	PT <sup>P</sup>
'grease, oil, fat'	*SimaR	—	*man A	*man A
'to hear'	—	*-ŋiR	—	*ŋin A

#### 4.2 *Tone B from \*-X*

The Kra-Dai tone-B category also includes words from Austronesian prototypes not ending in \*-R. Six examples are presented in Table 6.

The word 'shoulder' is typically reconstructed as PAN \*qabaRa, without a final consonant, even though Aklanon implies some kind of *-h* ending. All Kra-Dai forms show syncope of the penultimate vowel followed by reduction to \*b of the resulting \*-bR- cluster. At least in Buyang, this is the normal evolution for AN trisyllables. After reduction of the \*-bR- cluster, \*b is treated as *ʔb* as in 'fishing rod', Buyang *ʔbit D* < \*kabit 'hook'. Ostapirat (2005) ascribes tone B in the Kra-Dai cognates to medial \*-R-, assuming the evolution \*qabaRa > \*qabaŋ > *qabah*. This is accepted by Norquest (2013). For this they need a mechanism which allows pre-Kra-Dai medial \*-R- to become word-final under certain conditions. They do not explain what the conditioning is. Perhaps they have in mind an accentual condition, affecting words accented on the penult: Philippine languages tend to have penultimate stress in 'shoulder'. However, proposed comparisons matching an Austronesian word with medial \*-R- and Kra-Dai words in tone B do not tend to have penultimate stress in Philippine languages: for instance, Norquest (2013: 114) compares his PAN \*buRəç 'spurt (from mouth)' with his Proto-Hlai \*p<sup>h</sup>ufi 'id.' (where \*-fi is the precursor of tone B), but the Philippine reflexes usually show final stress. Moreover, such a mechanism is too powerful: it doubles the number of potential Kra-Dai cognates for each Austronesian word. Benedict (1975) put it to maximal use. In general the quality of comparisons of this type tends to be low.<sup>11</sup>

Blust assigned his reconstruction \*dapa 'palm of hand' to Proto-Malayo-Polynesian until c. 2015, when he raised it to Proto-Austronesian following my identification of the cited Atayal cognate (Sagart 2016). A Bunun cognate with in-

<sup>11</sup> An Austronesian cognate of Norquest's Proto-Hlai \*p<sup>h</sup>ufi that is phonologically more straightforward than his PAN \*buRəç was presented in Table 4.

TABLE 6: PAN \*X and Kra-Dai tone B

	‘shoulder’	‘palm of hand’	‘knee’	‘bran’	‘grandfather’	‘branch’
PSA <sup>S</sup>	*qabaRaX	*dapaX	*puquX	*qepaX	*apuX	*saŋaX
Buyang <sup>LJF</sup>	<i>qa ʔba B</i>	<i>pa B</i>	<i>qhu B</i>	<i>ta pha B</i>	<i>(puu B)</i>	—
Proto-Kra <sup>O</sup>	*m-ba B	—	*C-ku B	*m-pwa B	*m-pau B	—
Proto-Hlai <sup>No</sup>	*C-va:fi = B	—	—	—	—	—
Proto-Tai <sup>P</sup>	*C̣.ba: B	*hwa: B	*ɣow B	—	—	—
Proto-Kam-Sui <sup>Td</sup>	—	—	—	—	—	*ʔŋa B
Proto-Lakkja <sup>Tn</sup>	—	—	—	—	—	*tsh-Nɛ: B
Saisiyat <sup>F</sup>	<i>ʔäbaáʔ</i>	—	<i>pöʔöʔ</i>	—	—	—
Pazeh <sup>LT</sup>	<i>abaxa</i>	—	<i>pu</i>	—	<i>apu</i>	—
Thao <sup>B</sup>	<i>qalhafaxa</i>	—	<i>puqu</i>	—	<i>apu</i>	—
Favorlang <sup>F</sup>	—	—	<i>po-o</i>	—	—	—
Atayal <sup>E</sup>	—	<i>rapaʔ</i>	—	—	—	—
Bunun <sup>Lss</sup>	—	<i>d&lt;al&gt;apaʔ</i>	—	—	—	—
Amis <sup>Na</sup>	<i>ʔafala</i>	—	<i>poʔoq (&lt; M)</i> , <i>pokoh</i>	<i>əpah</i>	—	<i>caʔay (&lt; M)</i>
Aklanon <sup>SZN</sup>	<i>ʔabá:qah</i>	—	—	<i>opáh</i>	<i>apóh</i>	<i>sayáh</i>

fixed <al> is added here. A reconstruction for ‘palm of hand’ is not to be found in Pittayaporn’s dissertation but in a personal communication (e-mail communication dated February 4, 2018), Pittayaporn indicated that he would reconstruct it as “something like \*hwa: B”,<sup>12</sup> adding that “interestingly, this has \*B like what we expect from your proposal”.<sup>13</sup> The Bunun form *d<al>apaʔ* includes the paired-object infix <al>, from PAN \*<aR> or \*<ar>. Buyang *p-* from AN \*-p-, -a for AN \*a, are regular. Buyang unpredictably retains, or loses, the first syllable of Austronesian words beginning in a consonant: retention or loss of these syllables is best modeled as a lexical, non-phonological process. Here, as in several other examples cited in this paper, the first syllable is lost.

The word “knee” is reconstructed by Wolff (2010) as PAN \*puqu on the basis of northern and western Formosan forms (Saisiyat, Favorlang, Thao). The ‘knee’ semantics appear to be a specialization of a more general meaning, such as ‘joint bone’, ‘node, as of bamboo’, well reflected in Formosan, e.g. Thao *puqu* ‘bone’. Amis has a doublet: the metathesized form *poʔoq* ‘joint’, and *pokoh* ‘joint’ with reflection of medial \*-q- as -k- through dissimilation from final \*-X, also a uvular. Buyang *qhu B* has an aspirated initial, apparently reflecting a word-internal cluster \*-Cq-, where “C” is a voiceless stop. One needs to suppose a prefixed prototype such as \*ta-puquX, \*ka-pukuX or \*qa-puquX, with vowel syncope leading to \*-pq-. An anonymous reviewer cited Saek *kwaA A1* ‘leg’, cognate with PAN \*paqa ‘id.’, making the point that if the cited Kra-Dai forms were from AN \*puquX, one would expect a labiovelar initial in Saek, instead of the observed form *kʷʷ B*. It is possible that evolution of PSA \*p-q- to *kʷ-* in Saek is conditioned by the following vowel.

