



# Strata of standardization: the Phong Nha dialect of Vietnamese (Qung Bình Province) in historical perspective

Alexis Michaud, Michel Ferlus, Minh-Châu Nguyn

## ► To cite this version:

Alexis Michaud, Michel Ferlus, Minh-Châu Nguyn. Strata of standardization: the Phong Nha dialect of Vietnamese (Qung Bình Province) in historical perspective. *Linguistics of the Tibeto-Burman Area*, Dept. of Linguistics, University of California, 2015, 38 (1), pp.124-162. <<https://benjamins.com/#catalog/journals/ltba/main>>. <10.1075/ltba.38.1.04mic>. <halshs-01141389>

**HAL Id: halshs-01141389**

<https://halshs.archives-ouvertes.fr/halshs-01141389>

Submitted on 12 Apr 2015

**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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

**STRATA OF STANDARDIZATION:  
THE PHONG NHA DIALECT OF VIETNAMESE  
(QUẢNG BÌNH PROVINCE) IN HISTORICAL PERSPECTIVE\***

**Alexis Michaud\*, Michel Ferlus\*\*, Minh-Châu Nguyễn\*\*\***

\*International Research Institute MICA (CNRS – Hanoi University of Science and Technology – Grenoble INP), Vietnam

\*\*CNRS, France (*retired*)

\*\*\*Department of Linguistics, Vietnam National University, Hanoi, Vietnam

The present research, based on first-hand data, is intended as a contribution to the study of the present-day diversity of lesser-described Vietnamese dialects, and of the range of evolutionary paths to which they testify. The Vietnamese dialect of the hamlet of Phong Nha (commune of Sơn Trạch, county of Bố Trạch, Quảng Bình) is one of the “heterodox” dialects of Vietnamese, which are known to present considerable interest for the historical study of Vietnamese and of the Vietic group at large. These dialects are the product of the southerly expansion of Vietnamese over related (Vietic) languages, a process which involved various interferences. Comparative evidence reveals strata of standardization: some words are phonologically identical to Standard Vietnamese; others are of Southern Vietic stock, as demonstrated by the absence of telltale historical changes that took place in Vietnamese, such as the spirantization of medial stops; still others appear to be the result of hybridization.

Keywords: historical phonology; dialectology; spirantization; diphthongization; vowel systems; Vietnamese; Vietic languages; Quảng Bình.

## **1. INTRODUCTION**

The present study is based on first-hand data on the Vietnamese dialect of the hamlet of Phong Nha, in the commune of Sơn Trạch, county of Bố Trạch, Quảng Bình (thôn Phong Nha, xã Sơn Trạch, huyện Bố Trạch, tỉnh Quảng Bình). The aim of the present research is to contribute to the study of the present-day diversity of lesser-described Vietnamese dialects, and of the range of evolutionary paths to which they testify.

---

\* Many thanks to the language consultants, Mr. Trần Văn Hợp and Mr. Hoàng Minh Chiêm, and to Pr. Trần Trí Dõi (Department of Linguistics of Vietnam National University, Hanoi) for organizing the invitation of the consultants and for travelling to Phong Nha to accompany them on the journey to Hanoi. Many thanks to two anonymous reviewers, and to the editorial team, for the wealth of useful comments and suggestions that they offered.

Support from Agence Nationale de la Recherche (LabEx EFL, ANR-10-LABX-0083—Investissements d’Avenir) is gratefully acknowledged.

### 1.1. *The importance of heterodox dialects*

Vietnamese dialects exhibit considerable diversity, and remain relatively understudied. Emeneau (1951) provided a detailed description of the dialects of Vinh (Nghệ An) and Nam Định (the main city of the province of the same name); dialectal information is also found in Hoàng Thị Châu (1989; 2004) and in Chapter 4 of Thompson's *Vietnamese Grammar* (Thompson 1984). Apart from these and a few other notable exceptions, most research tends to focus on the varieties spoken in the major cities, with special emphasis on the two largest cities: Hanoi and Hồ Chí Minh City (Saigon). However, Maspero pointed out that, within the field of Vietnamese dialectology,

“from a historical point of view, comparison of Tonkinese [i.e. Red River delta Vietnamese: the Northern dialect of Vietnamese, including Hanoi Vietnamese] with Cochinchinese [i.e. Mekong delta Vietnamese, Southern dialect of Vietnamese, including Saigon Vietnamese], which belong in the same group, is least interesting, whereas comparison of these dialects with that of Haut Annam [spoken from the North of Nghệ An to the South of Thừa Thiên] is of much greater importance; unfortunately, this comparison has not yet been thoroughly carried out” (Maspero 1912: 3).<sup>1</sup>

The “Haut Annam” dialects show various irregular correspondences with Vietnamese as reflected in Rhodes's Dictionary (1651), referred to below as Middle Vietnamese. In view of these irregularities, they are referred to as “heterodox” (Ferlus 1995). This term will be used here, consequently using the term “orthodox dialects” for the varieties that exhibit regular relationships of correspondence with Middle Vietnamese. More than a century after the publication of Maspero's study, heterodox dialects survive in a few areas of Quảng Bình, Hà Tĩnh and southern Nghệ An; as emphasized by Alves (2007), there remains much progress to be done in the study of this crucial dialect area.

### 1.2. *Earlier work on heterodox dialects*

The first published description of a “Haut Annam” dialect is a report by Cadière (1902), who defines “Haut Annam” dialect as “the dialect spoken from Đà Nẵng to Vinh or thereabouts” (“le dialecte parlé depuis Tourane jusqu'à Vinh environ”; page X). As pointed out by Maspero (1912: 1, note 1), the data in this initial report suffer from some inconsistencies, due in part to the absence of sufficiently clear indications on where each cited form comes from. Maspero collected data

---

<sup>1</sup> “Au point de vue historique, la comparaison du tonkinois et du cochinchinois, qui appartiennent au même groupe, est la moins intéressante ; celle de ces dialectes avec celui du Haut-Annam est beaucoup plus importante : malheureusement elle n'a pas encore été faite de façon complète.”

on more than ten of the dialects of this area; some examples are adduced in his 1912 study, but to our knowledge, his field notes remain unpublished. Data on the dialect of Cao Lao Hạ (Quảng Bình) were collected by Michel Ferlus; they confirm the presence of some irregular correspondences with Middle Vietnamese, interpreted as resulting from the spread of Vietnamese on related language varieties – belonging to the Vietic<sup>2</sup> subbranch of Austroasiatic – with which there was sufficient closeness for cognate words in Vietnamese and in local Vietic languages to be perceived as such by the speakers (Ferlus 1995). Ferlus’s argument is that a phonetic compromise was reached between the realizations in Vietnamese and in local varieties, which were not replaced holus-bolus by Vietnamese. Some words were borrowed with their Vietnamese pronunciation; others were imitated in part, modifying the earlier (non-Vietnamese) form through the introduction of some of the phonological contents of the Vietnamese form.

Further additions to documentation and research on heterodox dialects include a study of the Thanh Chương variety, Nghệ An (Alves & Nguyễn Duy Hương 2007) and an experimental study of tones of several locations in Nghệ An and Hà Tĩnh (Honda 2008).

Understanding these dialects is a topic of interest (i) for Vietnamese dialectology, (ii) for the study of the Vietic subbranch as a whole, and (iii) for general linguistic models. Concerning point (ii), a better knowledge of local Vietnamese dialects is necessary for further progress in the analysis of layers of borrowings into Vietic languages that have been heavily influenced by language contact, such as Arem (Ferlus 2014) and Thổ (Ferlus 2001).

## **2. METHOD**

For administrative reasons related to research authorizations, elicitation sessions were conducted in Hanoi and not in Quảng Bình. Pr. Trần Trí Dõi, of the Department of Linguistics of Vietnam National University in Hanoi, kindly went to Phong Nha in person to invite two consultants: Mr. Trần Văn Hợp (hereafter M1), born in 1957, and Mr. Hoàng Minh Chiêm (hereafter M2), born in 1962. Both have lived continuously in Phong Nha.

The elicitation sessions took place at the recording studio of the International Research Institute MICA. The consultants were given explanations about the purpose of the work: to record the speech of their native village, as distinct from

---

<sup>2</sup> The Vietic group consists of (i) Maleng, (ii) Arem, (iii) Chứt, (iv) Aheu, (v) Hung, (vi) Thổ, (vii) Mường and (viii) Vietnamese. This group was referred to as “Việt-Mường” by Thomas and Headley (1970), followed by Ferlus (1996; 1999; 2004). Nguyễn Tài Căn (1995) aptly proposed the term “Việt-Chứt”, a compound of the names of the two most distant languages within the group; but the term “Vietic”, proposed by Hayes (1982; 1992), has now become commonly accepted for the entire group, and we follow this usage. Hayes’s proposal is to use “Việt-Mường” as a label for the lower-level grouping containing Vietnamese and Mường.

other varieties of Vietnamese of which they have some command, such as Hanoi Vietnamese – the current national standard – and dialects of Central Northern Vietnamese spoken elsewhere in Quảng Bình. Basic words were selected from Michel Ferlus’s word list (an expanded version of the EFEO-CNRS-SOAS word list for linguistic fieldwork in Southeast Asia, available online: Pain et al. 2014). The words were said by the third author (a speaker of Hanoi Vietnamese), and translated orally by the consultants.

An inherent difficulty under such a setup is that consultants have certain representations about their native dialect and its status relative to standards of correctness. It has long been known that, when an investigator from the city elicits data from rural people, they tend to modify their way of speaking, avoiding pronunciations which they think may sound ridiculous to the investigator. As a result of situations of non-egalitarian bilingualism, it can be difficult to elicit the full set of the oppositions present in the phonological system of the target dialect (see e.g. Haudricourt 1973: 23; on the importance of this issue when studying heterodox Vietnamese dialects: Thompson 1984: 79). A related difficulty is that the consultants have representations about the investigators’ expectations, methods and abilities, which also influence their behaviour. In the case of the present study, it is only on the second day that the speakers reported the existence of a local pronunciation – now frowned upon locally as coarse – which they had not indicated to us because they thought we would be unable to write it down: a voiced dental spirant, IPA [ð], found as a variant – alongside [z], currently more frequent – in certain words written with orthographic *d*. The consultants, who are literate in Vietnamese, had observed that there is no distinct character for this sound (as they interpret orthographic *d* as [z]). In the belief that “there is no way to write it down”, they considered that the investigators should dispense with it altogether, and they avoided this pronunciation. When finally heard, this sound proved to be of the greatest interest to the investigators: from a diachronic point of view, the spirant variant [ð] is a conservative realization (as explained in section 3.1.4). As a result, it was necessary to go through the word list again to establish which lexical items allow the spirant variant.

The odds are that the present description of the phonemic system achieves a good degree of accuracy, thanks in particular to (i) the presence of two speakers rather than one and (ii) sustained exchanges with the investigators during elicitation. It should nonetheless be kept in mind that the present report is based on a short (two-day) foray. There clearly remains room for a more in-depth study involving systematic vocabulary collection and a study of continuous speech.

We plan to make the recorded materials available online through the Pangloss Collection (Michailovsky et al. 2014) in 2016, if not before.

### 3. RESULTS

Phong Nha syllables have a (C<sub>i</sub>)(G)V(C<sub>f</sub>)+T structure, where C<sub>i</sub> is an initial consonant, G a glide /w/, V a vowel nucleus, C<sub>f</sub> a final consonant, and T a tone. Brackets indicate optional constituents.

#### 3.1. Consonants

##### 3.1.1. Inventory

An inventory of consonants is shown in Table 1.

	<i>coronal</i>						
	<i>bilabial</i>	<i>labiodental</i>	<i>dental</i>	<i>alveolar</i>	<i>retroflex</i>	<i>palatal</i>	<i>velar</i> <i>glottal</i>
<i>plosive</i>			t t <sup>h</sup>		ʈ	c	k k <sup>h</sup>
<i>nasal</i>	m		n			ɲ	ŋ
<i>trill</i>				r			
<i>fricative</i>	ð	f v	s	ʃ z	ʂ		x ɣ h
<i>approximant</i>				l			
<i>implosive</i>	ɓ				ɗ		

Table 1. Inventory of consonants

Alveolo-palatal ʃ and z are placed in the ‘alveolar’ column for convenience. Initial /ɛ/ was only found in the syllable /ɛɔ/ (/ɛɔ<sup>C1</sup>/ ‘to thread’, /ɛɔ<sup>C1</sup> la<sup>B1</sup>/ ‘dishonest’). It may turn out, upon further examination, to be no more than an allophone of /s/; Phong Nha /s/ is sometimes realized close to [ɛ] in various contexts. Initial /z/ is sometimes realized as a voiced palatal fricative [j].

The dental trill /r/ is realized as a retroflex fricative [ʂ] in the speech of consultant M1; in the speech of M2, it is sometimes realized as a trill, [r], and sometimes as a tap, [ɾ].

Speaker M2 often has a stop realization for /k<sup>h</sup>/, whereas speaker M1 generally realizes it as a fricative, [x~x<sup>h</sup>]. In the documented history of Vietnamese, aspirated \*p<sup>h</sup> evolved to /f/, and \*k<sup>h</sup> to /x/, leaving /t<sup>h</sup>/ as the only aspirated stop in the system. M2’s /k<sup>h</sup>/ is a conservative characteristic. In cases such as this one, where the two speakers have different phonetic realizations, the more conservative is chosen for phonemic notations.

