Annotation request for Vietnamese apex

Minh Nguyen, mxn, 1ec5.org
Kirk Miller, kirkmiller, gmail

2024 April 05

SAH: Our first suggestion is Latin combining *ur*, U+1DD1. The Cyrillic *titlo*, U+0483, is a possibility if it is redefined to support Latin use. There would be even less possibility of conflict with *titlo* than with combining *ur*, and it should have broader font support. (Though it would still be a good idea to change the representative glyph of combining *ur* in the charts.) However, Denis comments that the shape of the *titlo* "would need to be different in Middle Vietnamese as it's often too angular or too wavy and has the same hook turned on both ends instead of a more prominent one on the right." A font that supports both Latin and Cyrillic is unlikely to accommodate a Cyrillic character to Vietnamese. Combining *ur* is more representative graphically. Another possibility is U+0342 COMBINING GREEK PERISPOMENI, but that is historically inappropriate.

We request an annotation, to the medievalist character U+1DD1 COMBINING UR ABOVE ⟨ŏ̐⟩, that it is the appropriate Unicode character for what is known in modern Vietnamese as *dấu sóng* or *dấu lưỡi câu*, a diacritic ⟨o̐⟩ of Middle Vietnamese Latin orthography that marked a syllable-final nasal consonant. Since its description in Rhodes (1651), the diacritic has had the Latin name *apex*, from the Portuguese term for what is now called the *til* or *tilde*. The *dấu sóng* is distinct from the Middle and Modern Vietnamese tilde-shaped *dấu ngã* tone mark, ⟨õ⟩, which derives from the Greek *perispomeni*, ⟨õ̑⟩, not the tilde, nor is there any apparent connection to the Classical Latin *apex*, ⟨ó̆⟩, which marked long vowels. In this document, we will call the *dấu sóng* nasalization mark ⟨o̐⟩ the 'apex,' and both the *dấu ngã* tone mark ⟨õ⟩ and the Portuguese nasalization mark the 'tilde.'

We also request a change of the representative glyph in the Unicode code chart to one that is closer to Latin sources (Figure 22 *ff*) and also not dissimilar to some typesetting of Middle Vietnamese (Figure 6).

In Classical Latin, the *apex* ⟨ő̆⟩ marked long vowels (Figures 1–2) and was the ancestor of the acute. In modern typesetting it is generally replaced with the macron, which was borrowed from Greek poetic convention, though for graphic fidelity the acute is sometimes used in modern typesetting. We do not see need for a separate Unicode character for the Latin apex: a dedicated font should be able to handle its graphic peculiarities with the existing acute.

The Vietnamese apex marks a nasal consonant, and it may co-occur with the acute and tilde tone marks, *dấu sắc* and *dấu ngã*. Specifically, the Vietnamese apex transcribes a syllable-final [ŋ͡m], an allophone of /ŋ/ found after rounded back vowels in the Hanoi area (thus ⟨o̐⟩ and ⟨ŭ̐⟩). It was

unified with orthographic ⟨ng⟩ in the mid-18th century with survivals as late as Bỉnh (1822). The apex is of modern interest for reproducing and discussing seminal works of Vietnamese Latin orthography that almost any literate Vietnamese-speaker would know of. A modern description of the phonetics is given in Jacques (1998: 52), reproduced as Figure 11. De Rhodes' contemporary description is reproduced as Figure 12.

Graphically, the Vietnamese apex resembles a tilde or macron with a curled right terminal, ⟨ȏ̃⟩ or ⟨ȭ⟩. The curled-tilde allograph ⟨ȏ̃⟩ is seen in some modern publications (Figure 6), though the curled-macron allograph ⟨ȭ⟩ is more common. In manuscript, the apex may be a simple curl joined directly to the letter *o*, resembling the horn on the letter *ơ* but larger and contrasting with it (Figure 7). Indeed, in manuscript the apex may occur as an abbreviation for *ng* on the letters *ơ* and *ư*, for ⟨ơ̑⟩ and ⟨ữ⟩ (Figure 8 *ff*), though because *ơ* and *ư* aren't rounded vowels, the apex in such cases is simply /ŋ/ rather than the [ŋ͡m] allophone of /ŋ/.

Some digital texts use custom fonts to repurpose assigned Unicode points for the apex, such as the character *ả* that displays as *ū* in Bùi (2021) (Figure 17). Publications using typographic substitutions in non-custom fonts have been unsatisfactory. Nguyễn (1993), Nguyễn (1997) [no relation] and Jacques (1998), among others, use a tilde. This is unfortunate because both Middle and Modern Vietnamese orthography use the tilde as a tone mark, and this can co-occur with the apex in the same text and even on the same letter (Figure 3). An article by Lê & Nguyễn (2013) attempts a different substitution, but something went haywire during typesetting and the text was rendered gibberish (Figure 16).

Thanks to Denis Jacquesrye for research into the history of the Portuguese tilde, and to Harald Tveiten for research into the MUFI background of the Latin combining *ur*.