For ‘bran’, the MP word \*qepah (Blust) has a probable Formosan cognate, not previously identified, in Amis *ʔəpah* ‘wine’: bran can be used as an ingredient in rice wine (this, incidentally, allows one to restrict the meaning of \*qepaX to ‘bran’: chaff cannot be used in making wine). Under the phylogeny in Figure 1, the word reconstructs to Proto-Puluqish, earlier than Proto-Malayo-Polynesian. The aspirated stop in Buyang *ta pha B* reflects the \*-qp- cluster formed after penultimate vowel syncope in a trisyllabic prototype \*ta-qepaX, where \*ta-, directly reflected in the Buyang minor syllable, is perhaps the same inalienable possession prefix as in Saisiyat *ta-ʔælæh* < \*ta-quluH<sub>1</sub> ‘head’.

The item glossed as ‘grandfather’ also includes the meaning ‘grandchild’: it is one of several AN terms that designate both participants in a reciprocal kin rela-

12 An anonymous reviewer points out that Central Tai has tone C in this item. This is unexplained.

13 Another example of the correspondence PAN \*p: Proto-Tai \*hw- is PAN<sup>B</sup> \*qaCipa “kind of turtle”, more specifically, the soft-shelled turtle, in the cited Formosan languages: Proto-Tai \*hwuə A (Pittayaporn 2009).

tionship. The Aklanon form means ‘grandchild’; the Proto-Kra form is glossed as ‘grandfather’; the cited form *puu Bɪ* is from Écūn Buyang (Ostapirat 2000), closely related to Lángjià Buyang. Buyang *p-* for AN \*-p- and *u-* for AN \*-u- are regular developments (see the examples above). Lack of a minor syllable is, again, predictable, as the AN prototype \*apuX lacks an initial consonant.

Authors reconstruct the etymon ‘branch’ with a vowel ending: PAN<sup>B</sup> \*saŋa ‘bifurcation, fork of a branch’, PAN<sup>W</sup> \*caŋa ‘branch’: however the ending in Aklanon and the metathesis in Amis show that a consonantal ending was present. In the remarkable PLK form \*tsh-Nɛ: B, \*-N- is a nasal of uncertain place of articulation reconstructed on the basis of modern forms with nasalized vowels like *tshễ*. Theraphan (1992: 66) explains that Lakkja’s nasal vowels go back to earlier oral vowels preceded by a consonant+nasal cluster, and correspond to nasals plus oral vowels in Proto-Tai and Proto-Kam-Sui. For instance ‘pig’ is reflected as Proto-Lakkja<sup>Tn</sup> \*kh-Nu: A, PT<sup>P</sup> \*hmu: A, both < \*kmu A; ‘dog’, Proto-Lakkja<sup>Tn</sup> \*kh-Na: A, PT<sup>P</sup> \*hma: A, both < \*kma A. Here Proto-Lakkja \*tsh-N- is from an earlier \*tsɲ-, matching AN \*saŋ(aX) after vowel syncope. PKs<sup>Td</sup> shows the velar nasal. In order to explain the higher-series tonal reflex of tone B in ‘branch’ within a theory putting strong restrictions on onset consonant clusters, Thurgood (1989) reconstructed a preglottalized nasal \*ʔŋ. Reconstructing \*sɲ- or \*tsɲ- achieves the same result as far as tone register goes and makes better comparative sense.

In a world where Tsuchida’s \*H<sub>1</sub> and \*H<sub>2</sub> are the only possible final *h*-sounds, ‘shoulder’ and ‘grandfather’ could have ended in PAN \*-H<sub>2</sub>, ‘palm of hand’ in a vowel, ‘bran’ in \*-H<sub>1</sub> and ‘branch’ in either of \*H<sub>1</sub> and \*H<sub>2</sub>; ‘knee’, ‘thorn’, ‘chaff’ fit none of \*H<sub>1</sub>, \*H<sub>2</sub> or a vocalic ending. In such a world, the collection of words in Table 6 has no particular unity. Table 6 makes the agreement between these forms clear: Formosan languages outside of Amis show the same endings as PAN vocalic endings, while Amis (except in ‘shoulder’) and Aklanon have *-h* in word-final position. In Amis, moreover, metathesis into intervocalic position produces a glottal stop. This array is distinct from Tsuchida’s \*-H<sub>1</sub> and \*-H<sub>2</sub>: we need a third PAN ending, which I write as PAN \*-X. This was probably another back-of-the-vocal-tract voiceless fricative, given that it evolves into Aklanon and Amis *-h*. It was probably homorganic with \*q, since (a) it caused a preceding \*q to dissimilate to *-k-* in Amis (‘knee’), and (b) it has the same reflex *-ʔ-* as \*q in intervocalic position. It was also probably homorganic with PAN \*-R since (a) the two merged into the Proto-Kra-Dai tone B category, and (b) the absence of *-h* in the Amis word for ‘shoulder’ can be interpreted as a case of dissimilation between two consecutive uvular fricatives, \*-R- and \*-X, in the word’s final syllable. In other words, \*X was probably the uvular fricative [χ].

Naturally, as with \*-H<sub>1</sub> and \*-H<sub>2</sub>, some Austronesian words with \*-X do not



have a Kra-Dai cognate. Two examples are presented in Table 7. With ‘thorn’, Bunun excludes \*H<sub>1</sub> and \*H<sub>2</sub>, while Aklanon excludes a vocalic ending; with ‘chaff’, Amis excludes a vocalic ending and \*H<sub>2</sub>, while Saisiyat excludes \*H<sub>1</sub>.

TABLE 7: PAN \*-X without Kra-Dai cognates

	‘thorn’	‘chaff’
PAN <sup>S</sup>	*duRiX	*qeCaX
Saisiyat <sup>F</sup>	—	<i>kæ-ʔsæʔ</i>
Bunun <sup>L88</sup>	<i>duliʔ</i>	—
Amis <sup>Na</sup>	—	<i>ʔtah</i>
Aklanon <sup>SZN</sup>	<i>dúgih</i>	—

## 5 Tone A

Kra-Dai A-category words correspond primarily to Austronesian words ending in sonorants (vowels, semi-vowels, nasals, liquids). This was explicitly noted by Ostapirat (2005). A partial list of examples is presented in Table 8.

