NOTES ON GRAMMATICAL VOCABULARY IN CENTRAL VIETNAMESE

Mark J. ALVES
Montgomery College, USA
<Mark.Alves@montgomerycollege.edu>

Abstract
This study focuses on grammatical vocabulary in Central Vietnamese\(^2\) (CV, hereafter) with the goal of exploring historical-linguistic developments in Vietnamese in general. One purpose of focusing on grammatical vocabulary is to provide a sample of how this type of vocabulary can innovate over time in different dialectal regions. For research on Central Vietnamese in particular, this focus also allows some identification of historical patterns of lexical preservations (e.g., forms shared with other Vietic languages), innovations or general variation, and the effects of language contact. While CV is quite obviously Vietnamese, there are a number of distinctive traits in its grammatical vocabulary which highlight some of the differentiation over time between CV and standard Vietnamese\(^3\) (StV, hereafter). In a number of instances, data suggests reconstructability at the level of Proto-Việt-Mường or even Proto-Vietic forms. These historical points are dealt with throughout this paper and summarized in the conclusion.

Keywords: Vietnamese, Proto-Vietic, grammatical vocabulary

ISO 639-3 language codes: vie, mtq, seb

1. Central Vietnamese as a dialect
The historical-linguistic position taken in this paper follows that of Nguyễn Tài Cẩn (1979). Vietnamese belongs to the Vietic branch of Mon-Khmer, and within Vietic, there are the major sub-branches Pồng-Chứt (examples of languages include Rục, Thà Vựng, and Pồng) and Việt-Mường. Within Vietnamese, there are geographic regions where dialects can be distinguished. There are differences among them in pronunciation and lexical items which can be marked by isoglosses on a map of Vietnam.

Based on differences in both the lexicon and phonology, which can significantly impair the ability of Vietnamese from other regions from understanding CV speakers, CV must be considered a dialect, not merely a regional accent. This three-way regional distinction is well-known in Vietnam and has been documented in major linguistic works, such as that of Hoàng Thị Châu (1989: 90-103), as well as numerous shorter works (e.g., Friberg 1973, Vu 1982, etc.). Hoàng (Ibid.) even provided a list of lexical differences among several key grammatical types in the three regions of CV, Northern Vietnamese (NV, hereafter), and Southern Vietnamese (SoV, hereafter). Three areas—Bắc kỳ, Trung kỳ, and Nam Kỳ—and their lexical and phonological distinctions were recognized even in the cultural work of Phan K.B. at the beginning of the twentieth century (2011:4227435). Rather than discussing a distinction between NV, CV, and SoV, Thompson (1965:80-85) describes differences among Vietnamese dialects with major cities as the approximate dialect centers. In this view, Vĩnh and Huế, which are typically both considered to be CV dialects, are treated separately due to phonological distinctions.

1 This paper is a modified version of the article published in 2011 in Hội Thảo Quốc Tế Đào Tạo và Nghiên Cứu Ngôn Ngữ Học ở Việt Nam: Những Vấn ĐỀ Lưới và Th ực Thiện (International Conference on Linguistics Training and Research in Vietnam: Theoretical and Practical Issues); Trường Đại Học Khoa Học Xã Hội và Nhân Văn (Hanoi. The University of Social Sciences and Humanities). ed. by Nguyễn Hồng Cồn, et. al.

2 The following is a complete list of the abbreviations: CV=Central Vietnamese, NV=Northern Vietnamese, NCV=North-Central Vietnamese, SV=Southern Vietnamese, StV=Standard Vietnamese.

3 The term ‘standard Vietnamese’ is a translation of the Vietnamese term tiếng chuẩn, literally ‘standard speech,’ which refers to the generally accepted standard throughout Vietnam. It is generally similar to Northern Vietnamese. However, in some cases, even Northern Vietnamese has phonological and lexical traits which differ from the accepted standard.
While Thompson’s approach circumvents deciding on major dialect regions, it also recognizes the fact that the matter is more complex, and that considering a three-way division may not be sufficient. It has also been noted that North-Central Vietnamese (NCV, hereafter), in the region around the city of Vinh, can possibly be considered a sub-region of CV due to some additional lexical and phonological distinctions. Hoàng (Ibid. 261) notes the special grammatical vocabulary in the NCV region and provides a list of key lexical items. Vietnamese spoken in the region of Nghệ Tĩnh and Bình Trị Thiên have been posited by Nguyễn Tài Cẩn in his 1995 historical-linguistic textbook on Vietnamese as being part of a special region of Vietnamese, one with characteristics of both NV and Mường (1995:275-278). The substantial dialectal variety even within small distances in that region, such those described in Alves 2007, suggests that this area is perhaps among the oldest of Vietnamese. It is also worth noting that NCV is a region where Mường and Việt-Chứt languages were spoken. However, more investigation would be needed to strengthen such a position.