Conflict with the Unicode tilde

The Vietnamese apex dates to the Portuguese missionaries who devised the Vietnamese Latin alphabet in the early 1600s. *Apex* was the Latin term used at the time for the Portuguese tilde, as attested by dictionaries such as [J. Cardoso (1592) *Dictionarium Latino Lusitanicum*](#), which equates Latin 'apex' with Portuguese 'til da letra' (tilde). The Portuguese tilde of the time was flatter than it is today (Figure 21), and was indistinguishable in form from the contemporary Vietnamese apex ⟨ȭ⟩. (The modern wavy shape of the tilde dates from the late 18th century.) Thus the *til da letra*, under its Latin name *apex*, was the source of the Vietnamese apex in form, name and sound value (Jacques 1998: 52). This creates a problem for Unicode, because U+0303 COMBINING TILDE is used for the Portuguese tilde, but in Vietnamese it is used for the unrelated tilde-shaped tone mark, *dấu ngã*, which derives instead from the Greek *perispomeni* ⟨ȭ⟩ that can have that shape. De Rhodes (1651: 8–10) called the *dấu ngã* 'circumflexus' (circumflex), a calque of the Greek *perispomeni*. (The

Vietnamese vowel mark found on *â, ê, ô* that we now know as the 'circumflex' was called *accentus circumflexus latinorum* 'Latin circumflex accent' by de Rhodes.)

If Unicode had been based on 17th-century usage, the nasal Vietnamese *apex* would have been assigned U+0303 COMBINING TILDE along with its cognate in Portuguese orthography, and the tonic Vietnamese *circumflexus* would likely have been assigned to U+0342 COMBINING GREEK PERISPOMENI. Our problem is that it's the *circumflexus* that was encoded as U+0303.

Vietnamese mark	Encoding	Origin	Encoding
dấu sóng, ⟨ȏ̃⟩ (apex)	(TBD)	Portuguese tilde, ⟨õ⟩	U+0303
dấu ngã, ⟨õ⟩ (circumflexus)	U+0303	Greek *perispomeni*, ⟨õ⟩	U+0342

Table 1. Inconsistent use of U+0303 COMBINING TILDE for the Portuguese *til da letra / apex* and the unrelated Vietnamese *dấu ngã / circumflexus* ⟨õ⟩, rather than for the tildes of both Portuguese and Vietnamese (*dấu sóng / apex* ⟨ȏ̃⟩). Encoding the *dấu sóng* as the *perispomeni* (U+0342) would also be problematic, partly because that is the source of the *dấu ngã* but also because the *perispomeni* and tilde are shaped identically in many fonts.

The SAH has decided that creating a new character for the Vietnamese apex would be inadvisable, because doing so would call into question the identity of the 17th-century Portuguese 'apex,' which after all was the same character. It would thus be preferable for the UTC to declare some existing character to be the one to use for the Vietnamese apex, and to annotate it accordingly.

Character to annotate for use as the Vietnamese apex

The SAH is of the opinion that a distinct Unicode character for the Vietnamese apex is not justified, and that the the best solution would be to repurpose an existing character. Apart from the tilde itself and the Greek *perispomeni*, which is problematic for reasons mentioned in Table 1, the two most visually similar characters currently available in Unicode are the medievalist U+1DD1 COMBINING COMBINING UR ABOVE ⟨ȏ̃⟩ and U+0483 COMBINING CYRILLIC TITLO ⟨ȏ⟩.

With the acceptance of L2/06-027 by Unicode, the MUFI character M+F1CC COMBINING CURLY BAR ABOVE was conflated with M+F1C2 COMBINING ABBREVIATION MARK SUPERSCRIPT UR LEMNISKATE FORM into the Unicode character U+1DD1 COMBINING UR ABOVE. Janusz Bień (author of Bień 2021) comments on GitHub issue #11 (concerning Junicode font) that the lemniscate form of M+F1C2 should be deprecated. As noted in that GitHub discussion, the figures supplied by L2/06-027 match M+F1CC rather than M+F1C2. (L2/06-027 figures 60 and 61 are from Balbi (1460), which is shown at higher resolution in Figure 23 below. The other item claimed to justify U+1DD1 was figure 32, but the symbol doesn't actually occur there.) It is therefore unfortunate that the representative glyph

in the Unicode chart is based on M+F1C2 rather than on M+F1CC, and in our opinion it should be replaced. M+F1CC as attested in historical Latin sources such as Zaborowski (1514) and Balbi (1460), as well as in modern fonts such as Junocode, closely approaches the desired Vietnamese shaping. The glyphs for M+F1CC in Junicode and Cardo fonts are:

ũ ũ

The Cyrillic *titlo*, U+0483 COMBINING CYRILLIC TITLO ⟨◌҃⟩, has the advantage of broader font support, and a negligible chance of conflict in a Vietnamese Latin context, but the disadvantage of being too angular and having a turn at both ends rather than only on the right as the Vietnamese apex does. Specialized fonts that would give it its Vietnamese shape would be unlikely, especially online, given how widely it is supported for a Cyrillic context. The glyphs for the *titlo* in Brill and Noto Serif fonts are:

ũ ū

Wikimedia wikis substitute the IPA pitch-contour character U+1DC4 COMBINING MACRON-ACUTE ⟨◌᷄⟩. This workaround is present on 43 pages across three books in the Vietnamese Wikisource, 11 entries in the English Wiktionary and six in the Vietnamese Wiktionary (Figure 20). It would be trivial to replace these instances with whichever character the UTC decides on, and there is as yet no known usage of U+1DC4 for the apex outside Wikimedia. One problem with Unicode adopting U+1DC4 as a permanent solution is that a text about Middle Vietnamese could well include IPA transcription, since Middle Vietnamese orthography differs nontrivially from the modern and makes distinctions that have since disappeared, and pitch contours such as the one encoded in U+1DC4 are an essential element of the tones that Vietnamese is famous for. Using an IPA tone diacritic for an unrelated orthographic convention could thus be a problem with forward compatibility. Another problem is that U+1DC4 is clearly identified as a ligature of a macron and an acute, and that identity – combined with its broad font support as IPA – makes it unlikely that it will be given the proper Vietnamese form in fonts. The glyph in Gentium font is:

ű

We therefore request that the UTC annotate U+1DD1 ᴄᴏᴍʙɪɴɪɴɢ ᴄᴏᴍʙɪɴɪɴɢ ᴜʀ ᴀʙᴏᴠᴇ for use as the Vietnamese apex. If the UTC should decide instead on the Cyrillic *titlo*, U+0483, it will require redefinition in ScriptExtensions.txt for use as a Latin character.

Annotation

0303 COMBINING TILDE
    → 1DD1 combining ur *above*
1DD1 COMBINING UR ABOVE
    = Vietnamese *apex* (dấu sóng)

ScriptExtensions.txt [if the *titlo* is chosen]

U+0483 ᴄᴏᴍʙɪɴɪɴɢ ᴄʏʀɪʟʟɪᴄ ᴛɪᴛʟᴏ is defined only for Cyrillic script, with scx = Cyrl Perm. Latn will need to be added if it is to be used for the Vietnamese apex. (Glag should also be added, per GitHub issue #375.)

Representative glyph in the code chart

We request that the glyph for U+1DD1 ᴄᴏᴍʙɪɴɪɴɢ ᴄᴏᴍʙɪɴɪɴɢ ᴜʀ ᴀʙᴏᴠᴇ in the Unicode chart for Combining Diacritical Marks Supplement (1DC0–1DFF) be changed to ⟨◌̆⟩. This is both more representative of the medieval Latin form (Figure 22 *ff*) than is the current glyph on the code chart, and more appropriate for Middle Vietnamese. We will be happy to provide the font.

References

Balbi, Giovanni (1460) *Catholicon.* High-resolution scans available at:
    www.loc.gov/resource/rbc0001.2012rosen0031/?st=gallery
Bień, Janusz (2021). Traktat Stanisława Zaborowskiego i Skróty Brachygraficzne. *Poznański Półrocznik Językoznawczy*, no. 1 (30).
Bỉnh, Philipphê (1822). *Sách sổ sang chép mọi việc* [Book recording all things]. Page images available online at Vietnamese Wikisource.
Bùi Thị Minh Thùy (2021) *Từ điển VIỆT – BỒ – LA và các cứ liệu liên quan.* ['Dictionarium Annamiticum

Lusitanum et Latinum' and related documents.] Nhà Xuất Bản Tôn Giáo.

Đỗ Quang Chính (2008) [1972]. *Lịch sử chữ quốc ngữ 1620–1659, in lại nguyên văn theo ấn bản của Tủ Sách Ra Khơi – Saigon 1972 có sửa chữa của tác giả* [History of the national language of Vietnam, 1620–1659, reprinted as per the edition by Ra Khoi Books – Saigon 1972 with corrections by the author]. Nhà xuất bản Tôn giáo, Hanoi. 2nd ed.

Lê Khắc Cường & Nguyễn Thị Phương Trang (2013). Tiếng Việt trong 'Sách sổ sang chép các việc' của PHILIPHÊ BỈNH [Vietnamese in 'Book recording all things' by PHILIPHÊ BỈNH]. *Tạp chí Ngôn ngữ và Đời sống* [Journal of Language and Life] 12 (218): 2–14.

Jacques, Roland (1998). Le Portugal et la romanisation de la langue vietnamienne. Faut-il réécrire l'histoire? [Portugal and the romanization of the Vietnamese language. Should we rewrite history?] *Revue française d'histoire d'outre-mer*, 85 (318). Société française d'histoire d'outre-mer.