TABLE 8: Proto-Kra-Dai Tone A from Austronesian sonorant endings

	Buyang <sup>LJF</sup> (Kra branch)	PHL <sup>O</sup>	PT <sup>P</sup>	PAN	PMP <sup>S</sup>
‘eye’	<i>ma ta A</i>	*ata A	*p.ta: A	*maCa	*mata
‘eight’	<i>ma ɖu A</i>	*aRu A	—	—	*walu
‘seven’	<i>tu A</i>	*ʔtu A	—	—	*pitu
‘bear’ (n.)	<i>ta mɛ A</i>	*mui A	* <sup>h</sup> mwuj A	*Cumay	—
‘to kill’	<i>ma te A</i>	—	—	*paCay	*patay
‘fire’	<i>pui A</i>	*api A	*wɔj A	*[ ]apuy	*[ ]apuy
‘moon’	<i>ʔda:n A</i>	*ɲa:n A	*ɓluən A	*bulaN	*bulan
‘six’	<i>nam A</i>	*( <sup>ə</sup> )num A	—	—	*eʔenem
‘to eat’	<i>ka:n A</i>	—	*kuɲ A	*kaen	*kaen

Tsuchida’s PAN \*-H<sub>2</sub> is a second source of Kra-Dai category A (Table 9). The form \*kuCuH<sub>2</sub> ‘louse’ is a reconstruction by Tsuchida (1976), where \*H<sub>2</sub> is supported by *-h* in Takituduh Bunun, Aklanon, *-ʔu* in Saaroa, and absence of *-h* in Amis, Pazeh and Saisiyat. PAN \*k- is reflected by the minor syllable *qa* in Buyang,

while PHl<sup>O</sup> retains the first vowel of the AN word. Buyang *t-* for AN \*-C- is regular: compare *ma ta A* < \*maCa ‘eye’, Buyang *ma ta A* < \*maCa ‘kill’, Buyang *ma tɛ A* < \*paCay ‘kill’.

TABLE 9: Proto-Kra-Dai Tone A from PAN \*-H<sub>2</sub>

	Buyang <sup>LJF</sup> (Kra)	PHl <sup>O</sup>	PT <sup>P</sup>	PAN <sup>S</sup>	PMP <sup>S</sup>
‘louse’	<i>qa tu A</i>	*utu A	*traw A	*kuCuH <sub>2</sub>	*kutuh
‘ear’	<i>ta ða A</i>	*ilai A	*krwu: A	*CaŋilaH <sub>2</sub>	*taliŋah
‘three’	<i>tu A</i>	*utu C!	—	*tutuluH <sub>2</sub>	*tutuluh
‘bile’	<i>?di A</i>	*dai A	*bli: A	*qapejuH <sub>2</sub>	*qapejuh
‘hand’	—	*mi A	*mwu: A	*qa-limaH <sub>2</sub>	*qalimah

Tsuchida (1976) reconstructed ‘ear’ \*CaliŋaH<sub>2</sub> but there is strong evidence for a doublet PAN \*CaŋilaH<sub>2</sub>, with \*ŋ before \*l. Wolff (2010) convincingly argues that the latter is the original form: he claims that \*CaliŋaH<sub>2</sub> results from metathesis, the motivation for which was to align ‘ear’ with other <al>-infixes of paired body parts. A similar metathesis can be observed in the Thao word for ‘shoulder’, *qalhafa*, which reflects pseudo-infixes \*q<aR>abaH<sub>2</sub> rather than original \*qabaRaH<sub>2</sub>, and in Tausug *?agabah* ‘shoulder’. It is not clear which of the two variants is behind Buyang *ta ða A*. After penultimate vowel syncope, \*-ŋl-, or \*-lŋ-, as the case may be, simplified to *-l-*, regularly reflected by *ð-* in Buyang (cf. ‘head’, ‘eight’). The minor syllable *ta* regularly reflects AN \*Ca-. Note the curious-looking \*i at the end of Proto-Hlai \*ilai A. Proto-Hlai normally reflects PAN \*-a as \*-a. We may have here an indirect indication of the velar nature of \*H<sub>2</sub>, if a relatively fronted PSA \*a triggered the evolution \*-x > \*-c > \*-i in Hlai.

Tsuchida (1976) reconstructed ‘three’ \*teluH<sub>2</sub>, with \*H<sub>2</sub> supported by *-h* in Takituduh Bunun and Aklanon, combined with lack of *-h* in Seediq and Amis. A majority of Formosan languages have *u* in their first syllable, so I am amending \*teluH<sub>2</sub> to \*tuluH<sub>2</sub>. Tone C in Hlai is unexpected: based on PAN \*tuluH<sub>2</sub>, or (more probably) \*tu-tuluH<sub>2</sub> with CV-reduplication given the first vowel in \*utu, one expects tone A, not tone C, in the Hlai word. The irregularity may be due to list analogy since ‘three’ is part of a series of four consecutive numerals in tone C: \*ci C ‘one’, \*alau C ‘two’, \*utu C ‘three’, \*atəu C ‘four’. The Buyang form *tu A* shows penultimate vowel syncope characteristic of trisyllables: an AN \*CV-tuluH<sub>2</sub> would give an internal \*-tl- cluster, resolved as *t-* in Kra-Dai.

Concerning ‘bile’, gall, PAN \*-H<sub>2</sub> or \*-X are indicated in the AN word by lack of *-h* in Pazeh *apuzu* and Saisiyat *pæ?zo?* (*q/p* metathesis), and its presence in Aklanon *ápдох*. The word is not attested in Amis, the only language that could

discriminate between \*-H<sub>2</sub> and \*-X. Buyang *qa ?di A* ‘gall bladder’ has *?d* for PAN \*j, as in ‘navel’ *?duə A* < PSA \*puja. Ferlus (1996) reconstructs PKS \*pɔ- in this word. PT<sup>P</sup> \*ɓli: A has voiced \*pɔ- to \*bɔ- and reduced the cluster to \*ɓl-. The Buyang minor syllable moreover implies a first syllable beginning in a velar or uvular stop, implying an Austronesian prototype with a consonantal skeleton like \*{k,q}\_p\_j. Very few AN words fit this description, but PAN \*qapejuH<sub>2</sub> ‘gall’ does. The vowel correspondence in the last syllable is unexplained, however.