2. Method of data collection and analysis
For this study, Đặng Thanh Hòa’s entire 2005 Vietnamese dialect dictionary Từ Điển Phương Ngữ Tiếng Việt was studied, with grammatical vocabulary identified. The words and phrases were then entered into a spreadsheet with various categories noted, such as region, grammatical function, historical origins when identifiable, and other details. After that, other references were consulted, including Vương Lộc’s 2002 dictionary of archaic Vietnamese, Nguyễn Tài Cẩn’s 1995 comparative lists of Việt-Chترنت language, Nguyễn Văn Lợi’s 1995 description of the Rục language, Khang et.al.’s 2002 dictionary of the Mường language, and an unpublished manuscript by Michel Ferlus of reconstructions of Proto-Vietic, including complete lists of available vocabulary.

It is important to understand the method of selecting the vocabulary for this study. Determining which words and expressions are grammatical is not an absolute matter, and one may posit that there are degrees of grammaticality. In general, grammatical vocabulary tends to have more abstract meanings which function more to indicate syntactic relationships and even pragmatic functions, whereas content words, which are not included in this study, are the elements for which grammatical terms provide this semantico-syntactic context. For instance, the possessive marker của is clearly grammatical; hôm nay ‘today’ is slightly more semantically concrete but still expresses the grammatical function of time; while sách is a clearly content word.

Thus, words which express abstract relationships between nouns, such as locatives, conjunctions, and comparatives, are considered grammatical. Those expressing modality, negation, time, and degree to modify verbs and adjectivals are also treated as grammatical in this study. Quantity expressions and measure words, while not always as abstract as numbers and classifiers, are included in this study due to their generalized, open function and somewhat similar syntactic function to other clearly grammatical words. Pronouns and deictics are non-controversially considered grammatical, but also in this study, so are terms of address, which in Vietnamese have pronominal functions. The following non-exhaustive list is representative of the criteria used to select the vocabulary for this study.

- Pronominal reference (pronouns, deictics, and terms of address)
- Location (prepositions and locative nouns)
- Time and frequency
- Interrogative function
- Negation and modality
- Quantity and units of measurement
- Clause-connecting elements
- Comparison, degree, and intensification
- Pragmatic functions (sentence final particles, exclamations)
- Causation/causative verbs

In the database, over 140 items in CV were identified, of which over 100 are unique to CV and not generally shared with either SoV or NV. Of the 140, 35 are also seen in SoV, while only two are shared with NV. Overall, it can be said that there is a noticeably large number of CV grammatical vocabulary distinct from both NV or SoV, including at least a few crucial items. The significant overlap with SoV grammatical vocabulary, in addition to a number of phonological similarities, suggests a closer historical-linguistic relationship between those two regions than with NV despite the reduced mutual intelligibility between CV and either SoV and NV, a matter anecdotally attested among Vietnamese. More importantly, it must be assumed that the original approximate areas of where modern NV and CV are spoken had diverged a good deal before the migration of Vietnamese to Southern Vietnam. That migration occurred largely over the past
several centuries since the fall of the Champa Empire in the 1400s and resulted in expansion of Vietnamese settlements to modern day Southern Vietnam.

Based on the data collected, the types of CV grammatical vocabulary fall into the categories shown in Table 1. The first two categories are briefly explained in this section and then referred to in subsequent sections, while the third category, unique etyma, constitutes the bulk of the rest of the paper. A brief note on archaic Vietnamese preserved in CV is provided in a subsection below.

Table 1: Categories of CV Grammatical Vocabulary Relative to StV

<table>
<thead>
<tr>
<th>Category</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Compounds with one cognate in common with StV and one non-cognate morph</td>
<td>CV: <code>chữ</code> (word) vs. StV: <code>gì</code> (what)</td>
</tr>
<tr>
<td>2. Cognates in common with StV but with alternate pronunciations</td>
<td>CV: <code>không</code> (not) vs. StV: <code>không</code> (not)</td>
</tr>
<tr>
<td>3. Completely unique etyma (both archaic Vietnamese words and strictly CV forms)</td>
<td>CV: <code>mi</code> (you) vs. StV: <code>mày</code> (you)</td>
</tr>
</tbody>
</table>

Instances of mixed cognate morphemes show recurring patterns primarily among adverbial expressions. Among intensifying expressions with `hết` plus another morpheme, the second morpheme may differ in CV and StV, such as CV `hết cỡ` ‘extremely’ versus the synonymous StV item `hết mức`. In several compounds expressing time, CV uses `khi` in contrast with StV `lúc`, such as CV `khi đêm` ‘at night’ versus the StV equivalent `lúc đêm`. Among terms expressing time or quantity, several compounds in both CV and StV contain `một`, while the second morpheme alternates, such as CV `một tẹo` versus StV `một tí`. It is impossible to determine whether these are calques or simple alternates, but regardless, these word formation patterns are clearly part of the linguistic region.