Nguyễn Đình-Hoà (1997). *Vietnamese.* [*Tiếng Việt Không Son Phấn:* Vietnamese without veneer]. John Benjamins, Amsterdam.

Nguyễn Khắc Xuyên (1993). *Ngữ Pháp Tiếng Việt của Đắc Lộ 1651* [De Rhodes's Vietnamese Grammar of 1651]. Thời Điểm, Garden Grove, CA. OCLC 32129692.

Rhodes, Alexandre de (1991) [1651]. Về các dấu và dấu hiệu khác trên nguyên âm. In Hồ Lê, Cao Xuân Hạo & Hồ Tuyết Mai (eds.) *Từ điển Annam-Lusitan-Latinh (Thường gọi Từ điển Việt-Bồ-La).* Translated by Thanh Lãng, Hoàng Xuân Việt & Đỗ Quang Chính. Social Science Publishing House, Ho Chi Minh City. LCCN 91943138. High-resolution images of the original (de Rhodes 1651) available online at the Bayerische Staatsbibliothek.

———— (1651b) *Phép giảng tám ngày cho kẻ muấn chịu phép rứa tọi, ma bĕào đạo thánh đức Chúa blời / Cathechismus pro ijs, qui volunt suscipere baptismum in octo dies divisus* [Catechism for those who want to be baptized and enter the faith in eight days]. Sacrae Congr. de Propaganda Fide [Congregation for the Propagation of the Faith], Rome. High-resolution images available online at the Bayerische Staatsbibliothek.

Thụy Khuê (2023). Bối cảnh lịch sử Đàng Trong. In *Truyền giáo và quốc ngữ*. Reprinted in Truyền giáo và quốc ngữ: Chương 1 – Bối cảnh lịch sử Đàng Trong. *Da Màu*, Dec 08. Da Màu Foundation.

Figures
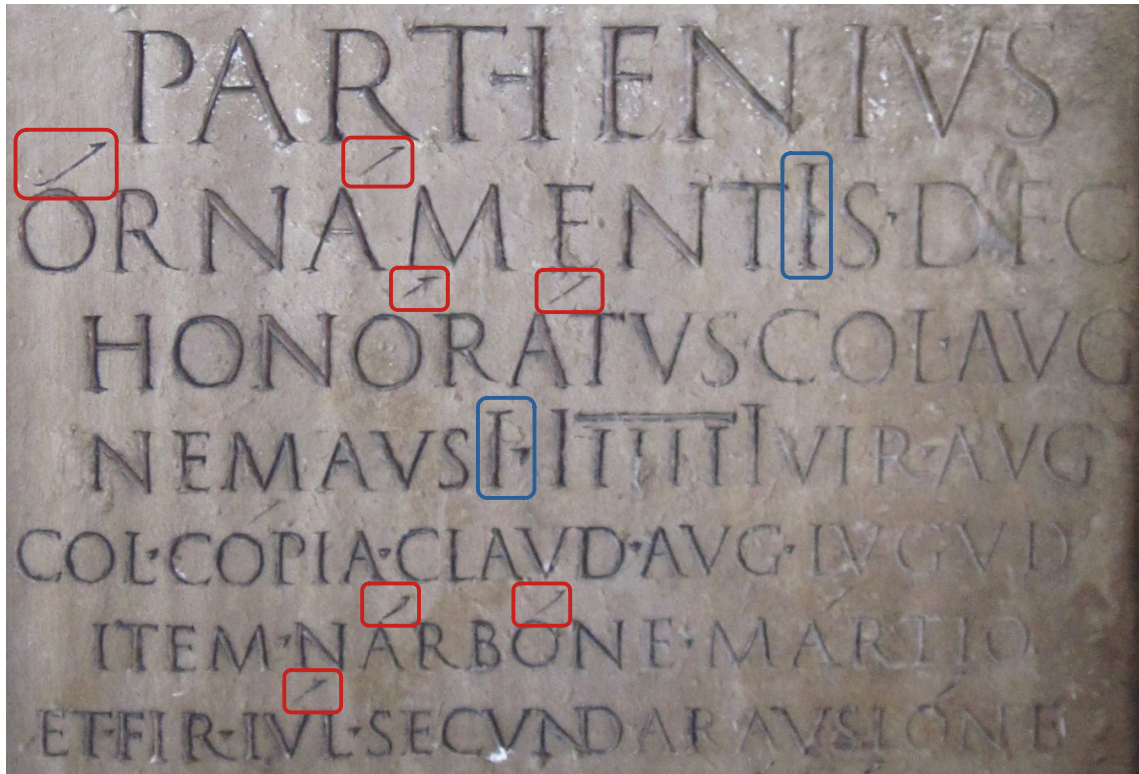
Figure 1. Late 1st – early 2nd-century Latin inscription with apices, ⟨◌́⟩. The diacritic is light-weight and has a serif at the end; the serif becomes a round hook in handwriting. The Latin apex is used on the letters A E O V but not I, which instead of taking the apex is substituted with *i longa*, U+A7FE ⟨Ɪ⟩ (blue). The apex tends to bear to the right, but in handwriting it may attach to the left stroke in the case of the letter V.
CIL XII 3203 in the Musée archéologique de Nîmes. Image from Wikimedia Commons.