Outside of Zorc’s PAN \*qa-límaH, ‘hand’ is typically reconstructed with a vocalic ending, e.g. Blust \*qa-*lima*, Wolff \*lima, \*qa-*líma*. There is actually little doubt that this word ended in \*-H<sub>2</sub>: Aklanon *?alímah* ‘hand’ and Takituduh Bunun (Li) *?ima* ‘id.’ independently argue for it (see Table 1). Lack of a final glottal stop in the Takituduh word is specifically noted by Li as a problem (Li 1988: 756), because he expects that ‘hand’ should have the same ending as ‘five’, Aklanon *lima*, Takituduh Bunun *hima?*. But Aklanon and Takituduh agree in pointing to \*lima ‘five’, \*qa-*lima*H<sub>2</sub> ‘hand’. For discussion of the vowel reflexes of ‘hand’ in the Kra-Dai branches, see Ostapirat (2005: 116) and more recently Vovin (2018).

## 6 Recapitulation

Outside of Austronesian stop endings, which remain stops in Kra-Dai (“tone” D), and sonorant endings, which usually give tone A (but see §9), four Austronesian consonant endings contribute to the formation of Kra-Dai tones: \*-H<sub>1</sub>, \*-H<sub>2</sub>, \*-R, and \*-X. PAN \*-R, \*-H<sub>1</sub> and \*-H<sub>2</sub> are classical reconstructions. \*-X is proposed here for the first time. Table 10 gives their reflexes in Formosan languages and in Aklanon. Tentative phonetic values for each are added, and my phonetic interpretation of these phonemes takes the following considerations into account.

As for \*-R and \*-X, austronesianists regard PAN \*R either as a uvular fricative/trill (Dempwolff 1934), a voiced velar fricative (Wolff 2010), or an alveolar trill (Blust 2009). Considering that outside of \*R, the PAN sound system lacks a full series of voiced fricatives,<sup>14</sup> and further, that the voiced stop series of PAN lacks a uvular member, I assume that one source of PAN \*-R is a voiced uvular stop. This suggests the value [ʁ] for PAN \*-R. A detail further supports the idea that \*-R was uvular: while Buyang retains the AN distinction between AN velars and uvulars, in ‘dry’ PMP \*kaRkaR, the Austronesian velar initial is reflected as a uvular in Buyang, i.e. *qha*. This is probably due to assimilation from the first consonant in the medial \*-Rk- cluster.

<sup>14</sup> The phoneme labelled \*z in Blust’s system was probably an affricate \*dz.

TABLE 10: PAN endings involved in Kra-Dai tone formation

PAN endings	vowels	*-H <sub>2</sub> [x]	*-X [χ]	*-R [ʁ]	*-H <sub>1</sub> [h]
Kra-Dai tone	A	A	B	B	C
Pazeh	-∅	-∅	-∅	-x	-h
Saisiyat	-ʔ	-ʔ	-ʔ	-L	-h
Thao	-∅	-∅	-∅	-lh	-∅
Atayal	-ʔ	-ʔ	-ʔ	-g, -∅	-h
Seediq	-∅	-∅	—	-g, -l, -∅	-h
Takituduh Bunun (Tsuchida)	-∅	-h	-∅	-l	-h
Takituduh Bunun (Li)	-ʔ	-∅	-ʔ	-l	-∅
Tsou	-∅	-∅	—	-r	-∅
Kanakanabu	-∅	-∅	—	-r	-∅
Saaroa	-∅	-∅	—	-r	-∅
Rukai	-∅	-∅	—	-∅	-∅
Amis	-∅	-∅	-h	-l	-h
Puyuma	-∅	-∅	—	-r	-∅
Paiwan	-∅	-∅	—	-∅	-∅
Aklanon	-∅	-h	-h	-g	-h
Number of examples	many	5	6	5	3

Was final \*-R a fricative or an approximant in the period leading to PKD? After a nasal consonant, \*-R may have had a less constricted variant, evolving to a final sonorant in tone A in Kra-Dai (Table 5). The variant of \*-R which occurred after other consonants was probably fricative: its merger with \*-X presumably occurred as a result of final devoicing in \*R. Reasons for reconstructing \*X as [χ] were given in section 4.2. A voiceless uvular fricative moreover fills an empty slot within the PAN consonant system.

Concerning \*-H<sub>1</sub> and \*-H<sub>2</sub>, the merger of \*-s with \*H<sub>1</sub> in the Hlai and Tai branches, presumably via a \*-h stage (see §8.1), makes sense if \*-H<sub>1</sub>, the main source of the Kra-Dai tone C, was itself [h]. That Mon-Khmer words ending in laryngeal sounds \*-h and \*-ʔ end up in tone C once borrowed into Buyang also makes sense if \*H<sub>1</sub> was a laryngeal: PMK \*ʔah ‘meat’, Buyang uə C; PMK \*t<sub>1</sub>əh ‘bean’, Buyang ma tai C; PMK \*t<sub>2</sub>rawʔ ‘taro’, Buyang ɕuə C; PMK \*snaʔ ‘cross-bow’, PT<sup>P</sup> \*h<sup>n</sup>wuə C; PMK \*kaaŋʔ ‘chin’, Buyang qa:ŋ C.<sup>15</sup>

As to \*-H<sub>2</sub>, its phonetic value may have been [x], another gap among the

<sup>15</sup> PMK forms are from Shorto (2006).

voiceless fricatives of PAN. As mentioned above in the discussion of ‘ear’, the vowel at the end of Proto-Hlai \*ilai A may be an indirect clue of the velar nature of \*-x, since \*-x is the more fronted of our three fricatives.

## 7 From Proto-Austronesian to Kra-Dai

The PAN five-way contrast of final consonants in Table 10 appears to have persisted into Proto-Southern Austronesian. From Proto-Southern Austronesian down to Proto-Kra-Dai, the formation of Kra-Dai tone categories is a history of mergers.

PAN \*-h maintained itself into PKD. Before PKD, PAN \*-H<sub>2</sub> was lost word-finally, exposing the preceding vowels: the newly-formed vocalic endings became indistinguishable from original vocalic endings, resulting in the Kra-Dai A category. In parallel, the non-nasalized allophone of PAN \*-R devoiced, merging with \*-X. The result was the B category.

Thus, aside from oral stop endings, which constitute a separate category (“tone D”), and sibilant endings, which were still sibilants in PKD, as argued in §8, PKD words belonged to either one of three categories: 1. ending in sonorants, including vowels; 2. ending in [-χ]; 3. ending in [-h].