A. Patterns of phonological correspondence

As historical phonological patterns between CV and StV are reasonably well described (Nguyễn Đình Hoa 1965, Friberg 1973, Vũ 1982, Hoàng 1989, Nguyễn Tài Cẩn 1995, etc.), it is generally easy to identify cognates among the dialects. Some of these alternations appear among high-frequency or significant CV grammatical words, as listed in Table 2, containing comparative samples of the phonemes shown with Vietnamese orthography. Cognates in the Vietic language Rục are included and highlight the conservative phonological aspects in CV. Some of the changes represent synchronic dialectal differences as well as diachronic changes, though that has been described elsewhere (c.f., Nguyễn Tài Cẩn 1995; Ferlus 1976 and 1981) and is beyond the scope of this paper.

Table 2: Common sound correspondences between CV and StV in grammatical vocabulary

<table>
<thead>
<tr>
<th>CORRESPONDENCES</th>
<th>CV</th>
<th>StV</th>
<th>Vietic</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ươ</code> vs. <code>a</code></td>
<td>đương</td>
<td><code>dang</code></td>
<td>NA</td>
<td>progressive marker</td>
</tr>
<tr>
<td><code>ây</code> vs. <code>ai</code></td>
<td><code>cây</code></td>
<td>cái</td>
<td>Rộc <code>ke</code></td>
<td>general classifier</td>
</tr>
<tr>
<td><code>–ng</code> vs. <code>–nh</code></td>
<td><code>eng</code></td>
<td>anh</td>
<td>NA</td>
<td>term of address (male, older, respectful)</td>
</tr>
<tr>
<td><code>i</code> vs. <code>ay</code></td>
<td><code>mi</code></td>
<td><code>mày</code></td>
<td>Rộc <code>mi</code></td>
<td>you (informal)</td>
</tr>
<tr>
<td><code>le</code> vs. <code>gi</code></td>
<td><code>le</code></td>
<td><code>gi</code></td>
<td>NA</td>
<td>what</td>
</tr>
<tr>
<td><code>tr</code> vs. <code>gi</code></td>
<td><code>trưa</code></td>
<td><code>giữa</code></td>
<td>Rộc <code>troah</code></td>
<td>between</td>
</tr>
<tr>
<td><code>ui</code> vs. <code>ôi</code></td>
<td><code>tui</code></td>
<td><code>tôi</code></td>
<td>NA</td>
<td>I</td>
</tr>
<tr>
<td><code>r</code> vs. <code>nh</code></td>
<td><code>rau</code></td>
<td><code>nhau</code></td>
<td>NA</td>
<td>reciprocal marker</td>
</tr>
</tbody>
</table>

Of the 140 plus items in the database, only about 20 entries fall into this category, with about ten additional forms which cannot be considered with high certainty to be phonologically related cognates. Still, considering the high grammatical functionality of some of these words, those phonological differences, combined with the many other distinct etyma, certainly contribute to decreased mutual comprehensibility between CV and speakers in other regions. CV `mi` ‘you (informal)’ for StV `mày` and CV `mi` ‘this’ for StV `này` are quite similar to Mường and Việt-Chữ languages, such as Rộc or Thà Vùng (Suwilai 1998). A few other highly functional words which are cognates with different pronunciations include SoV `chi` versus StV `gi` ‘what,’ CV `trưa` versus StV `giữa` ‘between/the middle,’ CV `tui` versus StV `tôi` ‘I,’ and the reciprocal in CV `rau` versus StV `nhau`. Many of these represent competing sound changes.
B. Archaic words

Throughout this paper, CV words which are related to Mường or Pọng-Chứt languages are noted, but in addition to these kinds of lexical preservations, there are various Vietnamese words which are not part of modern StV but which are listed in Vương Lộc’s 2002 dictionary of archaic Vietnamese words. These words and phrases are shown in Table 3. How frequently these are still used in parts of Central Vietnam and where they are used in that region may vary and would require additional research to clarify. Nevertheless, this overall list highlights some of the more archaic parts of the CV lexicon (archaic in that they are no longer used, though few have cognates with other Vietic languages), even in the restricted domain of grammatical vocabulary. This is not to say that NV and SoV do not have vocabulary which is preserved from early stages of Vietnamese, but in addition to the other older Vietic vocabulary, this data is worth considering in exploring the history of Vietnamese.

Table 3: Archaic Grammatical Vocabulary in CV No Longer Used in StV

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>bao lăm</td>
<td>bao nhiêu</td>
<td>how much (implied negative)</td>
</tr>
<tr>
<td>bâu</td>
<td>(từ nam giới dùng để gọi vợ)</td>
<td>term used by husband to refer to wife lovingly</td>
</tr>
<tr>
<td>in</td>
<td>y</td>
<td>exactly as</td>
</tr>
<tr>
<td>lung</td>
<td>nhiêu, quá mức</td>
<td>a lot</td>
</tr>
<tr>
<td>mụ bà</td>
<td>bà</td>
<td>grandmother</td>
</tr>
<tr>
<td>nằng</td>
<td>hay, thường</td>
<td>often</td>
</tr>
<tr>
<td>nghi</td>
<td>nó, hánn</td>
<td>he/she</td>
</tr>
<tr>
<td>nhủ</td>
<td>bào</td>
<td>make/command</td>
</tr>
<tr>
<td>phải chi</td>
<td>gia nhứ, gia mà</td>
<td>supposing that</td>
</tr>
<tr>
<td>phương chi</td>
<td>huống gì</td>
<td>much less/not to say</td>
</tr>
<tr>
<td>rày/rầy</td>
<td>nay</td>
<td>at this time</td>
</tr>
<tr>
<td>rơi</td>
<td>xong</td>
<td>finished</td>
</tr>
</tbody>
</table>

3. Types of Central Vietnamese grammatical vocabulary

As noted in the previous section, the words were selected based on their grammatical functions. This casts a wide net but ultimately provides a more complete picture of the range of grammaticalized vocabulary in CV. The rest of this section covers nine types of grammatical vocabulary.