Speaker M1 has some affrication in the realization of retroflex /ʈ/.

Syllable-final consonants will be dealt with as part of the discussion of rhymes.

### 3.1.2. Reminders about the spirantization of medial stops in Vietnamese

The initial consonant system of Vietnamese is characterized by the presence of spirants<sup>3</sup> (weak fricatives) resulting from the lenition of medial obstruents within sesquisyllables: the second consonant in CCV(C). This was noted by Haudricourt (1965: 171) and worked out in detail some time later (Ferlus 1982; supportive evidence is provided by Shimizu 2011). The following changes took place: [p] > [β], [t] > [ð], [c] > [j] and [k] > [ɣ]. These spirant consonants, intrinsically unstable, then evolved towards more stable units in Modern Vietnamese: [β] stabilized to [v], and [ð] and [j] merged to [z] in Northern Vietnamese and to [j] in Southern Vietnamese, as shown in Table 2. This consonantal stabilization is quite recent, as Alexandre de Rhodes' dictionary (1651) still shows evidence of the presence of the spirants.

	Proto-Vietic obstruents	p-b	t-d	s	c-ɟ	k-g
<i>non-spirantized</i> (initial stops)	modern phonetic value	ɓ	ɗ	t	c	k
	Vietnamese spelling (identical to Rhodes 1651)	<i>b</i>	<i>đ</i>	<i>t</i>	<i>ch</i>	<i>c/k</i>
<i>spirantized</i> (formerly medial stops)	Middle Vietnamese	β	ð	r	j	ɣ
	spelling in Rhodes (1651)	<i>ɸ, u</i>	<i>d</i>	<i>r</i>	<i>gi</i>	<i>g/gh</i>
	modern phonetic value	v	z/j	z/r	z/j	ɣ
	modern Vietnamese spelling	<i>v</i>	<i>d</i>	<i>r</i>	<i>gi</i>	<i>g/gh</i>

Table 2. A recapitulation of the evolution of stops in Vietnamese: initial stops in CV(C) monosyllables, and medial stops in sesquisyllables (C-CV(C), later simplified to CV(C))

Numerous exceptions were noted by Nguyễn Tài Căn (2000): correspondences between monosyllables in Vietic languages and spirant initials in Vietnamese, instead of what the model predicts. According to the model, spirants originate in the lenition of medial consonants. These counterexamples led Nguyễn Tài Căn to

<sup>3</sup> Martinet distinguishes fricatives, which involve distinctly audible friction, from spirants, which are identified through the “resonances of voice” (“la qualité des résonances de la voix”), i.e. a formant pattern. Diachronic studies provide compelling evidence that, among the sounds that fall into the IPA category of fricatives, some pattern with obstruents and others with continuants. Martinet recommends the use of Greek letters for spirants, e.g. [ð] and [θ] for interdental spirants, vs. [ð] and [β] for the corresponding fricatives (Martinet 2005: 100–101). Synchronically, however, it may turn out that the range of observed situations constitutes a continuum, and that a hard-and-fast encoding by different IPA symbols for fricatives vs. spirants would raise difficulties for transcribers. This would be parallel to the notion of fortis/lenis (Kohler 1979; Kohler 1984), likewise not encoded in IPA. In this article, we adhere to IPA symbols; the intended meaning can be clarified by phrases such as “the spirant [ɣ]” or “the spirant [ð]”.



the conclusion that the model must be wrong: that it is mistaken to reconstruct a pre-syllable for all the words that developed a spirant initial in Vietnamese.

A key point, however, is that the model predicts the existence of numerous exceptions. Presyllables can be lost without compensation – unlike the property of voicing in an initial consonant, for instance. In consonant shifts involving voicing properties of initial consonants (Haudricourt 1965; Ferlus 1979), great regularity is expected in the output; on the other hand, monosyllabicization can proceed in a more haphazard way. It is quite possible for presyllables to disappear from a set of words whereas other words (including homophonous words of different morphological makeup, e.g. those in which the presyllable is not morpho-semantically distinct from the main syllable) are unaffected. This is shown by the gradient of sesquisyllabicity among Vietic languages: no sesquisyllables in Vietnamese and Mường; 10% of sesquisyllables in Pong; about 35% in Thavung, Maleng and Sách/Rục; and over 50% in Arem (Ferlus 2014). Real counterexamples to the proposed generalization about spirantization in Vietnamese would come from words that have a presyllable (i.e. sesquisyllabic structure) in some Vietic languages, and a non-spirantized initial (orthographic *b-*, *đ*, *ch-*, *k-/c-*) in Vietnamese. In the absence of such counterexamples, the theory of spirantization of medial consonants cannot be considered disproved.

### 3.1.3. *The preservation of nonspirantized initials in Phong Nha Vietnamese*

Phong Nha Vietnamese preserves non-spirantized initials corresponding to spirantized initials in Middle Vietnamese, as shown in Tables 3 and 5-8. This means that, in these syllables, presyllables were lost without compensation. This is one of several possible evolutionary paths for presyllables, in the course of monosyllabicization: (i) loss without compensation; (ii) modification of medial consonants, as in orthodox Vietnamese; (iii) development of consonantal clusters, themselves yielding geminated consonants, then a tense/lax opposition among initials, which can later evolve into a split of the vowel system or of the tone system (Haudricourt 1991; Ferlus 1997a; for a synthesis: Michaud 2012: 119–120).

Henceforth, letters in superscript (B1, B2) indicate the etymological lexical tone categories; for information about tones, see section 3.3.

Proto-Vietic	gloss	Middle Vietnamese		Phong Nha Vietnamese
		spelling	IPA interpretation	
t.ke:	thorn	<i>gai</i>	ɣaj <sup>A1</sup>	kǎj <sup>A1</sup>
C.gi:ʔ	dibble	<i>gây</i>	ɣǎj <sup>B2</sup>	kǎj <sup>B2 / C2</sup>
t.ke:ʔ	girl	(con) <i>gái</i>	ɣaj <sup>B1</sup>	kǎj <sup>B1</sup>
t.ku:lʔ	knee	<i>gối</i>	ɣoj <sup>B1</sup>	kuj <sup>B1</sup>
	to carry on the back in a basket	<i>gùi</i>	ɣuj <sup>A2</sup>	kuj <sup>A2</sup>
kamʔ / gamʔ	to bite	<i>gặm</i>	ɣǎm <sup>B2</sup>	kǎm <sup>B1</sup>
C.ka:s	to scratch	<i>gãi</i>	ɣaj <sup>C1</sup>	k <sup>h</sup> aj <sup>C1</sup>
t.kɔ:c	to whittle	<i>gọt</i>	ɣot <sup>D1</sup>	k <sup>h</sup> ot <sup>D1</sup>

Table 3. Correspondences between Proto-Vietic \*k/g, Middle Vietnamese /ɣ/, and Phong Nha Vietnamese /k/

Aspiration in the last two examples, Phong Nha /k<sup>h</sup>aj<sup>C1</sup>/ ‘to scratch’ and /k<sup>h</sup>ot<sup>D1</sup>/ ‘to whittle’, corresponding to spirant initials in Middle Vietnamese, will catch the attention of linguists familiar with the Vinh dialect of Vietnamese, another heterodox dialect: see Table 4 (from Ferlus 1991).

Middle Vietnamese		Vinh dialect		gloss
Vietnamese spelling	IPA	with Vietnamese graphophonemic conventions	IPA	
<i>vụng</i>	vuŋ <sup>B2</sup>	<i>phúng</i>	p <sup>h</sup> uŋ <sup>B1</sup>	on the sly
<i>vặt</i>	văt <sup>D2</sup>	<i>phứt</i>	p <sup>h</sup> ut <sup>D1</sup>	to pick, to pluck
<i>vọt</i>	vot <sup>D2</sup>	<i>phót</i>	p <sup>h</sup> ot <sup>D1</sup>	to gush forth
<i>vỡ</i>	vɤ <sup>C2</sup>	<i>phở</i>	p <sup>h</sup> ɤ <sup>C1</sup>	to clear (shrubs)
<i>vỗ</i>	vo <sup>C2</sup>	<i>phỗ</i>	p <sup>h</sup> o <sup>C1</sup>	to clap one's hands
<i>dột</i>	đot <sup>D2</sup>	<i>thốt</i>	t <sup>h</sup> ot <sup>D1</sup>	to drip
<i>dầm</i>	đǣm <sup>A2</sup>	<i>thâm</i>	t <sup>h</sup> ǣm <sup>A1</sup>	drizzle
<i>dỗ</i>	đo <sup>C2</sup>	<i>thỗ</i>	t <sup>h</sup> o <sup>C1</sup>	to entice, to seduce
<i>gàn</i>	ɣan <sup>A2</sup>	<i>khan</i>	k <sup>h</sup> an <sup>A1</sup>	silly
<i>gọt</i>	ɣot <sup>D2</sup>	<i>khót</i>	k <sup>h</sup> ot <sup>D1</sup>	to peel
<i>gãi</i>	ɣaj <sup>C2</sup>	<i>khải</i>	k <sup>h</sup> aj <sup>C1</sup>	to scratch (with nails)
<i>gỡ</i>	ɣɤ <sup>C2</sup>	<i>khở</i>	k <sup>h</sup> ɤ <sup>C1</sup>	to clear up
<i>gõ</i>	ɣɔ <sup>C2</sup>	<i>khỏ</i>	k <sup>h</sup> ɔ <sup>C1</sup>	to knock
<i>gặp</i>	ɣăp <sup>D2</sup>	<i>khấp</i>	k <sup>h</sup> ăp <sup>D1</sup>	to meet
<i>gút</i>	ɣut <sup>D1</sup>	<i>khút</i>	k <sup>h</sup> ut <sup>D1</sup>	knot
<i>gảy</i>	ɣăj <sup>C1</sup>	<i>khảy</i>	k <sup>h</sup> ăj <sup>C1</sup>	to pluck (a string)

Table 4. Correspondences between aspirated initials in the Vinh dialect of Vietnamese and spirants in Middle Vietnamese

The interpretation of these correspondences proposed by Ferlus (1991) is that they result from a situation of contact in which local speakers tried to imitate spirant realizations by speakers of Northern Vietnamese, and ended up producing aspirated stops. (On the irregular tone correspondences, whereby the Standard Vietnamese word has a low-register tone and the Central-Northern Vietnamese word has a high register tone, see section 3.3.3.)

How come Phong Nha has similar correspondences, but in a handful of words only? This clearly looks like a case of borrowing, not regular sound change. The aspiration found in the last two examples in Table 3 are arguably due to the adoption of dialectal forms from Vinh or from another neighbouring dialect that possesses similar forms. While Vinh is inferior to Northern Vietnamese in terms of sociolinguistic prestige, it is a regionally influential variety, and it is therefore not implausible that such forms trickled into Phong Nha Vietnamese. The irregular tonal correspondences (for ‘to bite’ in Table 3, and for all items in Table 5 except ‘to hold’) will be explained in section 3.3.3.

<i>Proto-Vietic</i>	<i>gloss</i>	<i>Vietnamese spelling</i>	<i>Middle Vietnamese</i>	<i>Phong Nha Vietnamese</i>
	spiny amaranth	<i>giền</i> <sup>4</sup>	jen <sup>A2</sup>	cen <sup>A1</sup>
	middle	<i>giũa</i>	juə <sup>C2</sup>	tuə <sup>C1</sup>
k.ciɰ	to keep, to hold	<i>giữ</i>	ju <sup>C2</sup>	tu <sup>B2/C2</sup>
k.cu:j	to pierce, to bore; awl	<i>giùi (lỗ)</i>	juj <sup>A2</sup>	cuj <sup>A1</sup>
k.ra:	old	<i>già</i>	ja <sup>A2</sup>	ta <sup>A1</sup>

Table 5. Correspondences between Proto-Vietic \*c/ɟ, Middle Vietnamese /j/, and Phong Nha Vietnamese /c/ or /t/

<i>Proto-Vietic</i>	<i>gloss</i>	<i>Vietnamese spelling</i>	<i>Middle Vietnamese</i>	<i>Phong Nha Vietnamese</i>
p.se:ʔ	otter	(con) <i>rái</i>	raj <sup>B1</sup>	tɕj <sup>B1</sup>
m.rəpʔ	louse	<i>rận</i>	rɕn <sup>B2</sup>	tɕn <sup>B1</sup>
C.se:t	centipede	<i>rét / rit</i>	ret <sup>D1</sup> / rit <sup>D1</sup>	tit <sup>D1</sup>
p.səpʔ	snake	<i>rắn</i>	rɕn <sup>B1</sup>	tɕn <sup>B1</sup>
p.su:pʔ	navel	<i>rún</i>	run <sup>B1</sup>	ɗun <sup>B1</sup>

Table 6. Correspondences between Proto-Vietic \*s, Middle Vietnamese r, and Phong Nha Vietnamese /t/

‘Louse’ is an odd-man-out in this set, as it is not reconstructed with a \*s in Proto-Vietic. It may be due to a process of hypercorrection or borrowing. The initial /d/ in ‘navel’ is also unexpected; likewise, in Table 7, the initial correspondence between initials for ‘cushion’ is unexpected. The aspirated initial for ‘leak (in the roof)’ (last line in Table 7) is likely to be due to dialectal influence from a dialect sharing the characteristic of the Vinh dialect brought out in Table 4.