These three types, respectively, are the precursors of the Kra-Dai A, B and C categories. This three-way contrast is still in place in each of the Kra-Dai branches (Ostapirat 2005) though not necessarily in the same phonetic garb. Thus in the Tai branch, at the time of contact with Chinese, [-h] had shifted to [-ʔ] and [-χ] had shifted to [-h].<sup>16</sup> As a result, Chinese words ending in \*-h and in \*-ʔ were treated as tone B and C, respectively, in Tai.

To recapitulate, the present model derives the Kra-Dai tonal categories mainly from four ending conditions existing in PAN and maintained in PSA: vocalic endings and three fricatives \*-h, \*-χ, and \*-x. Evolution into the PKD A : B : C contrast was conditioned by place of articulation: the laryngeal \*-h gave the C category; the two uvulars \*-ɣ and \*-χ gave B, while the velar \*-x fell, merging with vocalic endings to form the A category. In contrast, evolution from PSA to Aklanon was conditioned by manner of articulation: the three fricatives merged into Aklanon -h, contrasting with PSA vocalic endings.

<sup>16</sup> Marc Miyake suggested a chain shift (email communication to the author, Nov. 21, 2017).

## 8 Post-PKD developments: Sibilant endings

The Proto-Austronesian sibilant endings \*-s and \*-S were still sibilants in Proto-Kra-Dai: they have have different tonal outcomes across Kra-Dai branches.

### 8.1 PAN \*-s

Haudricourt (1975: 254-255) explained Sui  $\eta e_3 < C$ , Tai  $hai_3 < C$ , Yay  $tai_3 < C$ , all ‘to weep’, as cognates of PAN \**taŋis* ‘id.’<sup>17</sup> A believer in Austro-Thai, he assumed that the first Austro-Thai vowel fell in Kra-Dai, resulting in a \**tŋ-* onset cluster, differently resolved in Sui, Tai and Yai; while final \*-s evolved to tone C via an \*-h stage. He pointed to parallel developments from a final [s] to [h], and ultimately to a tone, in Vietnamese and in Chinese. Haudricourt’s phonetic interpretation remains valid under the present understanding of Kra-Dai as a branch of Austronesian. Development to tone C has also occurred in Hlai: Proto-Hlai \**ŋai C* ‘to weep’ (Ostapirat 2004). In the Kra branch, however, \*-s has hardened to \*-t, joining the D category: Proto-Kra \**nit D* (Ostapirat 2000). Ostapirat (2005: 125) gives another example of a tone-C word corresponding to AN \*-s: Tai *huai C*, Kam-Sui \**kui C*, cognate with PMP \**qaRus* ‘current, stream’. This word is not attested in Kra or Hlai. A third example is added here: PMP<sup>B</sup> \**keRas* ‘hard’, for which Ferlus (1996) gives Common Thai \**kla C*.

TABLE 11: Evolution of final \*-s

	AN	PT <sup>P</sup>	PKS <sup>Td</sup>	PHI <sup>O</sup>	PKR <sup>O</sup>	Buyang <sup>LJF</sup>
‘to weep’	* <i>Caŋis</i>	* <i>t.haj C</i>	* <i>ŋne C</i>	* <i>ŋai C</i>	* <i>nit D</i>	<i>nit D</i>
‘current, stream’	* <i>qaRus</i>	* <i>qrwɔŋ C</i>	* <i>kru:i C</i>	—	—	—
‘hard’	* <i>keRas</i>	—	—	—	—	<i>qa la C</i>

Evolution of \*-s to tone C outside of the Kra branch agrees with the present reconstruction of tone C as ending in \*-h in Proto-Kra-Dai. Merger of \*-s with tone C would be more difficult to explain under a Gedney-inspired theory that projected the values of Tai tones at Chinese contact backwards, such as Ostapirat’s (2005: 124) or Norquest’s (2015: 168), for whom the Proto-Hlai tone C was \*-ʔ.

### 8.2 PAN \*-S

The first two words in Table 12 were identified by Ostapirat as examples of the correspondence of PAN \*-S to \*-t in PKD. Three examples are added here: ‘knee’,

<sup>17</sup> This word is now reconstructed with initial \*C.

‘(daughter’s) husband’ and ‘pubic hair’. These show tone A with loss of final \*-S, suggesting that evolution of \*-S to -t or to tone A was conditioned by the nature of the preceding vowel.

TABLE 12: *Evolution of final \*-S*

	PAN	PHL <sup>O</sup>	PHL <sup>N</sup>	PT <sup>P</sup>
‘flea’	*CumeS	*aməc	*hmə:t	*hmat
‘to winnow’	*tapeS	—	*wət	—
‘knee’	*tuduS	*aru A	*C-ru: A	—
‘husband, daughter’s husband’	*baNaS	*ali A	—	—
‘body hair, pubic hair’	*gumiS	—	—	*hmwu:j A

As to ‘knee’ PAN<sup>B</sup> \*tuduS, PHL<sup>O</sup> \*-r- reflects PAN \*-d- in ‘shrimp’, PHL<sup>O</sup> \*ura:ŋ A < PAN \*qudaŋ, ‘raw’ PHL<sup>O</sup> \*uri:p < PAN \*qudip. The vowel correspondence PHL<sup>O</sup> \*u to PAN \*u is also seen in ‘head louse’ and ‘three’.

Supporting evidence for ‘husband, daughter’s husband’ can be found in Blust & Trussell (2010): PAN<sup>B</sup> \*baNaS ‘male (of animals)’, PMP<sup>B</sup> \*banah ‘husband’. PAN \*-N- evolves to \*-l- in PHL<sup>O</sup>: ‘skin layer’ \*alic < PAN \*qaNiC ‘skin’; ‘child’ \*ali:k < PAN \*aNak. PHL<sup>O</sup> \*i for PAN \*a as in ‘hand’, PHL<sup>O</sup> \*mi A < PAN \*qalimaH<sub>2</sub> and ‘one’, PHL<sup>O</sup> \*ci A < PAN \*isa. It is significant that the word’s semantics in Kra-Dai show the same innovation from ‘male’ (of animals) to ‘husband’ as PMP.

Concerning ‘body hair, pubic hair’, PAN<sup>W</sup> \*gumis ‘down, body hair’<sup>18</sup> is based on Atayal, Bunun and Thao.<sup>19</sup> In PT, \*hm- presumably goes back to an earlier \*km-, itself reflecting \*gum- after devoicing of \*g and vowel syncope. PT \*-wu:j is a rare final, so that parallels cannot be cited.