A. Interrogative words

Table 2 above shows phonological distinctions of CV, with cognates in the Vietic language Rục, in grammatical vocabulary. Beyond phonological distinctions, CV is well known for having several distinctive interrogative words, separating it from both StV and SoV. These include CV mô ‘where’ versus StV đâu, CV rạng ‘how’ versus StV sao, CV chi ‘what’ versus StV gì. The latter form chi is pronounced essentially the same as in Mường, an apparent older Vietic pronunciation, at least to proto-Việt-Mường. CV mô is notable as having cognates in several other P慵-Chứt languages. While not an interrogative word (though it is a causative verb having a relatively grammatical function), CV mần ‘do’ versus StV làm is combined with CV rạng to express ‘how’. Mần is also a Việt-Chứt cognate seen in Rucción, suggesting a possible Proto-Vietic form.

Table 4: Interrogative words in various Vietic languages

<table>
<thead>
<tr>
<th>StV</th>
<th>CV</th>
<th>Rucción</th>
<th>Mường</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>đâu</td>
<td>mô</td>
<td>tumô</td>
<td>no</td>
<td>where (or which)</td>
</tr>
<tr>
<td>gì</td>
<td>chi</td>
<td>chamoŋ</td>
<td>chi</td>
<td>what</td>
</tr>
<tr>
<td>làm</td>
<td>màn</td>
<td>mun²</td>
<td>là</td>
<td>make/do (used in interrogative compounds)</td>
</tr>
</tbody>
</table>

This sort of distinction between CV and StV is among the most persuasive data, showing that CV and StV, or NV to be precise, diverged as dialects. Moreover, in some cases, such differences also illustrate lexical and phonological preservations that highlight the closer relationship between CV and other Vietic

---

4 It is also worth noting that, in this study, several CV grammatical words with deictic reference were found to have initial /r/. These include rạng ‘how’, ri/riə ‘thusly’, and rày/rây ‘at this time’. This is not unlike patterns in basic demonstratives noted in StV by Thompson in which such words systematically have initial /d/, /n/, and /b/ (1965:142).
languages. One might even hypothesize that NV diverged from earlier Vietic, as more closely represented by CV.

B. Negation and Modals

CV has a number of distinct negation and words expressing modality. In some cases, the difference is merely in pronunciation, as in items 3 and 4 in Table 5. Of the negation words, CV nô ‘not,’ which is more of a NCV form than CV in general, is an etymon worthy of attention. Whether it is a possible CV innovation is not clear. Cognates of the item have not been found in other Vietic languages. Other items on the list are interesting, though it is not clear based on the dictionary how common the other items are in CV.

Table 5: Negation words and modals

<table>
<thead>
<tr>
<th></th>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>nô</td>
<td>không, chẳng</td>
<td>no, not</td>
</tr>
<tr>
<td>2.</td>
<td>dọn</td>
<td>khỏi</td>
<td>no need</td>
</tr>
<tr>
<td>3.</td>
<td>làm chi</td>
<td>làm</td>
<td>nothing, not</td>
</tr>
<tr>
<td>4.</td>
<td>dương</td>
<td>dang</td>
<td>(progressive marker)</td>
</tr>
<tr>
<td>5.</td>
<td>rỗi</td>
<td>xong</td>
<td>finished</td>
</tr>
<tr>
<td>6.</td>
<td>thứ</td>
<td>nói, hâm ý phủ định</td>
<td>able</td>
</tr>
<tr>
<td>7.</td>
<td>lừa chi</td>
<td>thà ràng</td>
<td>had better</td>
</tr>
</tbody>
</table>

C. Pronouns and Terms of Address

The number of differences, both phonological and etymological, among pronouns in CV clearly highlights the distinct dialect status of CV. Items 1 through 4 are cognates with StV forms, though they differ phonologically (c.f., section 3A). The pronoun pluralizer bầy may be related to StV mijn. A likely cognate is seen in Ruc, the pluralizer bợl. Items 9 through 13 are all entirely distinct from StV. At this point, they are of uncertain etymological origin, without immediately recognizable cognates among Mường or Việt-Chữ languages.