Figure 2. In this Latin example the apex connects to the top-right corner of the letter E and overlies the following letter.
'Pilate stone' (cast), 1st century, Caesarea Maritima. Image from Wikimedia Commons.

Figure 3. de Rhodes (1651b: 14; 1651: 147). Some clear examples of the common word *cũ* "also" that combines the Vietnamese apex and tilde.

Figure 4. de Rhodes (1651b: 14). The apex on *cũ̽*, *khoũ* and *thoũ* (red) and the horn on *blời* and *những* (blue) appear to be typeset with the same element, but the apex lies above the letter whereas the horn connects to the side. Note also the recently unicodified letter ƀ in *ƀêy*.
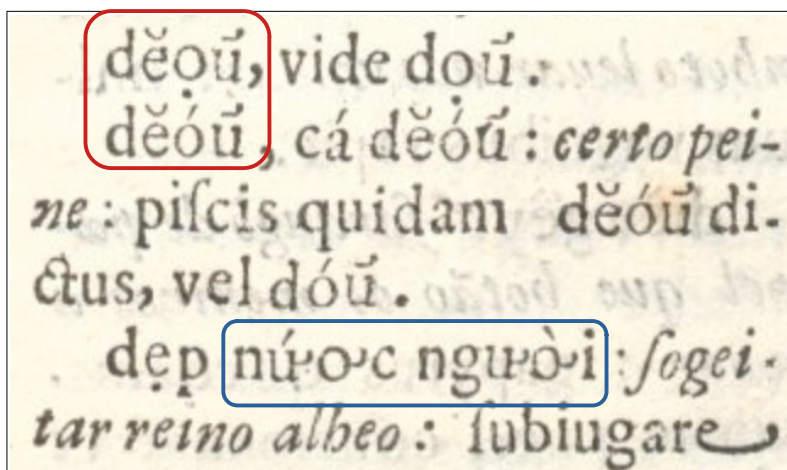
Figure 5. de Rhodes (1651: 170). Apex alongside other diacritics in *děọū́* and *děóū́*. The apex here is similar in shape to the hooks in *nước người* (blue).
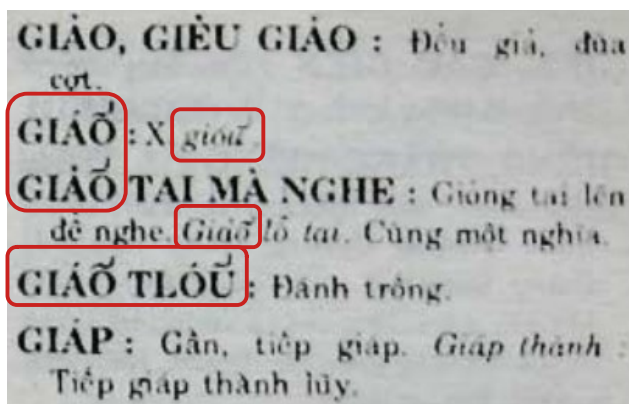
Figure 6. de Rhodes (1991 [1651]: 105). The Vietnamese apex entered by hand on the words *giấố tlốū̃* in a modern reprinting. In some tokens its origin in a Portuguese tilde is apparent.