<sup>4</sup> In present-day Vietnamese texts, the spelling *dền* is far more common than *giền* for ‘amaranth’, and *dùi* at least as common as *giùi* for ‘awl’. Since the two Middle Vietnamese phonemes at issue, /ð/ and /j/, merged in (standard) modern dialects, there tends to be hesitation as to which is the correct spelling (this is mentioned e.g. by Mai Ngọc Chừ et al. 2005: 123). The palatal found in Phong Nha constitutes strong evidence that the earlier forms had an initial palatal, not a dental, and that the spellings *giền* and *giùi* are etymologically appropriate.

<i>Proto-Vietic</i>	<i>gloss</i>	<i>Vietnamese spelling</i>	<i>Middle Vietnamese</i>	<i>Phong Nha Vietnamese</i>
k.niəw	gizzard	<i>diều (gà)</i>	điəw <sup>A2</sup>	điəw <sup>A2</sup>
?	eagle	<i>diều</i>	điəw <sup>A2</sup>	điəw <sup>A2</sup>
k.ta:lʔ	scrotum, testicles	<i>dái</i>	đaj <sup>B1</sup>	đaj <sup>B1</sup>
k.ta:l	tough (meat)	<i>(thịt) dai</i>	đaj <sup>A1</sup>	đaj <sup>A1</sup>
?	slobber, slaver	<i>dãi</i>	đaj <sup>C2</sup>	đǣj <sup>B1</sup>
	cushion	<i>(gối) dựa</i>	đuə <sup>B2</sup>	tuə <sup>B2/C2</sup>
k.ta:w	knife	<i>dao</i>	đaw <sup>A1</sup>	đaw <sup>A1</sup>
C.ta:	skin	<i>da</i>	đa	đa <sup>A1</sup> ; đǎ <sup>A1</sup>
?	leak (in the roof)	<i>(mái) dột</i>	đot <sup>D2</sup>	tʰot <sup>D1</sup>

Table 7. Correspondences between Proto-Vietic \*t/d, Middle Vietnamese /ð/, and Phong Nha Vietnamese /dʔ/

<i>Proto-Vietic</i>	<i>gloss</i>	<i>Vietnamese spelling</i>	<i>Middle Vietnamese</i>	<i>Phong Nha Vietnamese</i>
bu:ʔ / pu:ʔ	breast	<i>vú</i>	vu <sup>B1</sup>	bu <sup>B1</sup>
	to wash (rice)	<i>vo (gạo)</i>	və <sup>A1</sup>	bo <sup>A1</sup>

Table 8. Correspondences between Proto-Vietic \*p/b, Middle Vietnamese /v/, and Phong Nha Vietnamese /b/

Additionally, Table 9 presents various items that call for further investigation. Their peculiar characteristics may be due to hypercorrection, or to the phonetic interpretation of sounds from other dialects under situations of language contact.

<i>Proto-Vietic</i>	<i>gloss</i>	<i>Vietnamese spelling</i>	<i>Middle Vietnamese</i>	<i>Phong Nha Vietnamese</i>
k.rəŋ	horn	<i>sừng</i>	ʂuŋ <sup>A2</sup>	tʂuŋ <sup>A2</sup>
ʝo:n	weasel	<i>chồn</i>	con <sup>A2</sup>	tɔn <sup>A2</sup>
k.ro:ŋʔ	back of the blade	<i>(cái) sớng (dao)</i>	ʂoŋ <sup>B1</sup>	tɔŋ <sup>B1</sup>
	insect	<i>sâu (bọ)</i>	ʂǣw <sup>A1</sup> (bo <sup>B2</sup> )	tǣw <sup>A1</sup> (bo <sup>B2/C2</sup> )
p.la:ŋ	rack (over the fire)	<i>giàn</i>	jan <sup>A2</sup>	ʂan <sup>A2</sup>
gəl / kəl	tree	<i>cây</i>	kǣj <sup>A1</sup>	kǣn <sup>A1</sup>
k.rɛ:s / k.rɛ:s	root	<i>rễ</i>	re <sup>C2</sup>	rɛn <sup>B2/C2</sup>

Table 9. Items calling for further investigation

To venture some speculations about these items, the Phong Nha forms with initial /tʂ-/ suggest an evolution from \*k.rəŋ to \*k.ləŋ, the cluster \*kl eventually developing into a retroflex (\*k.rəŋ > \*k.ləŋ, and \*k.ro:ŋʔ > \*k.lo:ŋʔ).

### 3.1.4. The preservation of the spirant [ð] in Phong Nha Vietnamese

The Middle Vietnamese phoneme transcribed as *d* in the orthography was a spirant (weak fricative), /ð/. This phoneme is now realized as /z/ in Hanoi Vietnamese and /j/ in Southern Vietnamese. In Phong Nha Vietnamese, cognates of *d*-words can all be pronounced with initial /z/; additionally, some of these words can be pronounced with an initial voiced dental fricative [ð], now frowned upon locally as coarse and ridiculous. Table 10 shows the words at issue, distinguishing those that display synchronic variation (allowing a [ð] variant) from those that do not.

Proto-Vietic	gloss	Middle Vietnamese		Phong Nha Vietnamese
		spelling	IPA interpretation	
	slope	<i>dóc</i>	ðok <sup>D1</sup>	ðok <sup>D1</sup>
	to put out the fire	<i>dập</i>	ðɣp <sup>D2</sup>	ðɣp <sup>D2</sup>
ju:	daughter-in-law	( <i>con</i> ) <i>dâu</i>	ðɣw <sup>A1</sup>	ðu <sup>A1</sup> ; zɣw <sup>A1</sup>
C.ta:	skin, leather	<i>da</i>	ða <sup>A1</sup>	ða <sup>A1</sup> ; đã <sup>A1</sup>
	mother's elder sister	<i>dì</i>	ði <sup>A2</sup>	ði <sup>A2</sup>
	sausage	<i>dôi</i>	ðoj <sup>A2</sup>	ðoj <sup>A2</sup>
	to demolish	<i>dỡ</i>	ðɣ <sup>C2</sup>	ðɣ <sup>B2/C2</sup>
	easy	<i>dễ</i>	ðe <sup>C2</sup>	ðe <sup>B2/C2</sup>
C.ta:l <sup>p</sup>	under	<i>dưới</i>	ðuəj <sup>B1</sup>	ðɣj <sup>B1</sup> / ðuəj <sup>B1</sup>

Table 10. Phong Nha cognates of *d*-initial words in Middle Vietnamese

Finally, Phong Nha presents an interesting treatment of the Proto-Vietic voiced palatal implosive \*ɸ. Proto-Vietic is reconstructed with three voiced implosives: \*ɸ, \*ɸ' and \*ɸ'. They are reflected in Middle Vietnamese (and Modern Vietnamese) as /m/, /n/ and /ɲ/. In Phong Nha, the first two likewise become /m/ and /n/; on the other hand, Phong Nha words derived from Proto-Vietic \*ɸ-initial words display free variation between /ɲ-/ and /z-/, e.g. 'tinder', Middle Vietnamese *nhùi* /ɲuj<sup>A2</sup>/, Phong Nha /zuj<sup>A2</sup> ~ ɲuj<sup>A2</sup>/.

## 3.2. Rhymes

### 3.2.1. Inventory

The fourteen vowel nuclei of Phong Nha Vietnamese are the same as in Standard Vietnamese: nine vowels /i e ε a u u ɣ o ɔ/, two of which have phonemically distinct short counterparts: /ǎ/ and /ɣ̃/, and three diphthongs, /iə uə uə/. Short vowels are a majority in the system: there only exist two short vowels (Haudricourt 1952), and it is therefore economical to use a diacritic for these two short vowels and leave long vowels unmarked.

In syllable-final position, the following consonants are found: /-p -t -k -m -n -ŋ -j -w/. Combinations of vowels with the medial glide /<sup>w</sup>/ and with final consonants follow the same patterns as in Middle Vietnamese. For ease of reference, Table 11 provides a phonemic analysis of the rhymes of Middle Vietnamese. This table was created by Michel Ferlus in 1991; the analysis is essentially the same as that proposed by Cao Xuân Hạo (2007: 102). The hyphen (-) in the top row (in -k, <sup>w</sup>-m, etc.) materializes the position of the vowel within the consonantal structure of the syllable. The notion of rime is understood in a broad sense that includes the rounding of labialized initials. This rounding is transcribed as a superscript <sup>w</sup> to the left of the symbol C (for Consonant). Thus, the rime *uê*, as in the word *quê* ‘countryside, home village’, is located in column C<sup>w</sup>-, line e. A dash (—) indicates that the combination at issue is not found in the language. Some of the complexities of the orthography are not reflected in the table, such as the encoding of the contrast between the rhymes /<sup>w</sup>a/ and /uə/ by the use of different consonant symbols (and identical vowel symbols) in *qua* /kwa/ vs. *cua* /kuə/.

The choice of IPA symbols for vowels is based on Kirby (2011). It is not a narrow notation aiming at the greatest synchronic phonetic precision. For instance, the vowels *ơ* and *â* are transcribed as /ɤ/ and /ǣ/ respectively, reflecting their interpretation as a vowel pair distinguished by phonemic length, and overlooking the slight difference between them in terms of vowel quality – a difference which led Gsell (1980) to adopt a transcription by /ɤ/ and /ʌ/ respectively.

Velar finals after the back rounded vowels, /u o ɔ/, have an additional labial closure, as in Hanoi Vietnamese. For instance, ‘slope’, /ðɔk<sup>24</sup>/, is realized as [ðɔk<sup>P24</sup>], and ‘to use’, /ðuŋ<sup>32</sup>/, as [ðuŋ<sup>m32</sup>]. In Hanoi Vietnamese, the vowel quality of /o/ and /ɔ/ is affected: the actual pronunciation of /ok/ /ɔk/ /oŋ/ /ɔŋ/ is approximated as [ʌuk<sup>p</sup>] [ɛuk<sup>p</sup>] [ʌuŋ<sup>m</sup>] [ɛuŋ<sup>m</sup>] by Thompson (1984), and as [əuk<sup>p</sup>] [auk<sup>p</sup>] [əuŋ<sup>m</sup>] [auŋ<sup>m</sup>] by Henderson (1985: 21). It can also be transcribed as [ɤok<sup>p</sup>] [ʌɔk<sup>p</sup>] [ɤoŋ<sup>m</sup>] [ʌɔŋ<sup>m</sup>], “to reflect the intuition that lip rounding is moved to the right, the onset of the diphthong consisting in the unrounding of the original vowel: from [o] to [ɤo], from [ɔ] to [ʌɔ]” (Michaud 2004). In Phong Nha, on the other hand, the change is less advanced, and the quality of the vowel before velar finals is still readily identifiable.

	C-C <sup>w</sup> -	C-k C <sup>w</sup> -k	C-ŋ C <sup>w</sup> -ŋ	C-t C <sup>w</sup> -t	C-n C <sup>w</sup> -n	C-p —	C-m C <sup>w</sup> -m	C-j C <sup>w</sup> -j	C-w C <sup>w</sup> -w
i	i uy	ich uyeh	inh uynh	it uyt	in —	ip —	im —	— —	iu uyu
e	ê uê	êch uêch	ênh uênh	êt uêt	ên uên	êp —	êm —	— —	êu —
ɛ	e oe	ach oach	anh oanh	et oet	en oen	ep —	em —	— —	eo oeo
iə	ia uya	iêc —	iêng —	iêt uyêt	iên uyên	iêp —	iêm —	— —	iêu —
u	u —	uc —	ung —	ut —	un —	— —	— —	ui —	uu —
ɤ	ɔ uɔ	— —	— —	ɔt uɔt	on uon	ɔp —	om —	oi —	— —
ɤ̃	— —	âc —	âng —	ât uât	ân uân	âp —	âm —	ây uây	âu —
a	a oa	ac oac	ang oang	at oat	an oan	ap —	am oam	ai oai	ao oao
ă	— —	ăc ăc	ăng ăng	ăt ăt	ăn ăn	ăp —	ăm ăm	ay oay	au —
uə	ua —	uoc —	uong —	uot —	uon —	uop —	uom —	uoi —	uou —
u	u —	uc —	ung —	ut —	un —	up —	um —	ui —	— —
o	ô —	ôc —	ông —	ôt —	ôn —	ôp —	ôm —	ôi —	— —
ɔ	o —	oc —	ong —	ot —	on —	op —	om —	oi —	— —
uə	ua —	uôc —	uông —	uôt —	uôn —	— —	uôm —	uôi —	— —

Table 10. The rhymes of Middle Vietnamese



### 3.2.2. Comparative perspectives

Rhyme correspondences between Middle Vietnamese and “heterodox” dialects show a number of irregularities. The Vietnamese system is highly innovative: two Vietnamese vowels typically correspond to a single vowel in proto-Vietic (Ferlus 1997b). This two-way split does not correspond to the familiar effects of consonant shifts. It has been proposed that it results from dialect mixture (Ferlus 1997b:50): two dialects came in contact, one of which was conservative whereas the other had undergone opening diphthongization of low vowels /i:/, /u:/, /i/, /i/, followed by closing diphthongization of open vowels /a:/, /o:/ and /ɔ:/ . The coexistence and eventual merger of these two dialects – presumably in the Red River delta – resulted in a haphazard lexical distribution of the two vowel sets, and hence in a multiplication of the number of contrastive vowels. The resulting language variety endured to become dominant; all the language varieties resulting from its gradual spread, which constitute the core present-day Vietnamese dialects, possess this large inventory of vowels. Some “heterodox” dialects, however, do not show the diphthongized forms characteristic of Standard Vietnamese (from Middle Vietnamese to the present-day “orthodox” Northern and Southern dialects). Phong Nha is a case in point: it preserves reflexes of Proto-Vietic \*a:, \*ə:, \*ɔ: and \*ɛ: that appear essentially unchanged since Proto-Vietic. Examples are provided in Tables 12 to 15. Note that all Proto-Vietic reconstructions in the present article are from ongoing comparative work (Ferlus in preparation).