Table 6: Pronouns

<table>
<thead>
<tr>
<th></th>
<th>CV</th>
<th>StV</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>mi</td>
<td>Mỹ</td>
<td>you (informal)</td>
</tr>
<tr>
<td>2.</td>
<td>ni</td>
<td>này</td>
<td>this</td>
</tr>
<tr>
<td>3.</td>
<td>tau</td>
<td>tao</td>
<td>first-person singular, informal</td>
</tr>
<tr>
<td>4.</td>
<td>tui/bài</td>
<td>tôi</td>
<td>I</td>
</tr>
<tr>
<td>5.</td>
<td>bay/bây</td>
<td>chúng Mỹ</td>
<td>you, plural</td>
</tr>
<tr>
<td>6.</td>
<td>bài</td>
<td>bổn/tui</td>
<td>plural marker</td>
</tr>
<tr>
<td>7.</td>
<td>bài</td>
<td>bổn/tui tôi/tao</td>
<td>we</td>
</tr>
<tr>
<td>8.</td>
<td>bài tui</td>
<td>bổn/tui tôi</td>
<td>we</td>
</tr>
<tr>
<td>9.</td>
<td>nghi</td>
<td>nó, hân</td>
<td>he/she</td>
</tr>
<tr>
<td>10.</td>
<td>choa</td>
<td>chúng ta</td>
<td>we, I</td>
</tr>
<tr>
<td>11.</td>
<td>ri</td>
<td>thế này</td>
<td>thusly</td>
</tr>
<tr>
<td>12.</td>
<td>rứa</td>
<td>thế</td>
<td>thusly</td>
</tr>
<tr>
<td>13.</td>
<td>nô</td>
<td>ấy</td>
<td>that</td>
</tr>
</tbody>
</table>

As for terms of address, often derived from familial terms, again, CV shows a number of significant differences. Of note are the various forms which are not of Chinese origin, unlike in StV. CV â ‘elder sister’ shows a contrast with the likely Old-Sino-Vietnamese chị (standard Sino-Vietnamese 姐 (Mandarin jiē), mẹ and mụ ‘grandmother’ for the Chinese loanword bà 婆 (Mandarin pó), and CV o ‘sister of father’ in contrast with StV cô, also of Chinese origin 姑 (Mandarin gū). Are these preservations of older Vietic, innovations in CV, or something else? The latter item CV o in particular has a cognate in Ruc, suggesting

Admittedly, Chinese has phonologically similar bà ‘mother,’ standard Sino-Vietnamese mâu. However, having other forms meaning ‘mother’ (rather than ‘grandmother’), including StV mẹ and SV mã, altogether make clear identification of words with phonological differences, despite having initial /m/, problematic.
that this is possibly older in Vietic. Vương 2002 claims that CV bàu has a cognate in Mường pâu, which means ‘they/those people,’ though the semantic difference is significant and cannot be easily explained. It is a possible older Việt-Mường word, but the evidence must be considered tentative at this point. Vương also notes a possible Pông-Chứt cognate for CV tam, namely, Thà vũng saam, which matches a tendency seen in Sino-Vietnamese /t/ for Middle Chinese /*s/.

Clearly, these items are worth investigating further as they could shed light on what might not be Chinese elements in Vietnamese and reflect a pre-Chinese influenced era.

Table 7: CV Terms of Address

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ả</td>
<td>chị</td>
<td>sister, elder</td>
</tr>
<tr>
<td>bàu</td>
<td>(từ nam giới dùng để gọi vợ)</td>
<td>(term used by husband to refer to wife lovingly)</td>
</tr>
<tr>
<td>bọ</td>
<td>cha</td>
<td>father</td>
</tr>
<tr>
<td>ông</td>
<td>anh</td>
<td>(term of address for somewhat older man than the speaker)</td>
</tr>
<tr>
<td>mẹ</td>
<td>bà</td>
<td>grandmother</td>
</tr>
<tr>
<td>mụ</td>
<td>bà</td>
<td>grandmother</td>
</tr>
<tr>
<td>o</td>
<td>cô</td>
<td>sister of father</td>
</tr>
<tr>
<td>tam</td>
<td>em</td>
<td>(term to address younger person)</td>
</tr>
</tbody>
</table>

D. Locative and Deictic Terms

CV has a few distinct etyma for locative words and deictics, in addition to those with phonologically related cognates, as shown in Table 8. These forms are fairly conservative and appear to be early Vietic words or have more archaic pronunciations. The initial /tr/ in CV trúra ‘between’ versus the palatalized StV giữa form shows a conservative tendency in sound (cf. the cognate in the Vietic language Rục troař’). Similarly, CV nì ‘this’ versus StV này are essentially the pronunciation in Mường and Pông-Chứt languages. Finally, the CV locative côi ‘top’ versus StV trên, the latter of which is spread throughout Vietic languages, is noteworthy for having a likely cognate in Mường, synonymous côi (though the tone is readily explainable), showing at least a potential proto-Việt-Mường form. There is also an interesting parallel in the tone change from ngang to huyền in the apparent non-cognates for the two distals indicating ‘over there’ and ‘over there, farther.’ This is an interesting tonal, morphological calque, though which form is the earlier cannot be determined at this point. Altogether, these grammatical words again highlight the distinctiveness CV has as a dialect.