Figure 7. Bỉnh (1822: 2). A late manuscript with the apex in *tôi cử chưng* (together with a grave); *ơn trạớ Đức* ; *phá Dầớ Đ.C.J. ở trạớ địa* ; *Saớ le trước* ; *Cụ Tito cử Cụ Bảo* and various other tokens of *Dầớ* (red), where it joins with the letter *o*. It contrasts with the tilde in, for example, *Vĩnh lại huện, Ngãi* (blue). Transliteration from [Sach so sang chep cac viec.pdf/23](#) at Vietnamese Wikisource.
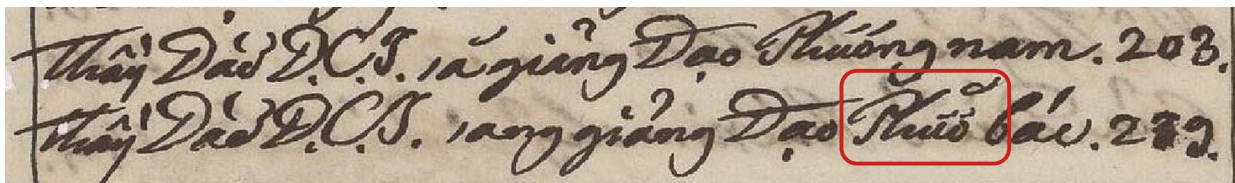
Figure 8. Bỉnh (1822: TOC, 23). Manuscript tokens of ⟨ở⟩, the apex on the horned letter *ơ*, in the word *phương* 'direction.' *Phương* is repeatedly abbreviated *phươ* in this document. Here the apex is a simple abbreviation for final *ng* and does not mark the allophonic value [ŋ͡m] that it would on *o* or *u*. The first context is *Thầy Dàõ D.C.J. sang giảng Đạo Phươ bác [page] 239*, where the apex also appears on ⟨õ⟩, and following an unusual token of *phương* spelled out in the preceding line. Note that the horns of the letters *ươ* are detached and resemble a typeset apex. The second context is *nừ Phươ Tây*, short for *người Phương Tây* 'Westerner.'
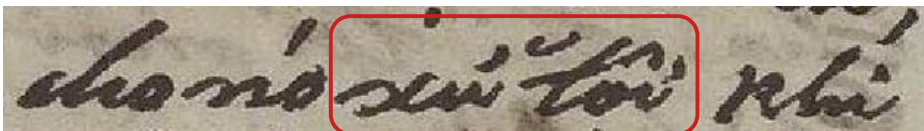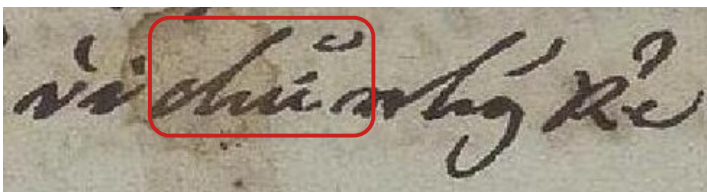
Figure 9. Bỉnh (1822: 50, 36). Manuscript tokens of ⟨ư⟩, the apex on the horned letter *ư*. contrasting with the tilde (on *nhữ*) in *vì chư nhữ kẻ*, where *vì chư* is short for *vì chưng* 'because of'; and in *cho nó xư tội khi*, where *xư tội* is short for *xưng tội* 'to confess.'

11

Sau hết dấu hiệu thứ ba là dấu lưới câu tối cần thiết trong ngôn ngữ *này, vì sự khác biệt của âm cuối mà dấu lưới câu làm nên, khác hẳn âm cuối mà,* **m,** *hay,* **n,** *làm nên, và, những tiếng mang dấu lưới câu thì ý nghĩa cũng khác hoàn toàn; vả lại dấu hiệu này hay dấu* lưới *câu chỉ cảm nhiễm* ố, và, ủ, *ở cuối tiếng thì dụ,* aố, *apis (con ong),* oủ, *auus, vel dominus (ông già, hay ông chủ) : nó được phát âm thế nào để môi không chạm vào nhau, mà lưới cũng không chạm tới của.*

Figure 10. de Rhodes (1991 [1651]: 11). Description of the diacritic in Vietnamese, with examples. A tilde is also used (blue).

Dans la version du quốc ngữ dont est témoin le dictionnaire d'Alexandre de Rhodes, les consonnes nasales finales dorso-prépalatale, dorso-vélaire et labio-vélaire font l'objet de trois notations graphiques différentes, respectivement « nh » (par ex. « *lành* », « bon, doux ») ; « ng » (par ex. « *làng* », « village »), et un signe dénommé *apex*, dérivé du *til* (tilde) portugais « ~ », placé au-dessus de la voyelle (par ex. « *làõ* », aujourd'hui « *lòng* », « ventre, cœur »). Ces notations reflètent bien la prononciation standard de la région de Hà Nội, tel qu'un auditeur portugais attentif peut l'analyser dans le système de notation auquel il est habitué. Dans la version fixée par le dictionnaire de Pigneau de Béhaine et Taberd (cf. note 110), les deux derniers types sont notés uniformément « ng », du fait qu'il ne sont pas en opposition, leur apparition dépendant de la voyelle précédente (dans le jargon de la phonologie, ils sont en « distribution complémentaire »). La distinction entre les deux est donc superflue pour un locuteur vietnamien. La simplification aurait pu aller jusqu'au bout, en confondant les notations « nh » et « ng » : leur distinction est peu rentable pour le système vietnamien, voire sans utilité si l'on note différemment les voyelles / a / et / æ /.