<i>Proto-Vietic</i>	<i>gloss</i>	<i>Vietnamese spelling</i>	<i>Middle Vietnamese</i>	<i>Phong Nha Vietnamese</i>
da:ŋ	path	<i>đường</i>	ɗuəŋ <sup>A2</sup>	ɗaŋ <sup>A2</sup>
ɗa:k	water	<i>nước</i>	nuək <sup>D1</sup>	nak <sup>D1</sup>
*C.la:s	fire	<i>lửa</i>	luə <sup>C1</sup>	la <sup>C1</sup>
ŋa:j	human being	<i>người</i>	ŋuəj <sup>A2</sup>	ŋaj <sup>A2</sup>
m.ra:ʔ	bush-hook	<i>rựa</i>	ruə <sup>B2</sup>	ra <sup>B2/C2</sup>
*C.la:jʔ	net	<i>lưới</i>	luəj <sup>B1</sup>	laj <sup>B1</sup>
la:s	tongue; ploughshare; blade <sup>5</sup>	<i> lưỡi</i>	luəj <sup>C2</sup>	laj <sup>B2/C2</sup>
ma:jʔ	to borrow	<i>mượn</i>	muəŋ <sup>B2</sup>	man <sup>B2/C2</sup>

Table 12. Proto-Vietic \*a:, Middle Vietnamese /uə/ (spelling: ua/uə), Phong Nha /a/

<sup>5</sup> These three words constitute developments from the same root, not just homophones.

<i>Proto-Vietic</i>	<i>gloss</i>	<i>Vietnamese spelling</i>	<i>Middle Vietnamese</i>	<i>Phong Nha Vietnamese</i>
C.ta:l	under	<i>dưới</i>	ðuəj <sup>B1</sup>	ðɤj <sup>B1</sup> / ðuəj <sup>B1</sup>
g.lə:t/b.la:t	to slip	<i>trượt</i>	tuət <sup>D2</sup>	tɤt <sup>D2</sup>

Table 13. Proto-Vietic \*ə:/\*a:, Middle Vietnamese /uə/ (spelling: ua/uə), Phong Nha /ɤ/

It may be that the vowel /ɤ/ in the Phong Nha examples in Table 13 represents the result of the reinterpretation of the vowel /uə/ of orthodox dialects, rather than a regular phonetic development.

<i>Proto-Vietic</i>	<i>gloss</i>	<i>Vietnamese spelling</i>	<i>Middle Vietnamese</i>	<i>Phong Nha Vietnamese</i>
rɔ:ŋ <sup>?</sup>	paddy field	<i>ruộng</i>	ruəŋ <sup>B2</sup>	rɔŋ <sup>B2/C2</sup>
m.rɔ:j	flea	<i>ruồi</i>	ruəj <sup>A2</sup>	rɔj <sup>A2</sup>
ə.rɔ:	tortoise	<i>rùa</i>	ruə <sup>A2</sup>	rɔ <sup>A2</sup>
rɔ:c	intestines	<i>ruột</i>	ruət <sup>D2</sup>	rɔt <sup>D2</sup>
ɓɔ:j <sup>?</sup>	salt	<i>muối</i>	muəj <sup>B1</sup>	mɔj <sup>B1</sup>
mɔ:s	mosquito	<i>muỗi</i>	muəj <sup>C2</sup>	mɔj <sup>B2/C2</sup>
ə.lɔ: <sup>?</sup>	paddy	<i>lúa</i>	luə <sup>B1</sup>	lɔ <sup>B1</sup>
ɗɔ:c	to swallow	<i>nuốt</i>	nuət <sup>D1</sup>	nɔt <sup>D1</sup>

Table 14. Proto-Vietic \*ɔ:, Hanoi Vietnamese uô/ua (IPA /uə/), Phong Nha /ɔ/

<i>Proto-Vietic</i>	<i>gloss</i>	<i>Vietnamese spelling</i>	<i>Middle Vietnamese</i>	<i>Phong Nha Vietnamese</i>
ʔeh ‘to defecate’ cas ‘to flow’	diarrhoea	<i>ỉa chảy</i>	iə <sup>C1</sup> teāj <sup>C1</sup>	ɛ <sup>C1</sup> cāj <sup>C1</sup>

Table 15. Proto-Vietic \*ɛ:, Middle Vietnamese /iə/ (spelling: iê/ia), Phong Nha /ɛ/ (the relevant syllable is the first in the disyllabic form)

There is only one example in Table 15: this is the only example found so far of Phong Nha /ɛ/ corresponding to Proto-Vietic \*ɛ: and Vietnamese /iə/ (*iê/ia*).

Another notable feature of the system is the presence of a short /ă/ or /ɤ/ in some syllables with a /-aj/ rhyme in Middle Vietnamese. Table 16 presents relevant examples, and clarifies that they derive from Proto-Vietic \*e: whereas Proto-Vietic \*a: plus final \*s or \*l yields a main vowel /a/ in Phong Nha. Note that Proto-Vietic \*ja:l ‘casting-net’ is an early borrowing from Sanskrit *jāla*.

Proto-Vietic		gloss	Vietnamese spelling	Middle Vietnamese	Phong Nha	Phong Nha main vowel
rhyme	form					
*e:	tke:	thorn	<i>gai</i>	ɣaj <sup>A1</sup>	kǎj <sup>A1</sup>	ǎ
	t.ke:?	female	<i>cái</i>	kaj <sup>B1</sup>	kǎj <sup>B1</sup>	ǎ
	ke:?	CLF	<i>cái</i>	kaj <sup>B1</sup>	kǎj <sup>B1</sup>	ǎ
	ple:?	fruit	<i>trái</i>	ʈaj <sup>B1</sup>	ʈǎj <sup>B1</sup>	ǎ
	pse:?	otter	<i>rái</i>	raj <sup>B1</sup>	tǎj <sup>B1</sup>	ǎ
	kde:?	stag, deer	<i>(con) nai</i>	naj <sup>A1</sup>	nǎj <sup>A1</sup>	ǎ
*a: plus	ja:l	casting-net	<i>chài</i>	caj <sup>A2</sup>	caj <sup>A2</sup>	a
final *s or	p.la:l	son	<i>(con) trai</i>	ʈaj <sup>A1</sup>	ʈaj <sup>A1</sup>	a
*l	C.ka:s	to scratch	<i>gãi</i>	ɣaj <sup>C2</sup>	k <sup>h</sup> aj <sup>C1</sup>	a

Table 16. Correspondences for syllables with a /-aj/ rhyme in Middle Vietnamese

The double set of reflexes of Proto-Vietic \*e in Phong Nha (/ǎj/ and /ǎj/) is unexplained, but not especially surprising. “Orthodox” Vietnamese did not undergo a linear development from Proto-Vietic; while Proto-Vietic \*e: mostly evolves into /aj/ (orthographic *ai*), there are also numerous instances of /ǎj/ (orthographic *ây*). Likewise, Proto-Vietic \*o: has two reflexes, /aw/ (orthographic *ao*) and /ǎw/ (orthographic *âu*).

As for Proto-Vietic \*o, only two examples are presented in Table 17; further data, comprising Phong Nha cognates for words such as *bao* ‘dream’, will be necessary to pursue the analysis. The mention ‘local voc.’ indicates that the word at issue (‘stream, river’) is part of local vocabulary that cannot be traced back to Proto-Vietic.

Proto-Vietic		gloss	Vietnamese spelling	Middle Vietnamese	Phong Nha Vietnamese	
vowel	form				main vowel	form
*o	local voc.	stream, river	rào	raw <sup>A2</sup>	a	raw <sup>A2</sup>
*o	t.ko:?	husked rice	gạo	ɣaw <sup>B2</sup>	a	ɣaw <sup>B2/C2</sup>
*o	s.po:	to dream	bao	ɓaw <sup>A1</sup>	?	?

Table 17. Correspondences for reflexes of Proto-Vietic \*o

### 3.3. The tone system

#### 3.3.1. Inventory

The tonal system is presented in Table 18. The table is arranged by etymological categories, A1 to D2 (Haudricourt 1972; Gedney 1972). The phonetic labels use Chao Yuen-ren’s scale from 1 (lowest) to 5 (highest); they approximate our auditory impression of the tones, pending further examination of their acoustic and perceptual properties. Tone A1 is rising, in the higher part of the speaker’s

range; A2 is lower, and slightly falling; B1 sounds falling-rising; the tone that results from the merger of B2 and C2 is very low; and tone C1 ends in glottal constriction. No tone sandhi was observed (on tone coarticulation in Vietnamese, see Brunelle 2003, 2009).

	A	B	C	D
1	45	324	2 <sup>2</sup>	24
2	32	1		53

Table 18. Phong Nha tones. Categories A, B, C: non-checked syllables (smooth syllables); category D: checked syllables (stop-final syllables)

### 3.3.2. Origin of the five tones: four tones plus one, or six tones minus one?

In the Vietic languages Pong/Phong (Laos) and Toum and Liha (spoken in the north-central provinces of Vietnam), there exist four-tone systems, following the confusion of etymological B1 with C1, and B2 with C2 (Ferlus 1998). A four-tone system (on non-checked syllables) is also found in the Cao Lao Hạ dialect of Vietnamese (Ferlus 1995), in which the etymological C category (originating in syllables with final \*-h) is now confused with etymological B2: i.e. Vietnamese B2, C1 and C2 are all reflected as the same tone in Cao Lao Hạ. Seen in this light, the five-tone system of Phong Nha Vietnamese may be due to the preservation of five of the six etymological categories; or it may be due to later borrowings that reintroduced a fifth category into a four-tone system (reintroducing a formerly lost opposition between B1 and C1). The way to test these two hypotheses is to evaluate the degree of regularity in tonal correspondences with Middle Vietnamese: if a category was reintroduced through borrowing, irregular correspondences are expected. Comparison argues in favour of the former hypothesis: Phong Nha has never been a member of the set of four-tone dialects.

### 3.3.3. Irregular tonal correspondences across dialects

Overall, tonal correspondences between Phong Nha and Middle Vietnamese are straightforward. There are a few cases of irregular tonal correspondences, falling in several categories. A reminder may be in order concerning about the three types of irregular tonal correspondences brought out by Ferlus (1999).

(i) The first type is due to the devoicing of sonorants in sesquisyllables in Proto-Việt-Mường, resulting in high-register reflexes in Northern Vietic (Vietnamese and Mường) vs. low-register in Southern Vietic (all the other languages of the group). For example, sesquisyllabic Proto-Vietic \*k.ma ‘rain’ yielded a high-series tone (A1) in Vietnamese: /muə<sup>A1</sup>/ (orthography: *mưa*). The proposed explanation is that, under the influence of Chinese (which had already undergone monosyllabicization), \*k and \*m came to be articulated as an initial consonant cluster \*km-, and their strong coarticulation detracted from the voicing of the nasal, so that it patterned together with voiceless onsets at the stage when a

consonant shift among initials took place (Haudricourt 1965, 1972). In Southern Vietic, on the other hand, voicing of the \*m was preserved through the sesquisyllabic realization of \*k.ma as [\*k<sup>a</sup>ma], hence the development of a low-series tone, e.g. Pong /k<sup>a</sup>ma<sup>A2</sup>/ and Liha /ma<sup>A2</sup>/.

(ii) The second type is due to the voicing of medial stops in Vietnamese after an initial \*r, resulting in low-register reflexes in Vietnamese vs. high-register in all the other languages (including Mường). For example, \*r.ka ‘chicken’ yields Vietnamese gà [ɣa<sup>A2</sup>], vs. Mường, Cuối, Pong and Thavung /ka<sup>A1</sup>/, Sách/Rục /r<sup>a</sup>ka<sup>A1</sup>/, Arem /l<sup>a</sup>kæ<sup>(1)</sup>/, and Mã Liềng /ũka<sup>A1</sup>/.