Table 8: Locative and Deictic Terms

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>trúra</td>
<td>giữa</td>
<td>between</td>
</tr>
<tr>
<td>côi</td>
<td>trên</td>
<td>top</td>
</tr>
<tr>
<td>tê</td>
<td>kia</td>
<td>over there</td>
</tr>
<tr>
<td>tê</td>
<td>kia</td>
<td>over there, farther</td>
</tr>
<tr>
<td>dâng tê</td>
<td>dâng kia</td>
<td>over there</td>
</tr>
<tr>
<td>dâng tê tê</td>
<td>dâng kia kia</td>
<td>over there, farther</td>
</tr>
<tr>
<td>dô tê</td>
<td>dô kia</td>
<td>over there</td>
</tr>
<tr>
<td>dươa</td>
<td>kê, canh</td>
<td>next to</td>
</tr>
<tr>
<td>nì</td>
<td>này</td>
<td>this</td>
</tr>
</tbody>
</table>

E. Measure Words

Measure words are not as grammaticalized as what are considered true classifiers, yet they have the function of serving as complements of numerals and taking other nouns as complements, a relatively grammatical function. There are several distinctive CV measure words, as listed in Table 9. CV đầu is also seen in Mường. CV has ré, while StV bánh is originally of Chinese origin (standard Sino-Vietnamese bánh, Chinese 粒, Mandarin bǐng). Similarly, the CV measure word for bread ó parallels the function of the StV Old Chinese loanword chiếc (Sino-Vietnamese chich, 糕, Mandarin zhē). These forms have clearly undergone semantic extension and grammaticalization. More research needs to be done to identify the etymological origins of these words and trace their historical relationships. Nevertheless, these examples highlight some differential patterns in this category.
Table 9: CV Measure Words

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ù</td>
<td>liên</td>
<td>pot (terracotta) of liquid</td>
</tr>
<tr>
<td>vầy</td>
<td>(liều cỏ/rác...)</td>
<td>a quantity of grass, trass, etc.</td>
</tr>
<tr>
<td>ngoại</td>
<td>vòng, bánh</td>
<td>a round of an activity</td>
</tr>
<tr>
<td>ố</td>
<td>chiếc</td>
<td>a loaf of bread, etc.</td>
</tr>
<tr>
<td>rè</td>
<td>bánh</td>
<td>a cake of tobacco</td>
</tr>
</tbody>
</table>

F. Quantity and Intensifier Words and Expressions

Basic numerals are the same in CV and StV, with the phonological exception of initial /l/ in NV năm ‘five’ in numeral compounds (e.g., ‘25’ hai mười lăm). However, in expressing degree and generalized quantity, CV shows a number of differences from StV, as listed in Table 10. The etymological sources of these are not immediately identifiable. One interesting word is the CV word lố for ‘dozen,’ in contrast with the Sino-Vietnamese cognate in StV tả, Chinese 7 打 (Mandarin dā). It may be a cognate with the Hokkien word for dozen, 羅 lô, which has also been borrowed by Cambodian, Thai, and Laotian (Pou and Jenner, 77).

Table 10: Distinct CV intensifying words and expressions

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>bảm</td>
<td>nhiều, làm</td>
<td>very, much</td>
</tr>
<tr>
<td>bồn</td>
<td>nhiều</td>
<td>much</td>
</tr>
<tr>
<td>ët</td>
<td>đơn, chỉ có một</td>
<td>only one</td>
</tr>
<tr>
<td>lén</td>
<td>đầy, nhiều</td>
<td>fully, a lot</td>
</tr>
<tr>
<td>lố</td>
<td>tả</td>
<td>a dozen</td>
</tr>
<tr>
<td>loạn</td>
<td>làm, nhiều</td>
<td>a lot</td>
</tr>
<tr>
<td>mê thiên</td>
<td>rất nhiều</td>
<td>very many</td>
</tr>
<tr>
<td>rèn</td>
<td>đầy, nhiều</td>
<td>fully, a lot</td>
</tr>
<tr>
<td>ba hột</td>
<td>chút ít, không đáng kể</td>
<td>very little</td>
</tr>
<tr>
<td>sáp</td>
<td>bọn, tụi</td>
<td>a group of X</td>
</tr>
</tbody>
</table>

Some intensifying expressions are partial cognates, including those with một, hết, and thấy (with the additional element referring to one’s grandmother). These pattern in an almost prefix-like, calque-like way. Other words in the list are completely distinct etyma, though their specific origin is not yet clear.