Figure 11. Jacques (1998: 52). Background information on the Vietnmese apex:
In the version of Quốc Ngữ that the dictionary of Alexander de Rhodes witnesses, the dorso-preplatal, dorso-velar and labio-velar final nasal consonants are the target of three different spellings, respectively "nh" (for example "lành," "good, sweet"); "ng" (for example "làng," "village"), and a symbol called apex, *derived from the Portuguese* til *(tilde) "~", placed above the vowel (for example "làõ," today "lòng," "belly, heart"). These spellings nicely reflect the standard pronunciation of the region of Hanoi, so that an attentive Portuguese listener can interpret them per the orthographic system that he is accustomed to. In the version established by the dictionary of Pigneau de Béhaine and Taberd (see note 110), the last two are uniformly transcribed "ng", due to the fact that they are not in opposition, their appearance depending on the preceding vowel (in the jargon of phonology, they are in "complementary distribution"). A distinction between the two is therefore not necessary for a Vietnamese-speaker. The reduction could have been completed by conflating the spellings "nh" and "ng": their differentiation is not helpful for the Vietnamese [orthographic] system, and would be unnecessary if we wrote the vowels /a/ and /æ/ differently.*

Figure 12. de Rhodes (1651; Appendix *Linguae annamiticae seu tunchinensis breuis declaratio*: 10). Contempory description of the apex. (The first two 'signs' mentioned are the circumflex and breve, which indicate vowel quality; these contrast with the 'accents' that mark tone and co-occur with the 'signs.')

Tertium denique ſignum est apex, qui in hac lingua eſt omninò neceſſarius propter differentiam terminationis, quam facit apex omnino diſtinctam à terminatione quam facit, m, aut, n, cum ſignificatione omninò diuerſa in dictionibus, quibus adhibetur; hoc autem ſignum ſeu apex ſolùm afficit ỗ, & ữ, in fine dictionis, ut, aỗ, apis; oữ; auus, vel dominus: pronunciatur autem ita, vt nec labia ſe inuicem contingant, nec lingua palatum attingat.

Finally, the third sign is the apex, which is absolutely necessary in this language because of the difference in the ending that the apex makes, which is entirely distinct from the ending that *m* or *n* makes, with a completely different meaning in the words in which it is used. However, this sign, namely the apex, only affects *ỗ* and *ữ* at the end of a word, as in *aỗ*, 'bee,' *oữ*, 'grandfather' or 'lord.' It is pronounced such that neither the lips touch each other nor the tongue touches the palate.

Figure 13. Nguyễn (1993: 39). The apex (substituted with a tilde) on Middle Vietnamese *aỗ* (modern aong, ong), *cữ* (cũng), *maỗ* (maóng, móng), *oữ* (oung, ông), *sóữ* (soúng, sống), *tổ tữ* (tổ tung, tổ tông). An orthographic tilde occurs on *Ngữ* and *cũng* (blue), so the substitution of tilde for apex conflates the diacritics on *cữ* and *cũng*.

Figure 14. Nguyễn (1997: 33). A well-cited linguistics book on Vietnamese that also substitutes a simple tilde for the apex.

Figure 15. The entry for Cochinchina (2753942) at OpenHistoricalMap.org. The Middle Vietnamese name is given as ⟨đàng tlaõ⟩, based on sources such as Đỗ (2008) and Thụy (2023). The only indication to a naive reader that the tilde in *tlaõ* is not the tilde diacritic that marks tone is that it appears over the *o* rather than the *a*; the other clue is that it corresponds to *-ng* in modern orthography, which is ⟨Đàng Trong⟩ (blue).

Figure 16. Lê & Nguyễn (2013: 4). A modern attempt to reproduce the apex, approximated as a *dấu ngắn* (˘), in a Vietnamese-language text without Unicode support. It comes out variously as *&&* and *((*.

Figure 17. Bùi (2021: 21). Typesetting of the apex (red), in the same text as the tilde (blue), in the publisher's preview of a recent publication. Copying and pasting the word *trũ* from the PDF, we see it is encoded as ⟨trả⟩.

Figure 18. Bùi (2021: 76–77). Although the resolution is a bit low, we can see the apex on the words *blở làỡ*, *đaỡ ɓơi đấu* and *tử ɓốn*, along with the tilde on *ngũ* and the recently unicodified Middle Vietnamese letter ⟨ɓ⟩.

Figure 19. Đỗ (2008 [1972]: 74). Substitution of the apex with a tilde (*đàng tlaõ*, red) in a text that also uses the tilde as a tone mark (*nghĩa*, blue).

Figure 20. Entries for *xaỗ* on English Wiktionary and for *làỗ dữ* on Vietnamese Wiktionary, with U+1DC4 ◌̈ COMBINING MACRON-ACUTE substituted for the apex.
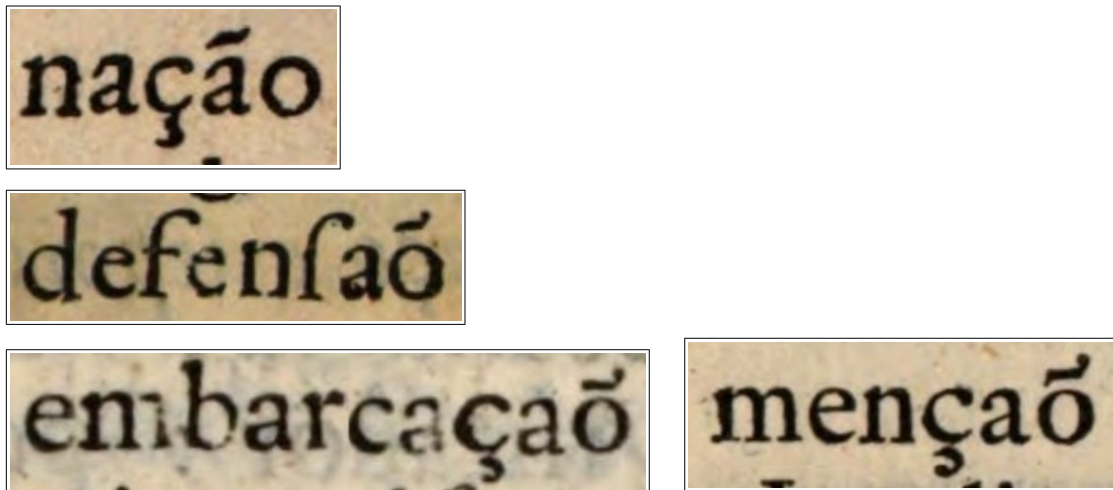
Figure 21. Flat tildes in 17th- and 18th-century Portuguese publication:

(a) *nação* in D. Nunes de Leão (1610) *Descripção do Reino de Portugal*.

(b) *defensaõ* in G. Estaço (1625) *Varias Antiguidades de Portugal*.

(c) *embaraçaõ* and *mençaõ* in L. de Menezes (1759) *Historia De Portugal Restaurado*, vol. 2.

These are the *apex* ⟨◌̃⟩ of Middle Vietnamese.

17
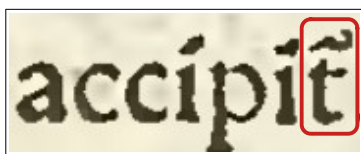
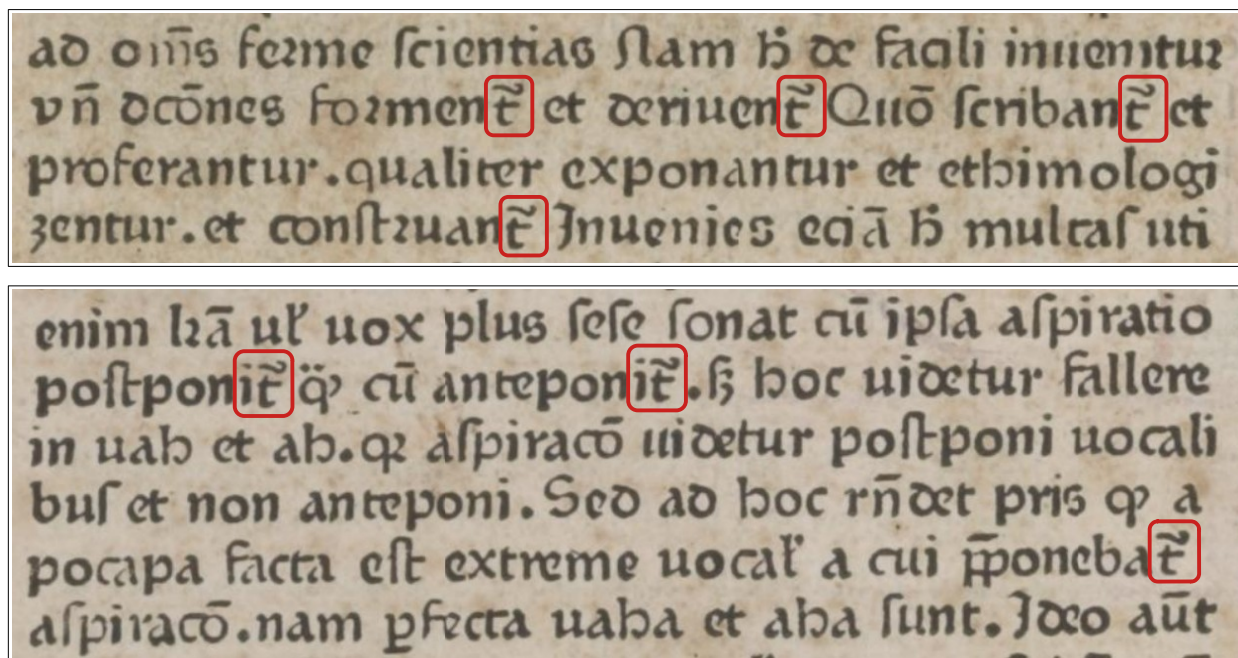Figure 22. Bień (2021: 34–35). Latin combining *ur* in *accipit^ur* from the 16th-century *Orthographia seu modus recte scribendi et legendi Polonicum idioma quam utilissimus* (Stanisław Zaborowski, 1514). The form is close to the Vietnamese apex seen in Figure 6.

Figure 23. Balbi (1460: 1–2 [of text; p. 9–10 of LOC pagination]). Baldi is a primary source for combining *ur* in L2/06-027, which requested it. The form is close to the Vietnamese apex seen in Figure 6.