(iii) The third type is due to the loss of Proto-Austroasiatic initial voicing in Northern Vietic languages, resulting in high-register reflexes, as against the expected low-register reflexes in Southern Vietic. This is interpreted by Ferlus (1999) as a substratum effect dating back to the time when Proto-Vietic spread northwards onto an Austroasiatic substratum of languages that lacked voiced stops (a set of languages of which the Khmuic language Ksing Mul arguably constitutes a remnant).

Some Phong Nha words illustrate the above types of irregular tonal correspondences categories. For instance, the word ‘louse’, Phong Nha /tʃn<sup>B1</sup>/, constitutes an instance of the second type of irregularity. It differs from Standard Vietnamese /rʃn<sup>B2</sup>/ *rận* by its initial and by its tone. In Vietnamese, the medial stop became voiced as a consequence of spirantization, resulting in the development of a low-series tone. Other examples include ‘to whittle, to cut out’, Phong Nha /k<sup>h</sup>ət<sup>D1</sup>/, and ‘to bite’, Phong Nha /kăm<sup>B1</sup>/, that have non-spirantized initials and high-register tones; compare Standard Vietnamese /ɣət<sup>D2</sup>/ *got* and /ɣăm<sup>B2</sup>/ *gặm*.

In addition to these three types, some mixed correspondences are found, however: cases where the word is a hybrid – a combination of phonological materials from Vietnamese and from another Vietic language. The Phong Nha word for ‘bedbug’, /rep<sup>D1</sup>/, is identical to the Standard Vietnamese word, except for its tone (Standard Vietnamese /rep<sup>D2</sup>/ *rệp*). Tone D1 is the regular tone of this word in all Vietic languages except Vietnamese: Mường /se:t<sup>D1</sup>/ or /t<sup>h</sup>e:t<sup>D1</sup>/, Pong /si:p<sup>D1</sup>/, Thavung, Sách and Rục /k<sup>a</sup>si:p<sup>D1</sup>/, Maleng Bro /krsɪ:p<sup>D1</sup>/). The Phong Nha word thus contains Standard Vietnamese segments (including the telltale spirant which caused the irregular tonal correspondence with the rest of Vietic) and a typically non-Vietnamese tone. Given the higher prestige of Vietnamese, it is a safe guess that /rep<sup>D1</sup>/ results from the modification of an earlier form, which had an initial /\*s-/ , by speakers who adopted the segments of the Standard Vietnamese word, while retaining the tone of the earlier form. The initial, and perhaps the vowel, were acquired through a process of standardization, which did not affect the tone.

Such a change is documented in other dialects. In Hanoi, ‘coffee’ was initially borrowed from the French /kafe/ as *trà phê* /tʰa<sup>A2</sup>.fe<sup>A1</sup>/ (Martini 1958: 337–338). The first syllable is a phonetic stretcher, as the tone is different from that (A1)

affected by default to borrowings from toneless languages, and the initial /t/ does not match the /k/ in /kafe/. This syllable is semantically motivated: it is none other than the Sino-Vietnamese word for ‘tea’, /t̚a<sup>A2</sup>/ *trà*. Labelling coffee as a type of tea acclimatizes the new beverage, introducing it into a set that previously included *trà tàu* ‘Chinese tea (black tea)’, *trà Huế* ‘Annamese tea (green tea)’, *trà hoa* ‘camellia tea’, *trà rìng* ‘three-seeded mercury (herbal) tea’, and so on. As for the second syllable, /fe<sup>A1</sup>/ *phê*, it constitutes both a straightforward phonetic rendering of the syllable /fe/ and a semantically appropriate syllable, as coffee is apt to produce a sensation of elation not entirely unlike that produced by opium – another commodity associated with foreigners, which was actively promoted by colonial authorities (Le Failler 2001) –, whose psychotropic effects are evoked in Vietnamese through the expressive form *phê phê*. On the other hand, Vietnamese people with a command of French would pronounce the foreign word as /ka<sup>A1</sup> fe<sup>A1</sup>/, with tone A1 (a level, non-low tone, used as the default for foreign syllables without a final stop). A hybrid of the two forms emerged: /ka<sup>A2</sup> fe<sup>A1</sup>/ *cà phê*, correcting the initial to /k/ as in the donor language, but retaining the tone of /t̚a<sup>A2</sup>/ *trà* ‘tea’. This form became standard (see e.g. Nguyễn Như Ý 1999), to the puzzlement of the linguist Emeneau, who tentatively hypothesized that the A2 tone in /ka<sup>A2</sup> fe<sup>A1</sup>/ must be due to an (implausible) assimilation to /ka<sup>A2</sup>/ *cà* ‘eggplant’ (Emeneau 1951: 4, 158). The process is similar to that found in Phong Nha for the word ‘bedbug’: the partial correction of a form to make it sound closer to a perceived standard form, without replacing it altogether.

#### 4. NOTES ABOUT THE MIGRATION HISTORY OF PHONG NHA

This section proposes notes on ways to analyze the historical situation that resulted in the linguistic patterns currently observed in Vietnamese dialects. A key issue in studying dialects is language contact among related language varieties.

##### 4.1. *The paucity of available information about migrations*

The history of migration can shed light on situations of language contact. But historical records about population movements in the area of present-day Northern and Central-Northern Vietnam mostly concern the southerly flux of settlers from present-day South China into present-day Northern Vietnam, reported in Chinese chronicles.<sup>6</sup> These sources are recapitulated in the form of a

---

<sup>6</sup> In the 3<sup>rd</sup> century BC, a group from the Shu (蜀) kingdom (Sino-Vietnamese reading: Thục), in the Sichuan basin, moved south and established a dynasty in the Red River delta; this attempt to escape from the domination of the kingdom of Qin (秦), which had conquered Shu, paradoxically resulted in the eventual integration of the Red River delta to the emerging Chinese empire. Ma Yuan 馬援, sent by the Eastern Han dynasty in 42 to 43 AD, crushed a

six-volume *History of Migrations in China* (Ge Jianxiong 1997). No compendium of comparable scope is available at present concerning the history of population movements inside Vietnam. The overall pattern during the past millenium is clear – the settling of the centre and south of present-day Vietnam through rural migration from North to South – but the process remains little-documented in its details. It seems as if few large-scale efforts had been made, during the successive Vietnamese dynasties, to compile detailed documentation on topics of local history – one of the consequences of the preference of successive Vietnamese emperors for Chinese culture and Chinese history. (See e.g. Langlet 1990: 105–184 on the constant emphasis placed on Chinese cultural models by emperor Gia Long, and the consequences for the last Vietnamese dynasty.) Useful information can occasionally be gleaned from imperial annals, and regional monographs by mandarins (for a detailed example, see Ferlus 2003); but these sources essentially testify to the fact that remote areas were ill-understood by imperial authorities, and of little interest to them (Le Failler 2014: 16). By the first half of the 20<sup>th</sup> century, the assessment of the situation by the geographer Pierre Gourou (1936: 179) was that “the official statistics are based on too fanciful data to provide any adequate idea of the evolution of the population”. Data collection was complicated by the fact that local communities were distrustful of population surveys as foreboding increases in taxes and other demands. The 1989 census has been considered as the first truly reliable census (Banens 2000), and contemporary studies of migration mostly concern developments since the 20<sup>th</sup> century (e.g. about the settlement of the highlands by migrants from the Red River delta: Hardy 2003; and present-day Chinese immigration: Nguyen 2013).

#### ***4.2. Piecemeal information about the settlement of Phong Nha***

As a first step towards understanding the history of the settlement of Phong Nha, information provided by the Phong Nha speakers clarifies that the Phong Nha area used to be known by the name of Chùa Nghe, and inhabited by Chút populations, i.e. speakers of languages of the Southern Vietic subgroup of Austroasiatic (“Chút” includes the following subgroups: Rục, Máy, Sách, Arem, and Mã Liềng). Ethnic Vietnamese (Kinh) settlers only moved into the area four to five generations ago. The first to settle in Phong Nha are reported to have carried the family name Dương (楊, or perhaps 陽 or 羊), and to have moved in from a neighbouring village, Cù Lạc, located in the same county (xóm Cù Lạc, xã

---

rebellion against the Chinese administration, and established the Red River delta as a Chinese province (Maspero 1918: 14–24), “thereby initiating the process whereby it acquired the solid Chinese framework that allowed it to play the leading role in East Indochina from the 10<sup>th</sup> century onward” (Maspero 1918: 28).

Sơn Trạch, huyện Bồ Trạch, Quảng Bình). Other settlers from various areas of Nghệ An and Hà Tĩnh provinces gradually joined the village, by small groups of about one to three families. Family names found in Phong Nha are, by reported order of successive arrival: Dương, Lê (黎 or 梨), Đinh (丁), Nguyễn (阮), Trần (陳) and Hoàng (黃). Mr. Hoàng Minh Chiêm reported that his ancestors had moved to Phong Nha three to four generations ago, coming from the commune of Quảng Văn, Quảng Trạch county (xã Quảng Văn, huyện Quảng Trạch, tỉnh Quảng Bình). Language contact appears to have taken place predominantly among the Vietnamese dialects of the settlers, as no marriages are reported between the “majority” (Kinh) and “minority” (Arem and Chứt) populations. The first Kinh settlers in Quảng Bình are reported to have arrived around the year 1300 AD, with an increase in migratory inflow in the 15<sup>th</sup> and 16<sup>th</sup> centuries, mostly occupying the more level lands (Nguyễn Văn Lợi & Nguyễn Văn Mạnh 2010: 27–34).

The history of the settlement of Phong Nha by speakers from different areas of North-Central Vietnam would call for detailed comparison with these dialects, in order to understand phonological correspondences one after the other. The road ahead is clear: a fine-grained study of processes of diffusion across dialects would require in-depth descriptions of the dialects spoken in the places of origin of the migrants – and a reconstruction of the state of these dialects at the time of migration.

## 5. CONCLUSION

Examination of Phong Nha data suggests the presence of several strata of standardization: layers of influence from prestige varieties of the language. Some words are currently identical with Standard Vietnamese, presumably through relatively recent adoption (borrowing). The word for ‘husked rice’, Phong Nha /yaw<sup>B2/C2</sup>/, displays spirantization and a low-series tone, like in Standard Vietnamese: /yaw<sup>B2</sup>/ *gạo*. Others are not of Vietnamese stock, but originate in another Vietic language, for instance ‘to plant (a tree)’, Phong Nha /loŋ<sup>A1</sup>/, Middle Vietnamese *trồng* [toŋ<sup>A2</sup>]. The word is reconstructed to Proto-Vietic \*m.loŋ; among Vietic languages, Vietnamese alone has a high-series tone for this word.<sup>7</sup> No explanation can be proposed at present concerning the development of

---

<sup>7</sup> The comparative word list proposed by Nguyễn Văn Tài (2005: 283) has tone A2 for almost half of the 29 language varieties for which a form for ‘to plant’ is reported. In detail, tonal notations in the late Pr. Nguyễn Văn Tài’s volume need to be used with caution, however, as it seems that this author found it difficult to leave aside the prism of his native language’s tone system when categorizing aurally the tones of other Vietic languages. There may also have been an issue of “contaminated tonal transcriptions” (Phan 2012: 6; Phan 2013: 315–316) due to the consultants’ production of words with their (“orthodox”) Vietnamese tone (used by the investigator at elicitation) instead of the tone in the language variety under study.



tone for this item in Vietic; this is an example of the many issues that remain to be investigated in the field of Vietic studies. A concern here is that, in view of the small number of linguists at work in this area, it is far from certain that in-depth linguistic surveys can be carried out before the influence of more prestigious dialects erases the features that constitute keys to historical reconstruction (for an overview: Trần Trí Dõi 2000, 2003). Here as in other parts of the world, the importance of classical linguistic fieldwork (Dixon 2007) cannot be overemphasized.

## REFERENCES

- Alves, Mark. 2007. A look at North-Central Vietnamese. In Ratreé Wayland, John Hartmann & Paul Sidwell (eds.), *SEALS XII: Papers from the 12th meeting of the Southeast Asian Linguistics Society (2002)*, 1–7. Canberra: Pacific Linguistics.
- Alves, Mark & Nguyễn Duy Hương. 2007. Notes on Thanh-Chương Vietnamese in Nghệ-An Province. In Mark Alves, Paul Sidwell & David Gil (eds.), *SEALS VIII: Papers from the 8th meeting of the Southeast Asian Linguistics Society (1998)*, 1–9. Canberra: Pacific Linguistics.
- Banens, Maks. 2000. *Vietnam: a reconstitution of its 20th century population history*. Tokyo: Institute of Economic Research, Hitotsubashi University. <https://hal.archives-ouvertes.fr/hal-00369251/>.
- Brunelle, Marc. 2003. Tone Coarticulation in Northern Vietnamese. *Proceedings of the 15th International Congress of Phonetic Sciences*, 2673–2676. Barcelona.
- Brunelle, Marc. 2009. Northern and Southern Vietnamese tone coarticulation: A comparative case study. *Journal of Southeast Asian Linguistics* 1. 49–62.
- Cadière, Léopold. 1902. *Phonétique annamite (dialecte du Haut-Annam) [Annamese phonetics (Haut-Annam dialect)]*. (Publications de l'École Française d'Extrême-Orient, vol. III.) Paris: Ernest Leroux.
- Cao Xuân Hạo. 2007. Thêm mấy giải pháp âm vị học cho các vận mẫu có nguyên âm ngắn của tiếng Việt [Additional solutions for the phonemic analysis of Vietnamese rhymes containing short vowels]. *Tiếng Việt: mấy vấn đề ngữ âm, ngữ pháp, ngữ nghĩa [Vietnamese: issues of phonetics, grammar and semantics]*, 103–115. T.P. Hồ Chí Minh: NXB Giáo dục.
- Dixon, Robert M. 2007. Field linguistics: a minor manual. *Sprachtypologie und Universalienforschung* 60(1). 12–31.
- Emeneau, Murray Barnson. 1951. *Studies in Vietnamese (Annamese) grammar*. (University of California Publications in Linguistics.) Berkeley: University of California Press.
- Ferlus, Michel. In preparation. Proto-Viet-Muong (Proto-Vietic).