Evidence exists that the Kra-Dai word for ‘body hair, pubic hair’ ended in a sibilant in PKD, and even much more recently in Proto-Central-Eastern Kra: Lǐ (1999: 118) observes a word-final correspondence  $t :: t :: i :: i$  between Baha, Yalang, Ecu and Langjia (Table 13) for which the most likely reconstruction is a sibilant. The correspondence for PKD final \*t is  $t :: t :: t :: t$  (Lǐ 1999: 113).

Why did \*-S not result in tone C? There are two competing explanations. In one, \*S was lost through an \*-x stage, which gives tone A; in the other, although \*S was lost through an \*-h stage, that loss occurred at a time when PKD \*h had already changed to \*ʔ.

<sup>18</sup> Wolff’s \*s is identical with \*S as understood by other Austronesianists.

<sup>19</sup> Blust reconstructs PAN \*kumiS, PMP \*gumi with an unexplained voicing alternation; however since PAN \*g goes to *k* in each of the relevant Formosan languages, Wolff’s PAN \*g- accounts more naturally for the evolution to PMP.

TABLE 13: *A final sibilant correspondence in Kra (source: Li 1999: 118)*

	Baha	Yalang	Ecun	Langjia
'body hair, down'	<i>mut n</i>	<i>mət 31</i>	<i>mui 55</i>	<i>mui n</i>
'armspan'	<i>pjat 33</i>	<i>pja:t 53</i>	<i>pja:i 55</i>	<i>pja:i n</i>
'one step'	<i>ga:t n</i>	—	—	<i>kai n</i>

## 9 Unexplained aspects of Kra-Dai tone formation

Among the points not explained by the present model are words having different tones in different Kra-Dai branches. This is the case of 'head' and 'palm' discussed above. Additional examples include the numeral 'nine', PMP \*siwa, reflected by tone C \*C-βuu:ʔ in Proto-Hlai (Norquest 2015), but by tone B \*s-ɣwa in Proto-Kra (Ostapirat 2004).<sup>20</sup> This alternation is not the result of list analogy. The word for 'floodland, wet field', PAN \*bena() (Favorlang *bonna*, PMP \*bena), is in tone B in Proto-Hlai<sup>0</sup> (\*na B), but in tone A elsewhere in Kra-Dai: Proto-Kra<sup>0</sup> \*na A 'rice-field', Proto-Tai<sup>P</sup> \*na: A. The word 'excrement' is generally in tone C: Proto-Kra<sup>0</sup> \*kai C, Proto-Tai<sup>P</sup> \*C.quuj C, Proto-Hlai<sup>0</sup> \*aka:i C, Proto-Hlai<sup>N</sup> \*ɦa:y C, Proto-Lakkja<sup>Tn</sup> \*gwei C; but it is tone B in Proto-Kam-Sui<sup>Td</sup>: \*ke B. This set may reflect PAN \*Caqi() but there are phonological uncertainties, both on the Austronesian and Kra-Dai sides of the comparison. Wolff (2010) plausibly attributes these irregularities to euphemism.

Also in need of clarification is the occurrence of tones other than A in a small number of words originally ending in nasals, liquids and semivowels. In practice this mostly concerns tone-C words.<sup>21</sup> See the examples in Table 14, adapted from Ostapirat (2005: 124, Table 7.26). In other instances we have tone C and tone A alternating across branches (Table 15).

TABLE 14: *Kra-Dai Tone C in words with sonorant endings*

	PT <sup>P</sup>	PKS <sup>Td</sup>	PH <sup>0</sup>	PLk <sup>Tn</sup>	PSA
'water'	*C.nam C	*n̩am C	*nam C	*num C	*nanum
'sour'	*sɔm C	*khjum C	—	*khloṃ C	*qaRsem

<sup>20</sup> The Tai and Kam-Sui branches have a Chinese loanword for 'nine'.

<sup>21</sup> Instances of pan-Kra-Dai tone-B words having these endings and an Austronesian etymology are rare, if they exist at all. A rare example in the Tai branch is PMP \*belaj 'spotted, dappled' (Blust & Trussell 2010) vs. Proto-Tai (Pittayaporn 2009) \*bla:ŋ B 'spotted'.



TABLE 15: *Kra-Dai Tone C alternating with tone A in words with nasal endings*

	PT <sup>P</sup>	PK <sup>O</sup>	PH <sup>I</sup> <sup>O</sup>	PLk <sup>Tn</sup>	PSA
'black'	*C̥.dam A	*hl/dəm A	*(ʔ)dam C	*ʔlam A	*qudem
'to plant, transplant'	*t.nam A	*təm C	—	—	*CaNem

Contrary to a widely held belief, the Kra-Dai word for 'water' is not directly cognate with PAN \*daNum 'id.': despite the symbol used for it in the standard orthography, PAN \*-N- was a lateral. In word-medial position it generally corresponds to a lateral in Kra-Dai. The Kra-Dai form in Table 14 is a match for a PSA \*nanum, cognate with Amis *nanom* 'water', Puyuma *nanum* 'id.' (male ritual term), Kavalan *nanum*, *m-nanum* 'to drink', Amis *mi-nanom* 'id.'

A theoretical explanation for tone-C words with sonorant endings would be that PSA possessed a contrast between plain sonorant endings, producing tone A in PKD, and cluster endings composed of a sonorant and a consonant capable of producing tone C in Kra-Dai. There is some evidence for a formative \*-s of unknown function producing doublets in the AN languages by substituting itself to a word's consonantal ending, if one is present. Here are some examples:

- (1) 'salt': PAN<sup>B</sup> \*timu, PAN<sup>B</sup> \*timus < \*timu-s
- (2) 'edge': PMP<sup>D</sup> \*tep[i]; 'to walk along the edge': PMP<sup>D</sup> \*tepis < \*tepi-s
- (3) 'cut off': PAN<sup>W</sup> \*putun/\*putul, \*putuc<sup>22</sup> < \*putun-s or \*putul-s
- (4) 'to smear': PMP<sup>B</sup> \*lumeŋ, \*lumes < \*lumeŋ-s
- (5) 'slippery': PMP<sup>B</sup> \*lamaR, \*lamas < \*lamaR-s
- (6) 'to be outside': PMP<sup>D</sup> \*luwar; 'go out' luwas < \*luwar-s
- (7) 'to taste, try': PMP<sup>B</sup> \*tamiq, \*tamis < \*tamiq-s
- (8) 'to scratch': PMP<sup>D</sup>: \*garut, \*garus < \*garut-s
- (9) 'to split': PMP<sup>B</sup> \*baqak; 'to split lengthwise': PMP \*baqas < \*baqak-s.