Table 11: CV intensifying words and expressions with mixed etyma

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>hết cỡ</td>
<td>hết mức</td>
<td>excessively</td>
</tr>
<tr>
<td>một tẹo</td>
<td>một tí</td>
<td>a little</td>
</tr>
<tr>
<td>một thì</td>
<td>một tí</td>
<td>a little</td>
</tr>
<tr>
<td>thấy mụ nội</td>
<td>thấy bà</td>
<td>terribly</td>
</tr>
<tr>
<td>nhất</td>
<td>nhất</td>
<td>most</td>
</tr>
<tr>
<td>đề sợ</td>
<td>cực kỳ</td>
<td>extremely</td>
</tr>
<tr>
<td>đuì</td>
<td>rất, khi</td>
<td>very</td>
</tr>
<tr>
<td>gắt</td>
<td>ở mức độ rất cao</td>
<td>extremely</td>
</tr>
<tr>
<td>lung</td>
<td>nhiều, quá mức</td>
<td>a lot</td>
</tr>
<tr>
<td>ôm</td>
<td>quá, hết mức</td>
<td>extremely</td>
</tr>
<tr>
<td>bất</td>
<td>rất, quá</td>
<td>very</td>
</tr>
</tbody>
</table>
unit of measurement becomes a generalizing expression (c.f., English approximative expression ‘in the area of X’).

**Table 12: CV adverbs expressing approximation**

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>cỡ</td>
<td>khoảng</td>
<td>about</td>
</tr>
<tr>
<td>cùng</td>
<td>khắp cả, hết tất cả</td>
<td>everywhere completely</td>
</tr>
<tr>
<td>lời</td>
<td>chúng, khoảng chúng</td>
<td>about, approximately</td>
</tr>
<tr>
<td>lời chúng</td>
<td>chúng, khoảng chúng</td>
<td>about, approximately</td>
</tr>
</tbody>
</table>

**G. Time Words and Expressions**

In this category, many of the words are compounds, and of many of those compounds, there are many instances of shared cognates, as shown in Table 13. For example, CV khi parallels StV lúc and indicates time in many compounds. The other morphs in the compounds have the same meaning in many instances, such as CV khi sớm ‘while early’ versus StV lúc sớm.

**Table 13: CV Khi versus StV Lúc**

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>khi đầu</td>
<td>lúc đầu</td>
<td>at the beginning</td>
</tr>
<tr>
<td>khi đêm</td>
<td>lúc đêm</td>
<td>at night</td>
</tr>
<tr>
<td>khi hồi</td>
<td>lúc nãy</td>
<td>just now</td>
</tr>
<tr>
<td>khi nơi</td>
<td>lúc này</td>
<td>at this time</td>
</tr>
<tr>
<td>khi sớm</td>
<td>lúc sớm</td>
<td>while early</td>
</tr>
<tr>
<td>khi tê</td>
<td>lúc trước</td>
<td>previously</td>
</tr>
</tbody>
</table>

One morph of historical note is StV giờ, which has been posited (Pulleyblank, 283) as an Old Sino-Vietnamese loanword (standard Vietnamese thì, Chinese 时, Mandarin shí) originally meaning ‘time’ and which has come to mean in Vietnamese ‘hour’ or more generally ‘at this time.’ CV has the clearly related cognate giừ, but also the likely cognate chíù. The two readings giừ and chíù appear to be phonological alternants in CV, but are still related to the StV form of the same etymon. The likely older form is chíù considering its non-palatalized initial, which, if accurate, would not be an insignificant fact in understanding Sino-Vietnamese historical linguistics.

**Table 14: CV words meaning ‘now’**

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>bây chíù</td>
<td>bây giờ</td>
<td>now</td>
</tr>
<tr>
<td>bây giù</td>
<td>bây giờ</td>
<td>now</td>
</tr>
<tr>
<td>giừ</td>
<td>giờ</td>
<td>now</td>
</tr>
</tbody>
</table>

The large number of time words and expressions in CV which differ from those in StV clearly draws attention. Table 15 shows this substantial list, some of which involves a contrast between CV bữa and StV hôm in several compounds.
Table 15: Various CV time words and expressions

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>bữa điệp hôm kia</td>
<td>hôm kia</td>
<td>day before yesterday</td>
</tr>
<tr>
<td>bữa nỉ ngày hôm nay</td>
<td></td>
<td>today</td>
</tr>
<tr>
<td>bữa ráy bữa hôm ấy đến nay</td>
<td></td>
<td>from that day to today</td>
</tr>
<tr>
<td>bữa tế hôm kia</td>
<td></td>
<td>day before yesterday</td>
</tr>
<tr>
<td>giấc dặ, khoảng, lúc</td>
<td></td>
<td>around this time</td>
</tr>
<tr>
<td>hồi tế hồi sau, lúc sau</td>
<td></td>
<td>at a previous time</td>
</tr>
<tr>
<td>mai lối ngày trước nói khái quát</td>
<td></td>
<td>a previous day, non-specific</td>
</tr>
<tr>
<td>may hôm chăng may lúc</td>
<td></td>
<td>not long</td>
</tr>
<tr>
<td>may nã chăng được may chỗ</td>
<td></td>
<td>not even a moment</td>
</tr>
<tr>
<td>mới ngày kia</td>
<td></td>
<td>two days later</td>
</tr>
<tr>
<td>xém súyt</td>
<td></td>
<td>nearly, about to</td>
</tr>
<tr>
<td>rê vày</td>
<td></td>
<td>at this time</td>
</tr>
</tbody>
</table>

CV also shows a number of distinct time adverbs of frequency. Notable among these are some reduplicants, such as CV **hoài hủy** and **lần lần**, both of which parallel reduplication in StV. As noted in section 3B above, CV **năng** is an archaic form of Vietnamese.