- Ferlus, Michel. 1979. Formation des registres et mutations consonantiques dans les langues mon-khmer [Development of phonation-type registers and consonant shifts in Mon-Khmer languages]. *Mon-Khmer Studies* 8. 1–76.
- Ferlus, Michel. 1982. Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien [Spirantization of medial obstruents and development of the Vietnamese consonant system]. *Cahiers de linguistique - Asie Orientale* 11(1). 83–106. Updated version available from <https://halshs.archives-ouvertes.fr/halshs-01063845>
- Ferlus, Michel. 1991. Le dialecte vietnamien de Vinh. *24th International Conference on Sino-Tibetan Languages and Linguistics*. Ramgambaeng University, Bangkok. <https://halshs.archives-ouvertes.fr/halshs-00922737>
- Ferlus, Michel. 1995. Particularités du dialecte vietnamien de Cao Lao Hạ (Quảng Bình, Vietnam) [Peculiarities of the Vietnamese dialect of Cao Lao Hạ (Quảng Bình, Vietnam)]. *Dixièmes Journées de Linguistique d'Asie Orientale*. Paris. <http://halshs.archives-ouvertes.fr/halshs-00922735>
- Ferlus, Michel. 1996. Langues et peuples viet-muong [Viet-Muong languages and peoples]. *Mon-Khmer Studies* 26. 7–28.
- Ferlus, Michel. 1997a. Compte rendu d'exposé: Evolution vers le monosyllabisme dans quelques langues de l'Asie du Sud-Est [Evolution towards monosyllabism in some languages of Southeast Asia; short report on talk given at the Société de Linguistique de Paris]. *Bulletin de la Société de Linguistique de Paris* 92(1). XVII–XVIII. Full handout available from: <http://halshs.archives-ouvertes.fr/halshs-00927456>
- Ferlus, Michel. 1997b. Problèmes de la formation du système vocalique du vietnamien [Issues in the development of the Vietnamese vowel system]. *Cahiers de linguistique - Asie Orientale* 26(1). 37–51.
- Ferlus, Michel. 1998. Les systèmes de tons dans les langues viet-muong [Tone systems in Viet-Muong languages]. *Diachronica* 15(1). 1–27.
- Ferlus, Michel. 1999. Les disharmonies tonales en viet-muong et leurs implications historiques [Irregular tone correspondences among Vietic languages and their historical implications]. *Cahiers de Linguistique - Asie Orientale* 28(1). 83–99.
- Ferlus, Michel. 2001. Les hypercorrections dans le thỏ de Làng Lỡ (Nghệ An, Vietnam) ou les pièges du comparatisme [Hypercorrections in the Thỏ dialect of Làng Lỡ (Nghệ An, Vietnam): an example of pitfalls for comparative linguistics]. *Quinzièmes Journées de Linguistique de l'Asie Orientale*. Ecole des Hautes Etudes en Sciences Sociales, Paris. <http://halshs.archives-ouvertes.fr/halshs-00922722>
- Ferlus, Michel. 2003. L'intérêt linguistique du Hưng Hóa Ký Lược (Monographie de la province de Hưng Hóa) de Phạm Thân Duật [The usefulness of Phạm Thân Duật's Hưng Hóa Ký Lược (Monograph of Hưng Hóa province) for linguistic research]. Paris. <https://halshs.archives-ouvertes.fr/halshs-00922777>

- Ferlus, Michel. 2004. The Origin of Tones in Viet-Muong. In Somsongue Burusphat (ed.), *Papers from the Eleventh Annual Conference of the Southeast Asian Linguistics Society 2001*, 297–313. Tempe, Arizona: Arizona State University Programme for Southeast Asian Studies Monograph Series Press.
- Ferlus, Michel. 2014. Arem, a Vietic language. *Mon-Khmer Studies* 43. 1–15.
- Ge Jianxiong (葛剑雄). 1997. *Zhōngguó yímín shǐ* 中国移民史 [A History of Migrations in China]. 6 vols. Fújiàn rénmin chūbǎnshè 福建人民出版社 (Fujian People's Publishing House).
- Gedney, William J. 1972. A checklist for determining tones in Tai dialects. In Estelle Smith (ed.), *Studies in Linguistics in honor of George L. Trager*, 423–437. The Hague: Mouton.
- Gourou, Pierre. 1936. *Les paysans du delta tonkinois. Etude de géographie humaine [The peasants of the Tonkin delta: a study in human geography]*. Paris: Publications de l'Ecole Française d'Extrême-Orient.
- Gsell, René. 1980. Remarques sur la structure de l'espace tonal en vietnamien du sud (parler de Saïgon) [Observations about the structure of the tone space in Southern Vietnamese (Saigon dialect)]. *Cahier d'études vietnamiennes, Département de Langues et Civilisations de l'Asie Orientale de l'Université Paris 7* 4.
- Hardy, Andrew. 2003. *Red hills: migrants and the state in the highlands of Vietnam*. Copenhagen: Nordic Institute of Asian Studies.
- Haudricourt, André-Georges. 1952. Les voyelles brèves du vietnamien [The short vowels of Vietnamese]. *Bulletin de la Société de Linguistique de Paris* 48(1). 90–93.
- Haudricourt, André-Georges. 1965. Les mutations consonantiques des occlusives initiales en môn-khmer [Mon-Khmer consonant shifts: initial stops]. *Bulletin de la Société de Linguistique de Paris* 60(1). 160–72.
- Haudricourt, André-Georges. 1972. Two-way and three-way splitting of tonal systems in some Far Eastern languages (Translated by Christopher Court). In Jimmy G. Harris & Richard B. Noss (eds.), *Tai phonetics and phonology*, 58–86. Bangkok: Central Institute of English Language, Mahidol University.
- Haudricourt, André-Georges. 1973. La linguistique panchronique nécessaire à la linguistique comparée, science auxiliaire de la diachronie sociologique et ethnographique [On the necessity of panchronic linguistics for comparative linguistics, itself an auxiliary science of sociological-ethnographical (/anthropological) diachronic sociology]. *Ethnies* 3. 23–26.
- Haudricourt, André-Georges. 1991. Compte rendu d'exposé: La monosyllabisation et l'apparentement des langues du Sud-Est asiatique [Monosyllabicization and family relationships among languages of Southeast Asia; short report on talk given at the Société de Linguistique de Paris]. *Bulletin de la Société de Linguistique de Paris* 86(1). XX–XXI.
- Hayes, La Vaughn H. 1982. The mutation of \*r in pre-Thavung. *Mon-Khmer Studies* 11. 83–100.

- Hayes, La Vaughn H. 1992. Vietic and Việt-Mường: a new subgrouping in Mon-Khmer. *Mon-Khmer Studies* 21. 211–228.
- Henderson, Eugénie J.A. 1985. Feature shuffling in Southeast Asian languages. In Suriya Ratanakul, David Thomas & Premsrirat Suwilai (eds.), *Southeast Asian Linguistic Studies presented to André-G. Haudricourt*, 1–22. Bangkok: Mahidol University.
- Hoàng Thị Châu. 1989. *Tiếng Việt trên các miền đất nước: phương ngữ học [Vietnamese in the various areas of the motherland: a dialectological study]*. Hà Nội: NXB Khoa học Xã hội.
- Hoàng Thị Châu. 2004. *Phương ngữ học tiếng Việt [Vietnamese dialectology]*. Hà Nội: NXB Đại học Quốc gia Hà Nội.
- Honda, Koichi. 2008. Tone in the Lam River speech of North-Central Vietnamese: an acoustically-based multi-speaker description and analysis. Canberra: Australian National University Ph.D.
- Kirby, James. 2011. Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association* 41(3). 381–392.
- Kohler, Klaus J. 1979. Dimensions in the perception of fortis and lenis plosives. *Phonetica* 36. 332–343.
- Kohler, Klaus J. 1984. Phonetic explanation in phonology: the feature fortis/lenis. *Phonetica* 41(3). 150–174.
- Langlet, Philippe. 1990. *L'ancienne historiographie d'État au Vietnam. Raisons d'être, conditions d'élaboration et caractères au siècle des Nguyen [State historiography in Vietnam: its motivation, conditions of elaboration and characteristics in the century of the Nguyen]*. Paris: École française d'Extrême-Orient.
- Le Failler, Philippe. 2001. *Le pilori des chimères. Monopole et prohibition de l'opium en Indochine [From opium monopoly to opium prohibition in Indochina]*. Paris: L'Harmattan.
- Le Failler, Philippe. 2014. *La Rivière Noire: l'intégration d'une marche frontière au Vietnam [The Black River: integration of a frontier area into Vietnam]*. Paris: CNRS Editions.
- Mai Ngọc Chừ, Vũ Đức Nghiệu & Hoàng Trọng Phiến. 2005. *Cơ sở ngôn ngữ học và tiếng Việt [Essentials of general and Vietnamese linguistics]*. Hanoi: NXB Giáo dục.
- Martinet, André. 2005. *Economie des changements phonétiques: traité de phonologie diachronique [The economy of phonetic changes: a treaty in diachronic phonology]*. Paris: Maisonneuve & Larose.
- Martini, François. 1958. Notices bibliographiques - langue vietnamienne: Emeneau, Studies in Vietnamese (Annamese) Grammar, et Lê Văn Lý, Le parler vietnamien [book reviews]. *Bulletin de l'École Française d'Extrême-Orient* 49. 337–348.

- Maspero, Henri. 1912. Etude sur la phonétique historique de la langue annamite: les initiales [Studies in Annamese historical phonetics: initial consonants]. *Bulletin de l'Ecole Française d'Extrême-Orient* 12. 1–127.
- Maspero, Henri. 1918. Etudes d'histoire d'Annam [A study of the history of Annam]. *Bulletin de l'Ecole française d'Extrême-Orient* 18. 1–36.
- Michailovsky, Boyd, Martine Mazaudon, Alexis Michaud, Séverine Guillaume, Alexandre François & Evangelia Adamou. 2014. Documenting and researching endangered languages: the Pangloss Collection. *Language Documentation and Conservation* 8. 119–135.
- Michaud, Alexis. 2004. Final consonants and glottalization: new perspectives from Hanoi Vietnamese. *Phonetica* 61(2-3). 119–146.
- Michaud, Alexis. 2012. Monosyllabicization: patterns of evolution in Asian languages. In Nicole Nau, Thomas Stolz & Cornelia Stroh (eds.), *Monosyllables: from phonology to typology*, 115–130. Berlin: Akademie Verlag. <http://halshs.archives-ouvertes.fr/halshs-00436432>
- Nguyễn Như Ý. 1999. *Đại từ điển tiếng Việt [Comprehensive Vietnamese dictionary]*. Hanoi: NXB văn hoá-thông tin.
- Nguyễn Tài Cẩn. 1995. *Giáo trình lịch sử ngữ âm tiếng Việt [A Course in Vietnamese historical phonology]*. Hanoi: NXB Giáo dục.
- Nguyễn Tài Cẩn. 2000. Bàn thêm về hiện tượng xát hóa [Additional discussion on the phenomenon of spirantization; translated by Jim Owen]. In *Proceedings of: Pan-Asiatic Linguistics: the fifth international symposium on languages and linguistics*, vol. 1, 170–177.
- Nguyen, Van Chinh. 2013. Recent Chinese migration to Vietnam. *Asian and Pacific Migration Journal* 22(1). 7–30.
- Nguyễn Văn Lợi & Nguyễn Văn Mạnh. 2010. *Lễ hội dân gian của người Việt ở Quảng Bình [Folk festivals of the Vietnamese people in Quang Binh]*. Hanoi: NXB Văn hóa Dân tộc.
- Nguyễn Văn Tài. 2005. *Ngữ âm tiếng Mường qua các phương ngữ [The phonetics of the Mường language across its various dialects]*. Hanoi: NXB Từ điển Bách khoa.
- Pain, Frédéric, Michel Ferlus, Alexis Michaud & Thu Hà Phạm. 2014. *EFEO-CNRS-SOAS word list for linguistic fieldwork in Southeast Asia*. Hanoi: International Research Institute MICA. <https://halshs.archives-ouvertes.fr/halshs-01068533>
- Phan, John. 2012. Mường is not a subgroup: phonological evidence for a paraphyletic taxon in the Viet-Muong sub-family. *Mon-Khmer Studies* 40. 1–18.
- Phan, John. 2013. Lacquered words: the evolution of Vietnamese under Sinitic influences from the 1st century BCE to the 17th century CE. Cornell University Ph.D.
- Rhodes, Alexandre de. 1651. *Dictionarium Annamiticum Lusitanum et Latinum*. Rome.