If forms like \*lumeŋ-s, \*lamaR-s and \*luwar-s existed in PSA, they could evolve to \*lumes, \*lamas and \*luwas in PMP, and, *via* \*lumeŋh, \*lamaRh and \*luwarh, to tone C in PKD. This is parallel to the process giving the Departing tone *qùshēng* in Chinese, where Old Chinese sonorant endings plus suffixed \*-s evolve to Middle Chinese sonorant endings in tone C. We are not able to point to any actual examples of sonorant vs. \*-s alternation in Austronesian coinciding with evolution to tone C in Kra-Dai, however.

<sup>22</sup> Wolff writes PAN \*s as "c".

### Conclusion

While the number of Kra-Dai tonal etymologies in this paper is higher than in previous attempts, it must be reiterated that the hypotheses proposed here are (of necessity) based on small numbers of examples: this confers to them a somewhat tentative character.

It should also be made clear that the present account is not dependent upon the hypothesis that Kra-Dai is a branch of Austronesian (Sagart 2004). Rather, it assumes genetic relatedness between the two taxa and relies on sound correspondences between the tones of one and the endings of the other. Should the Kra-Dai languages turn out to be the non-Austronesian branch of putative Austro-Thai, the sound correspondences would remain: instead of belonging to post-PAN levels, the relevant phonemes would be assignable to Proto-Austro-Thai.

At the same time, while the present model “works” under both views of the Austronesian-Kra-Dai relationship, it is significant that the additional time depth provided by the Austro-Thai hypothesis is of little help in accounting for Kra-Dai tones: the building blocks are in place during the Formosan phase of Austronesian phonological history.

### Abbreviations and notations

AN Austronesian, KD Kra-Dai, OC Old Chinese, PAN Proto-Austronesian, PAT Proto-Atayalic, PHL Proto-Hlai, PKD Proto-Kra-Dai, PKr Proto-Kra, PKS Proto-Kam-Sui, PLk Proto-Lakkja, PMP Proto-Malayo-Polynesian, PSA Proto-southern Austronesian, PT Proto-Tai, PWMP Proto-Western Malayo-Polynesian.

Superscript capital letters following these abbreviations or modern language names identify authors: <sup>A</sup> Adelaar (2011), <sup>B</sup> Blust & Trussell (2010), <sup>D</sup> Dempwolff (1938), <sup>E</sup> Egerod (1980), <sup>F</sup> Ferrell (1969), <sup>LJF</sup> Li (1999), <sup>LT</sup> Li & Tsuchida (2001), <sup>L88</sup> Li (1988), <sup>L04</sup> Li (2004b), <sup>Na</sup> Namoh (2013), <sup>No</sup> Norquest (2013), <sup>O</sup> Ostapirat (2000; 2004; 2005), <sup>P</sup> Pittayaporn (2009), <sup>S</sup> Sagart (this paper), <sup>Sn</sup> Sneddon (1984), <sup>SZN</sup> Salas Reyes et al. (1969), <sup>Td</sup> Thurgood (1989), <sup>Tn</sup> Theraphan (1992), <sup>Ts</sup> Tsuchida (1976).

Parentheses around a segment in reconstructed forms indicate that it may have been present or not, and square brackets that its identity is unsure. Empty parentheses and brackets are used when the identity of the segment is unknown.

## Acknowledgments

This paper is a revised version of Sagart (2017a,b,c), a series of three short blog posts on the origins of Kra-Dai tones. I thank Bill Baxter, Marc Miyake, Peter Norquest, Weera Ostapirat, Pittayawat Pittayaporn and two anonymous reviewers for their valuable comments. Errors, needless to say, are mine.

## References

- Adelaar, Alexander. 2011. *Siraya: Retrieving the phonology, grammar and lexicon of a dormant Formosan language*. Berlin: De Gruyter Mouton.
- Benedict, Paul K. 1942. Thai, Kadai and Indonesian: A new alignment in Southeastern Asia. *American Anthropologist* 44(4). 576–601.
- Benedict, Paul K. 1975. *Austro-Thai language and culture, with a glossary of roots*. With a foreword by Goodenough, Ward. New Haven: HRAF Press.
- Blust, Robert. 2009. *The Austronesian languages*. Canberra: Research School of Pacific and Asian Studies.
- Blust, Robert. 2018. The “mystery aspirates” in Philippine languages. *Oceanic Linguistics* 57(1). 221–247.
- Blust, Robert & Trussell, Stephen. 2010. Austronesian comparative dictionary: Web edition. <http://www.trussell2.com/acd/> (accessed 12 May 2018).
- Dempwolff, Otto. 1934. *Vergleichende Lautlehre des Austronesischen Wortschatzes*. Vol. 1: *Induktiver Aufbau*. Berlin: Dietrich Reimer.
- Dempwolff, Otto. 1938. *Vergleichende Lautlehre des Austronesischen Wortschatzes*. Vol. 3: *Austronesisches Wörterverzeichnis*. Berlin: Dietrich Reimer.
- Egerod, Søren. 1980. *Atayal-English dictionary*. London: Curzon Press.
- Ferlus, Michel. 1996. Remarques sur le consonantisme du proto kam-sui. *Cahiers de Linguistique Asie Orientale* 25(2). 235–278.
- Ferrell, Raleigh. 1969. *Taiwan aboriginal groups: Problems in cultural and linguistic classification*. Taipei: Institute of Ethnology, Academia Sinica.
- Gedney, William J. 1986. Speculations on early Tai tones. In McCoy, John & Light, Timothy (eds.), *Contributions to Sino-Tibetan studies*, 144–156. Leiden: E. J. Brill.
- Greenhill, Simon & Blust, Robert & Gray, Russell. 2003. Austronesian basic vocabulary database. <https://abvd.shh.mpg.de/austronesian/> (accessed 8 April 2019).
- Haudricourt, André-Georges. 1954a. Comment reconstruire le chinois archaïque. *Word* 10(2–3). 351–364.
- Haudricourt, André-Georges. 1954b. De l’origine des tons en vietnamien. *Journal Asiatique* 242. 69–82.
- Haudricourt, André-Georges. 1975. À propos du puzzle de W.J. Gedney. In Harris, Jimmy G. & Chamberlain, James R. (eds.), *Studies in Tai linguistics in honour of W.J. Gedney*, 257–258. Bangkok: Central Institute of English Language.