Table 16: Time adverbs of frequency

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>lút tuột, thẳng một mạch</td>
<td></td>
<td>continuously</td>
</tr>
<tr>
<td>thết mãi</td>
<td></td>
<td>continuously</td>
</tr>
<tr>
<td>hoài hủy mãi, luôn luôn</td>
<td></td>
<td>continuously, forever</td>
</tr>
<tr>
<td>lần dần</td>
<td></td>
<td>gradually</td>
</tr>
<tr>
<td>lần lần dần dần</td>
<td></td>
<td>gradually</td>
</tr>
<tr>
<td>năng hay, thường</td>
<td></td>
<td>often</td>
</tr>
</tbody>
</table>

H. Conjunctives

Among connective words in CV are some phonological differences and some compounds with CV etyma, such as **răng** and **rứa**, as noted above. As seen in other cases, these compounds appear almost like calques, though it is impossible to determine whether CV or NV first developed the terms and which variety calqued which originally. Regardless, the semantic patterns in these words are suggestive of a kind of linguistic area.

Table 17: CV Connective Words

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>dù răng</td>
<td></td>
<td>although</td>
</tr>
<tr>
<td>dùng</td>
<td></td>
<td>but</td>
</tr>
<tr>
<td>bởi rứa</td>
<td></td>
<td>thus</td>
</tr>
<tr>
<td>giả thì như</td>
<td></td>
<td>for example</td>
</tr>
<tr>
<td>hèn chi</td>
<td></td>
<td>that is why</td>
</tr>
<tr>
<td>nên chi</td>
<td></td>
<td>thus</td>
</tr>
<tr>
<td>phải chi</td>
<td></td>
<td>supposing that</td>
</tr>
<tr>
<td>phương chi</td>
<td></td>
<td>much less</td>
</tr>
</tbody>
</table>

I. Sentence Particles

The dialect dictionary contains at least ten CV sentence particles, though it is difficult to determine whether some of these are cognates with phonological variants or distinct etyma. It is tempting, for instance, to consider CV **nà**, an emphatic particle, to be related to StV **nào**, and the CV question particle **hi** appears phonologically similar to StV **hả**. However, neither of these forms fall under a common pattern of phonological correspondences between CV and StV. Sentence particles are semantically highly abstract and less readily perceived phonetically. It is thus, perhaps, not surprising that many of the dialectal forms, while
having similar functions, vary phonologically but only slightly. Consider the semantic and phonological similarities between CV particles nghe and nha with StV né, all of which express urging, though again without phonological patterns to clearly related them.

Table 18: CV sentence particles

<table>
<thead>
<tr>
<th>CV</th>
<th>StV</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>hãy</td>
<td>hà</td>
<td>emphatic question</td>
</tr>
<tr>
<td>hãy</td>
<td>nhi</td>
<td>emphatic question</td>
</tr>
<tr>
<td>hỉ₁</td>
<td>nhi</td>
<td>emphatic question</td>
</tr>
<tr>
<td>hỉ₂</td>
<td>hà</td>
<td>question</td>
</tr>
<tr>
<td>hè</td>
<td>nào</td>
<td>urging</td>
</tr>
<tr>
<td>nà</td>
<td>nào</td>
<td>urging</td>
</tr>
<tr>
<td>nghe</td>
<td>nhé</td>
<td>urging</td>
</tr>
<tr>
<td>nha</td>
<td>nhé</td>
<td>urging</td>
</tr>
<tr>
<td>nì</td>
<td>này</td>
<td>urging</td>
</tr>
<tr>
<td>nờ</td>
<td>não, nào</td>
<td>urging</td>
</tr>
</tbody>
</table>

It is perhaps best to assume that there are various cognates among these words, but that due to their more peripheral nature in the lexicon, being nearly devoid of concrete meaning and most often occurring in sentence-final position, it is often the case that cognates cannot be claimed with certainty. It may be more realistic to posit a loose or even phonaestheme-like element for this category of words.

4. Concluding Thoughts

Despite the focus on grammatical vocabulary, the data repeatedly demonstrate how CV is distinct from StV as a dialect, thereby suggesting how it is a distinct dialect region historically. It also shows in many instances—both phonologically and lexically—how close CV is to the older stages of Việt-Mường and Vietic in Vietnamese linguistic history. At the same time, the data are illustrative in various ways of a kind of linguistic area. In numerous instances, compounds with mixed cognates are suggestive of how CV and StV parallel each other in word-formation strategies. Indeed, in some cases, one may posit that these word-formation strategies are suggestive of derivational morphology, a counter-argument to the position that Vietnamese has strictly isolating morphology. This small study illustrates how grammatical vocabulary can serve to clarify the relationships among dialects.

References