- Shimizu, Masaaki. 2011. The order of spirantisation and voicing in the history of Vietnamese initials - evidence from Chu Nom materials. 829–843. Hanoi, Vietnam: University of Social Sciences and Humanities, VNU.
- Thomas, David & Robert Headley. 1970. More on Mon-Khmer subgroupings. *Lingua* 25. 398–418.
- Thompson, Laurence C. 1984. A Vietnamese Reference Grammar. *Mon-Khmer Studies* 13-14(1-367).
- Trần Trí Dõi. 2000. *Nghiên cứu ngôn ngữ các dân tộc thiểu số ở Việt Nam [Linguistic research about the ethnic minorities of Vietnam]*. Hanoi: NXB Đại học Quốc gia Hà Nội.
- Trần Trí Dõi. 2003. *Chính sách ngôn ngữ văn hoá dân tộc ở Việt Nam [Policies concerning ethnic languages and cultures in Vietnam]*. NXB Đại học Quốc gia Hà Nội.

### APPENDIX: WORD LIST

This list is arranged by alphabetic order of IPA forms. Tones are noted phonetically; for the correspondences with the etymological categories A1 to D2, see section 3.3.1 of the article. The glosses constitute translations, not necessarily cognates. In many cases the Phong Nha forms are cognate with the Vietnamese form used for elicitation, e.g. ‘three’ /b̩a45/ (tone: A1) corresponding exactly to Vietnamese *ba* /b̩a<sup>A1</sup>/; but no attempt was made to provide cognates systematically in the list below, e.g. by verifying whether Phong Nha /b̩ǎj45/ for ‘to jump (tiger)’ is a semantic development from Standard Vietnamese *bay* /b̩ǎj<sup>A1</sup>/ ‘to fly’.

<i>English</i>	<i>Vietnamese orthography</i>	<i>Phong Nha</i>
elder sister	<i>chị</i>	a2 <sup>?</sup> d̩o32
three	<i>ba</i>	b̩a45
to jump (tiger)	<i>cọp vồ, tát</i>	b̩ǎj45
silver	<i>bạc</i>	b̩ak53
to chop up	<i>bằm, vằm</i>	b̩ǎm32
to say	<i>nói, bảo</i>	b̩aw2 <sup>?</sup>
flea	<i>bọ chét</i>	b̩o1 c̩et24
ox	<i>bò</i>	b̩o32
zebu	<i>bò u</i>	b̩o32 u45
to wash (rice)	<i>vo (gạo)</i>	b̩o45
to milk	<i>vắt sữa</i>	b̩oɸ53 ʃuə2 <sup>?</sup> , n̩aŋ1 ʃuə2 <sup>?</sup>
to break, to interrupt	<i>bẻ gãy, làm đứt, làm vỡ</i>	b̩e2 <sup>?</sup> l̩ɔj1
seven	<i>bảy</i>	b̩eɟ2 <sup>?</sup>
we	<i>chúng tôi</i>	b̩eɟ32 tuɟ45
sea	<i>biển</i>	b̩iəŋ2 <sup>?</sup>

butterfly	<i>bươm bươm</i>	buəm45 buəm324
four	<i>bốn</i>	bon324
breast (female)	<i>vú</i>	bu324
mud	<i>bùn</i>	bun32
to run	<i>chạy</i>	căj1
casting net	<i>chài</i>	caj32
bamboo strips	<i>dây lát</i>	cak53 lat53
to thread	<i>xâu kim</i>	ɕə2 <sup>?</sup>
dishonest	<i>gian xảo</i>	ɕə2 <sup>?</sup> la324
dog	<i>chó</i>	cə324
to split bamboo into tape	<i>chẻ lát</i>	cə2 <sup>?</sup> lat53
to wedge	<i>chêm</i>	cem45
spiny amaranth	<i>rau giền gai</i>	cen45
lightning	<i>chớp</i>	cɤp24
lead	<i>chì</i>	ci32
head louse	<i>chí, chấy</i>	ci324
bird	<i>chim</i>	cim45
to sing (of birds)	<i>(chim) hát</i>	hət24
nine	<i>chín</i>	cin324
to take	<i>lấy</i>	cui1
bee	<i>ong mật</i>	ɔŋ45 mət53
weasel	<i>chồn</i>	con32
banana	<i>chuối</i>	cuəj324
broom	<i>cái chổi</i>	cuj2 <sup>?</sup>
to pierce, to bore	<i>giùi lỗ, khoan lỗ</i>	cuj45
rat	<i>chuột</i>	cwot53
stone	<i>đá</i>	ɗa324
testicles	<i>hòn dái, trứng dái</i>	ɗaj324
tough (meat)	<i>thịt dai</i>	ɗaj45
knife	<i>dao con</i>	ɗaw45 / ɗaw45
to lead by a lunge	<i>dắt đi</i>	ɗăk24
way, path	<i>đường</i>	ɗaŋ32
to give birth; to lay (eggs)	<i> đẻ ( đẻ sinh; đẻ trứng)</i>	ɗə2 <sup>?</sup>
urine	<i>nước tiểu</i>	ɗəj324
slobber, slaver	<i>nước dãi, bọt miếng</i>	ɗɤj324
to pound rice	<i>giã gạo</i>	ɗɤm45 ɣaw1
earth, land	<i>đất</i>	ɗɤt24
footprints	<i>dấu, vết chân</i>	ɗɤw324
to go	<i>đi</i>	ɗi45
water leech	<i>con đĩa, tác</i>	ɗiə2 <sup>?</sup>
beans (general)	<i>đậu, đỗ</i>	ɗo1
groundnut, peanut	<i>lạc, đậu phộng</i>	ɗo1 ɲo2 <sup>?</sup>
footprint	<i>dấu chân</i>	ɗo324

plain	<i>đồng bằng</i>	ɗoŋ31 bǎŋ31
copper	<i>đồng</i>	ɗoŋ32
bronze	<i>đồng thanh, đồng điều</i>	ɗoŋ32 ɗɛn45
ridgepole	<i>xà nóc</i>	ɗoŋ45
to light a fire	<i>đốt lửa</i>	ɗot24 la2 <sup>?</sup>
red	<i>đỏ</i>	ɗɔ2 <sup>?</sup>
carrying pole	<i>đòn gánh</i>	ɗɔn32 t̄iəŋ45
east	<i>(phía) đông</i>	ɗɔŋ45
tail	<i>đuôi</i>	ɗuəj45
to pack	<i>gói lại</i>	ɗum45
navel	<i>cái rún, rón</i>	ɗun324
to draw up, to set up	<i>dựng lên, sắp đặt</i>	ɗuŋ1
leather, hide	<i>da thuộc</i>	ɗa45
long	<i>dài</i>	ɗaj32
slope	<i>dốc</i>	ɗɔk24
easy	<i>dễ</i>	ɗe1
to demolish a house	<i>phá, dỡ nhà</i>	ɗɔ1
below	<i>bên dưới</i>	ɗɔj324 ; ɗuəj324
to put out a fire	<i>dập (lửa)</i>	ɗɔp53
mother's elder sister	<i>đì</i>	ɗi32
kite, falcon, eagle	<i>diều hâu</i>	ɗiəw32
sausage	<i>xúc xích, dôi</i>	ɗoj32
to use	<i>dùng</i>	ɗuŋ32
<i>proper name: Dung (sometimes spelt Dzung in English)</i>		ɗuŋ45
wet rice field bunds	<i>bờ ruộng</i>	ɗuəŋ32 roŋ1
fern	<i>(cây) tổ rồng, dương xỉ</i>	ɗuəŋ45
frog	<i>ếch, ngóe</i>	ɛc324
diarrhoea	<i>ỉa chảy</i>	ɛ2 <sup>?</sup> cǎj2 <sup>?</sup>
noise	<i>tiếng ồn, âm</i>	ɔm45
incubate	<i>ấp trứng</i>	ɔp24
red pepper, chilli	<i>ớt</i>	ɔt324
chicken	<i>gà</i>	ɣa32
to sing (rooster)	<i>(gà) gáy</i>	ɣǎj324
chicken	<i>gà trống/mái</i>	ɣa32 t̄oŋ4, ɣa32 maj324
cast iron, pig iron	<i>gang</i>	ɣaŋ45
husked rice	<i>gạo</i>	ɣaw1
glutinous rice	<i>gạo nếp</i>	ɣaw1 nep24
ginger	<i>gừng</i>	ɣɔŋ32
hoof	<i>móng guốc, guốc</i>	ɣuək324
sun bear	<i>gấu chó</i>	gu1
two	<i>hai</i>	haj45
deep cave	<i>hang sâu</i>	haŋ45 su45
cave	<i>hang, động</i>	haŋ45, haŋ45 ɗoŋ1



to cough	<i>ho</i>	hɛn324
wild boar, wild pig	<i>heo, lợn rừng</i>	hɛw45 ri45
pig (domestic)	<i>heo, lợn</i>	hɛw45, lɔn1
rapid	<i>ghênh, thác ghênh</i>	ho32
flower	<i>bông, hoa</i>	hwa45
to marry (of a woman)	<i>lấy chồng</i>	joŋ324
fish	<i>cá</i>	ka324
spawning fish	<i>cá đẻ</i>	ka324 đɛ2ʔ
crocodile	<i>cá sấu</i>	ka324 sɔ̃w324
to bite	<i>(chó) cắn, gặm</i>	kẵm324
paws, feet (of animals)	<i>bàn chân</i>	kẵŋ2ʔ
to hold in the jaws	<i>gặm, ngậm</i>	kap53
sand	<i>cát</i>	kat24
to harvest, gather (general)	<i>gặt hái</i>	kăt24 haj324
to harvest rice	<i>gặt lúa</i>	kăt24 lɔ324
grass	<i>cỏ</i>	ko2ʔ
toad	<i>cóc</i>	kõk24
dibble	<i>dụng cụ chọc lỗ tra hạt</i>	kɔ̃j1
male (animal)	<i>đực, trống</i>	kõn45 đɔ̃uk53
female (animal)	<i>cái, mái</i>	kõn45 kɔ̃j324
stag, deer	<i>hươu</i>	kõn45 nɔ̃j45
wing	<i>cánh</i>	kẽp324
to cry (animals)	<i>kêu (động vật)</i>	kew45
ant	<i>kiến đen</i>	kiɛn324 đɛn45
excrements, faeces	<i>cứt</i>	kit24
walking staff	<i>(gậy) ba toong</i>	kɔ̃j1
ladder	<i>cái thang</i>	kɔ̃j324
cooked rice	<i>cơm</i>	kɔ̃m45
tea plant	<i>cây chè</i>	kɔ̃n45 ɕɛ32
banana (plant)	<i>cây chuối</i>	kɔ̃n45 ɕuəj324
areca palm tree	<i>cây cau</i>	kɔ̃n45 kăw45
tree	<i>cây cối</i>	kɔ̃n45 koj324
rice seedlings	<i>cây mạ</i>	kɔ̃n45 ma324
lemongrass	<i>sả chanh</i>	kɔ̃n45 ɕa45 ɕa2ʔ
coconut palm	<i>cây dừa</i>	kɔ̃n45 ɕuə32
nettle-hemp	<i>cây gai</i>	kɔ̃n45 ɕuə45
to plant out rice seedlings	<i>cấy lúa</i>	kɔ̃j324 lɔ324
thorn	<i>gai</i>	kɔ̃j45
girl	<i>gái</i>	kɔ̃j324
plough	<i>cái cày</i>	kɔ̃n45 kăj32
rainbow	<i>cầu vồng</i>	kɔ̃w31 vɔ̃ŋ31
tuber	<i>củ</i>	ko2ʔ
mortar (for pounding rice)	<i>cối giã gạo</i>	koj324 đɔ̃m45 yaw1