- Li, Jǐnfāng (李锦芳). 1999. *Bùyāngyǔ yánjiū* (布央语研究). Beijing: Zhōngyāng Mínzú Dàxué Chūbǎnshè.
- Li, Paul Jen-kuei. 1988. A comparative study of Bunun dialects. In *Selected papers on Formosan languages*, 743–766. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei. 2004a. Reconstruction of Proto-Atayalic phonology. In *Selected Papers on Formosan Languages*, vol. 1, 625–692. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei. 2004b. Basic vocabulary for Formosan languages and dialects. In *Selected papers on Formosan languages*, vol. 2, 1483–1530. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei. 2004c. Men's and women's speech in Mayrinax. In *Selected papers on Formosan languages*, vol. 2, 1045–1053. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei & Tsuchida, Shigeru. 2001. *Pazih dictionary*. Taipei: Institute of Linguistics, Academia Sinica.
- Mei, Tsu-lin. 1970. Tones and prosody in Middle Chinese and the origin of the Rising tone. *Harvard Journal of Asian Studies* 30. 86–110.
- Namoh, Rata. 2013. *O pidafo'an to sowal misanopangcah*. Taipei: Nántiān Shūjú.
- Norquest, Peter. 2013. A revised inventory of Proto Austronesian consonants: Kra-Dai and Austroasiatic evidence. *Mon-Khmer Studies* 42. 102–126.
- Norquest, Peter. 2015. *A phonological reconstruction of Proto-Hlai*. Leiden: Brill.
- Ostapirat, Weera. 2000. Proto-Kra. *Linguistics of the Tibeto-Burman Area* 23(1). 1–251.
- Ostapirat, Weera. 2004. Proto-Hlai sound system and lexicons. In Lin, Ying-ching & Hsu, Fang-min & Lee, Chun-chih & Sun, Jackson T.S. & Yang, Hsiu-fang & Ho, Dah-an (eds.), *Studies in Sino-Tibetan languages: Papers in honor of Professor Hwang-cherng Gong on his seventieth birthday*, 121–175. Taipei: Institute of Linguistics, Academia Sinica.
- Ostapirat, Weera. 2005. Kra-Dai and Austronesian: Putting together archaeology, linguistics and genetics. In Sagart, Laurent & Blench, Roger & Sanchez-Mazas, Alicia (eds.), *The peopling of East Asia*, 109–133. London: RoutledgeCurzon.
- Pecoraro, Ferdinando. 1977. *Essai de dictionnaire taroko-français*. Paris: SECMI.
- Pittayaporn, Pittayawat. 2009. *The phonology of Proto-Tai*. Ithaca: Cornell University. (Doctoral dissertation).
- Pulleyblank, Edwin. 1962. The consonantal system of Old Chinese, part 2. *Asia Major* 9. 206–265.
- Reid, Lawrence A. 1971. *Philippine minor languages: Word lists and phonologies*. Honolulu: University of Hawaii Press.
- Sagart, Laurent. 1999. The origin of Chinese tones. In Kaji, Shigeki (ed.), *Cross-linguistic studies of tonal phenomena, tonogenesis, typology, and related topics*, 91–103. Tokyo: Research Institute for Languages and Cultures of Asia and Africa.
- Sagart, Laurent. 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43(2). 411–444.
- Sagart, Laurent. 2005. Sino-Tibetan-Austronesian: An updated and improved argument. In Sagart, Laurent & Blench, Roger & Sanchez-Mazas, Alicia (eds.), *The peopling of*

- East Asia: Putting together Archaeology, Linguistics and Genetics*, 161–176. London: RoutledgeCurzon.
- Sagart, Laurent. 2008. The expansion of setaria farmers in East Asia: A linguistic and archaeological model. In Sanchez-Mazas, Alicia & Blench, Roger & Ross, Malcolm & Peiros, Ilya & Lin, Marie (eds.), *Past human migrations in East Asia: Matching archaeology, linguistics and genetics*, 133–157. Abingdon: Routledge.
- Sagart, Laurent. 2016. The wider connections of Austronesian: A response to Blust (2009). *Diachronica* 33(2), 255–281.
- Sagart, Laurent. 2017a. The origins of Kra-Dai tones, current status (I): Tone C. 20 November 2017. <https://stan.hypotheses.org/123>.
- Sagart, Laurent. 2017b. The origins of Kra-Dai tones, current status (II): Tone B. 20 November 2017. <https://stan.hypotheses.org/120>.
- Sagart, Laurent. 2017c. The origins of Kra-Dai tones, current status (III): Tone A. 20 November 2017. <https://stan.hypotheses.org/117>.
- Salas Reyes, Vicente & Zorc, R. David & Prado, Nicolas L. 1969. *A study of the Aklanon dialect, volume 2: Dictionary*. Washington, DC: Peace Corps.
- Shorto, Harry. 2006. *A Mon-Khmer comparative dictionary*. Sidwell, Paul & Cooper, Doug & Bauer, Christian (eds.). Canberra: Pacific Linguistics.
- Sneddon, James N. 1984. *Proto-Sangiric and the Sangiric languages*. Canberra: Pacific Linguistics.
- Theraphan, L. Thongkum. 1992. A preliminary reconstruction of Proto-Lakkja (Cha Shan Yao). *Mon-Khmer Studies* 20, 57–89.
- Thurgood, Graham. 1989. Notes on the reconstruction of Proto-Kam-Sui. In Edmondson, Jerold & Solnit, David (eds.), *Comparative Kadai: Linguistic studies beyond Tai*, 179–218. Arlington: University of Texas at Arlington.
- Tsuchida, Shigeru. 1976. *Reconstruction of Proto-Tsouic phonology*. Tokyo: ILCAA.
- Vovin, Alexander. 2018. Proto-K(r)adai 'hand', 'eye', and 'bird' revisited. *Cahiers de Linguistique Asie Orientale* 47(1), 150–164.
- Wolff, John U. 2010. *Proto-Austronesian phonology with glossary*. With a foreword by Sagart, Laurent. 2 vols. Ithaca: Cornell University Press.
- Zorc, R. David. 1995. A glossary of Austronesian reconstructions: An introduction to Austronesian studies. In Tryon, Darrell T. (ed.), *Comparative Austronesian dictionary*, part 1, fascicle 2, 1106–1197. Berlin: Mouton de Gruyter.