owl (general)	<i>cú mèo</i>	ku45
crab	<i>cua</i>	kuə45
to carry in a basket	<i>gùi</i>	kuj32 ; koŋ32
knee	<i>đầu gối</i>	kuj324
spur (of cock)	<i>cựa (gà trống)</i>	kuə2 <sup>?</sup>
crow	<i>quạ</i>	kwa1
to scratch (with one's nail)	<i>gãi</i>	k <sup>h</sup> aj2 <sup>?</sup>
tiger	<i>khái</i>	k <sup>h</sup> aj324 ~ xaj324
pass, defile	<i>đèo; hẻm núi, khe núi</i>	k <sup>h</sup> ε45
stream	<i>suối</i>	k <sup>h</sup> ε45
monkey (general)	<i>khỉ</i>	k <sup>h</sup> i2 <sup>?</sup>
smoke	<i>khói</i>	k <sup>h</sup> ɔj324
to whittle, to cut out	<i>gọt, chạm gọt</i>	k <sup>h</sup> ət24
sweet potato	<i>khoai lang</i>	k <sup>h</sup> waj45 laŋ45
liar	<i>người nói dối</i>	k <sup>h</sup> wet24
fire	<i>lửa</i>	la2 <sup>?</sup>
leaf	<i>lá</i>	la324
gum	<i>lợi, nôi</i>	laj1
tongue	<i>lưỡi</i>	laj1
blade	<i>lưỡi dao</i>	laj1 đaw45
net	<i>lưới</i>	laj324
village	<i>làng</i>	laŋ32
to set (the sun sets)	<i>lặn (mặt trời)</i>	lăn53
fly, bluebottle	<i>nhặng</i>	lăŋ32
embers	<i>mẫu củi cháy dở</i>	lɤ1
to flower, bloom, blossom	<i>ra hoa, nở (hoa), trổ bông</i>	lo2 <sup>?</sup> hwa45
feather	<i>lông</i>	loŋ45
to plant a tree	<i>trồng cây</i>	loŋ45 kɤn45
paddy	<i>lúa</i>	lɔ324
eel	<i>lươn</i>	luən45
rice seedlings	<i>(cây) mạ</i>	ma324
shell (of tortoise)	<i>mai rùa</i>	maj45 rɔ32
to borrow (object)	<i>mượn (vật)</i>	man1
moon	<i>mặt trăng</i>	măt53 t̚aŋ45
sun	<i>mặt trời</i>	măt53 t̚ɤj32
sesamum	<i>mè, vừng</i>	mε32
partition, bamboo panel	<i>phên</i>	men45
cat	<i>mèo</i>	mew32
cloud	<i>mây</i>	mɤj45
to cultivate a rice field	<i>làm ruộng</i>	mɤn32 rɔŋ1
honey	<i>mật</i>	mɤt53
tomb	<i>mồ, mộ</i>	mo32 ; ma2 <sup>?</sup>
where?	<i>(đi) đâu?</i>	mo45

termite	<i>con mối</i>	moj24
mouth	<i>miệng</i>	mom1
taro	<i>khoai sọ, môn</i>	mon45
cockscorn, crest	<i>mào (gà)</i>	moŋ32 ɣa32
one	<i>một</i>	mot53
beak	<i>mỏ</i>	mo2 <sup>?</sup>
mosquito	<i>muỗi</i>	moj1
salt	<i>muối</i>	moj324
to rise (the sun rises)	<i>mọc (mặt trời)</i>	mo:k53
weevil	<i>mọt</i>	mot53
lip	<i>môi</i>	muj45
cinders, ashes	<i>tro</i>	mun45
rain	<i>mưa</i>	muə45
water	<i>nước</i>	nak24
year	<i>năm</i>	năm45
five	<i>năm</i>	năm45
this year	<i>năm nay</i>	năm45 ni45
stopples, cork, plug	<i>cái nút</i>	năp24
branch	<i>cành, nhánh</i>	neŋ324
anchor	<i>mỏ neo</i>	new45
boat, ship	<i>thuyền, tàu</i>	nok24
to swallow	<i>nuốt</i>	not24
to give birth	<i>sinh đẻ</i>	no2 <sup>?</sup>
to cook rice (water)	<i>nấu cơm</i>	no:w324 ko:m45
orchard	<i>vườn hoa quả</i>	nuəŋ45
dark sky	<i>(trời) râm</i>	no:m45 ~ no:m45
to weed	<i>nhổ, nhổ cỏ</i>	no2 <sup>?</sup> ko2 <sup>?</sup>
small	<i>nhỏ, bé</i>	no2 <sup>?</sup>
tusk, ivory	<i>ngà</i>	ŋa32
man (homo)	<i>người ta, con người</i>	ŋaj32
far, distant (from)	<i>xa (nơi)</i>	ŋaj45
day	<i>ngày</i>	ŋăj45
calf	<i>(con) bê, nghé</i>	ŋe324
young buffalo	<i>nghé</i>	ŋe324
to be pregnant	<i>có bầu, có thai</i>	ŋen324
near to	<i>gần (nơi)</i>	ŋin45
horse	<i>ngựa</i>	ŋuə1
star	<i>ngôi sao</i>	ŋoj45 sa:w45
in front	<i>đàng, đằng trước</i>	ŋo1 tuək24
to look at	<i>nhìn</i>	ŋo324
nest	<i>ổ, tổ chim</i>	o2 <sup>?</sup>
father's elder sister	<i>cô, bác, bá</i>	o45
billhook	<i>rựa nhỏ</i>	ra1

cattle pen, cow pen	<i>chuông trâu</i>	ran32
river	<i>sông</i>	raw32
fence	<i>rào, giậu</i>	raw32
tooth	<i>răng</i>	răŋ45
placenta (after birth)	<i>nhau</i>	răw45
bedbug	<i>rệp</i>	rep24
root	<i>rễ cây</i>	ren1 kɤn45
axe	<i>rìu</i>	rɛw32
dragon	<i>rồng</i>	ron32
tortoise	<i>rùa</i>	rɔ32
fly	<i>ruồi</i>	rɔj32
gnat, midge	<i>muỗi mắt, muỗi nhỏ</i>	rɔj32 kɔn45
whip, lash	<i>cái roi</i>	rɔj45
wet rice field (in the plains)	<i>ruộng nước</i>	rɔŋ1
dry paddy field	<i>ruộng lúa khô, ruộng cạn</i>	rɔŋ1 k <sup>h</sup> o45
intestines	<i>ruột</i>	rɔt53
fishnet	<i>rớ, vó</i>	rɤ324
dry rice field	<i>rẫy</i>	rɤj1
sieve	<i>rây</i>	rɤj45
mountain, hill	<i>núi, đồi</i>	ru324
forest	<i>rừng</i>	ru324
to husk rice	<i>xay lúa</i>	săj45 lɔ324
to wash (clothes)	<i>giặt quần áo</i>	săt24
rack (over the fire)	<i>giàn trên bếp</i>	şan32
wax	<i>sáp</i>	şap24
to winnow paddy	<i>sây lúa</i>	şăj2 <sup>7</sup> lɔ324
manioc, tapioca, cassava	<i>sắn</i>	şăn324
to hunt	<i>săn mồi</i>	şăn324
iron	<i>sắt</i>	şăt24
face to face	<i>đổi mặt</i>	şăw24
six	<i>sáu</i>	şăw324
land leech	<i>con vắt</i>	şen45
lotus	<i>(hoa) sen</i>	şen45
to hug	<i>ôm</i>	şijn324
adhesive	<i>dính</i>	şijn324
wolf	<i>chó sói, sói đỏ</i>	şɔj324
squirrel	<i>sóc</i>	şɔk24
pot (for cooking rice)	<i>nồi nấu cơm</i>	şɔŋ45
thunder	<i>sấm</i>	şɤm324
milk	<i>sữa</i>	şuə1
to carry on a pole	<i>gánh</i>	şuəŋ45
deep	<i>sâu</i>	şu32
ugly	<i>xấu</i>	şu45

to bark	<i>chó sủa</i>	ʃuə2ʔ
hand-held pestle	<i>chày tay</i>	tǎj32 dʒm45 ɣaw1
snake	<i>rắn</i>	tǎn324
to put out a fire	<i>tắt lửa</i>	tǎt24 la2ʔ
plantain leaf	<i>tàu lá chuối</i>	tǎw32 la324 cwej324
rhinoceros	<i>tê giác</i>	te45 jak24
crayfish	<i>tôm sông, tôm nước ngọt</i>	teɸ324
garlic	<i>tỏi</i>	təj2ʔ
to fight	<i>đánh trả, chống trả</i>	tǎ2ʔ
son	<i>con trai</i>	taj45
python	<i>trăn</i>	tǎn45
white	<i>trắng</i>	tǎŋ324
above	<i>bên trên</i>	tɛn45
to weave straw wall	<i>đan/bện tấm tranh</i>	tɛɸ45
centipede	<i>con rết, rít</i>	tit24
prawn	<i>tôm</i>	tom45
cushion	<i>gối dựa, gối đệm</i>	tuə1
sky	<i>trời</i>	tʃj32
otter	<i>con rái cá</i>	tʃj324
fruit	<i>trái, quả</i>	tʃj324
ear lobe	<i>dái tai</i>	tʃj324
areca nut (betel nut)	<i>trái cau</i>	tʃj324 kǎw45
coconut	<i>trái (quả) dừa</i>	tʃj324 zuə32
body louse	<i>rận</i>	tʃn324
to slip	<i>trượt, tuột</i>	tʃt53
buffalo	<i>trâu</i>	tʃw45
insect	<i>sâu bọ</i>	tʃw45 bə1
hare	<i>thỏ rừng</i>	tʰə2ʔ ru324
trunk (of trees)	<i>thân cây</i>	tʰɛn45 kɛn45
tin	<i>thiếc, kẽm</i>	tʰiək24
leak (from the roof)	<i>dột (dột mái, mái dột)</i>	tʰot24
head	<i>đầu, tróc</i>	tɔk24
back of the blade	<i>cái sống dao</i>	tɔŋ324
drum	<i>trống</i>	tɔŋ324
betel	<i>trầu</i>	tɔ32
mole	<i>chuột chũi</i>	tɔ32
earthworm	<i>con trùn, giun</i>	tun32
pangolin	<i>con tê tê, trút</i>	tut24
to keep (buffalo)	<i>giữ trâu</i>	tɔw1 tʃw45
middle, between	<i>ở giữa</i>	tɔə2ʔ
father's elder sister's husband	<i>đượng</i>	tɔəŋ1
horns (of buffalo)	<i>sừng</i>	tɔŋ32
antlers	<i>gạc nai</i>	tɔŋ32 nɛj45

egg	<i>trúng</i>	tʉŋ324
to broadcast rice	<i>gieo vãi lúa</i>	vaj1 lɔ324
scale (of fish...)	<i>vảy (cá, con tê tê...)</i>	văj2ʔ
gold	<i>vàng</i>	vaŋ32
lime	<i>vôi</i>	voj45
(wood) shavings	<i>vỏ bào</i>	vɔ1
elephant	<i>voi</i>	vɔj45
elephant's trunk	<i>vòi voi</i>	vɔj32 vɔj45
to thin out bamboo tape	<i>vót lát</i>	vɔt24 lat53
gibbon	<i>vượn</i>	vʉəŋ1
marsh	<i>đầm lầy</i>	vʉŋ32 lun45
yes	<i>dạ</i>	ʒa1
false, untrue	<i>giả</i>	ʒa2ʔ
house	<i>nhà</i>	ʒa32
weeds	<i>cỏ dại</i>	ʒaj1
mad dog	<i>chó dại</i>	ʒaj1
to chew	<i>nhai</i>	ʒaj45
to dare (to do sth.)	<i>dám</i>	ʒam324
to imprison	<i>giam, bỏ tù ai</i>	ʒam45
to paste, glue	<i>dán lại</i>	ʒan324
to disperse	<i>xích, dang ra</i>	ʒaŋ45
to stroll	<i>đi dạo, đi chơi</i>	ʒaw45
to teach	<i>dạy</i>	ʒăj1
mountains	<i>dãy núi</i>	ʒăj2ʔ
to jump, to leap (e.g. frog)	<i>nhảy</i>	ʒăj2ʔ ~ năj2ʔ
thick	<i>dày</i>	ʒăj32
cricket	<i>con đé</i>	ʒe324
goat	<i>đê</i>	ʒe45
to hoard	<i>để dành, dành dụm, tiết kiệm</i>	ʒeŋ32
to weave	<i>dệt, đan</i>	ʒet53
tear grass, job's tears	<i>bo bo, ý dĩ, cườm thảo</i>	ʒi1
divorced (husband)	<i>bỏ chồng, li dị</i>	ʒi1
sulfur	<i>diêm sinh</i>	ʒiəŋ45
to speak, make a speech	<i>diễn thuyết</i>	ʒiəŋ1
hedgehog	<i>nhím</i>	ʒim324
bad	<i>xấu, dở</i>	ʒɤ2ʔ
cucumber	<i>dưa chuột, dưa leo</i>	ʒɤ45 lɛw45
to play	<i>chơi</i>	ʒɤj1
to get up	<i>đứng dậy, dậy</i>	ʒɤj1
to struggle	<i>vùng vẫy, dẫy dụa</i>	ʒɤj1
bat	<i>dơi</i>	ʒɤj45
to lead	<i>dẫn ai</i>	ʒɤŋ1
villager	<i>dân làng</i>	ʒɤŋ45

to make a sacrifice	<i>cúng, hiến dâng</i>	ʒɤŋ45
to rise	<i>dâng lên</i>	ʒɤŋ45
castor oil tree	<i>cây thầu dầu</i>	ʒɤw32
daughter in law, son's wife	<i>con dâu</i>	ʒɤw45 ; ɔu45
to observe	<i>dò xét, quan sát</i>	ʒo32
wind	<i>gió</i>	ʒo324
rose-apple	<i>roi, mận</i>	ʒoɟ45
thin	<i>ốm gầy</i>	ʒom324
somnambulism	<i>mộng du</i>	ʒu45
Rhizomys sinensis	<i>dũi</i>	ʒuj1
tinder	<i>bùi nhùi (để mồi lửa)</i>	ʒuj32
evil	<i>dữ tợn</i>	ʒu1
fruit shell	<i>vỏ cứng, gáo dừa</i>	ʒuə32
antelope goral	<i>ban linh, linh dương</i>	ʒuəŋ45

---