蹺蹭蹪婆胡春香
Theo gót chân bà Hồ Xuân Hương
In the footsteps of the Queen of Nôm poetry, Hồ Xuân Hương

Di sản Hồ Xuân Hương đối với dân tộc và nhân loại
[National and International Heritage of Hồ Xuân Hương]
Ban Vận động UNESCO vinh danh nữ sĩ Hồ Xuân Hương
& Hội đồng Gia tộc Họ Hồ Việt Nam

by Ngô Thanh Nhàn
Linguistic String Project, New York University, and
Nôm Studies, *Center for Vietnamese Philosophy, Culture and Society*, Temple University
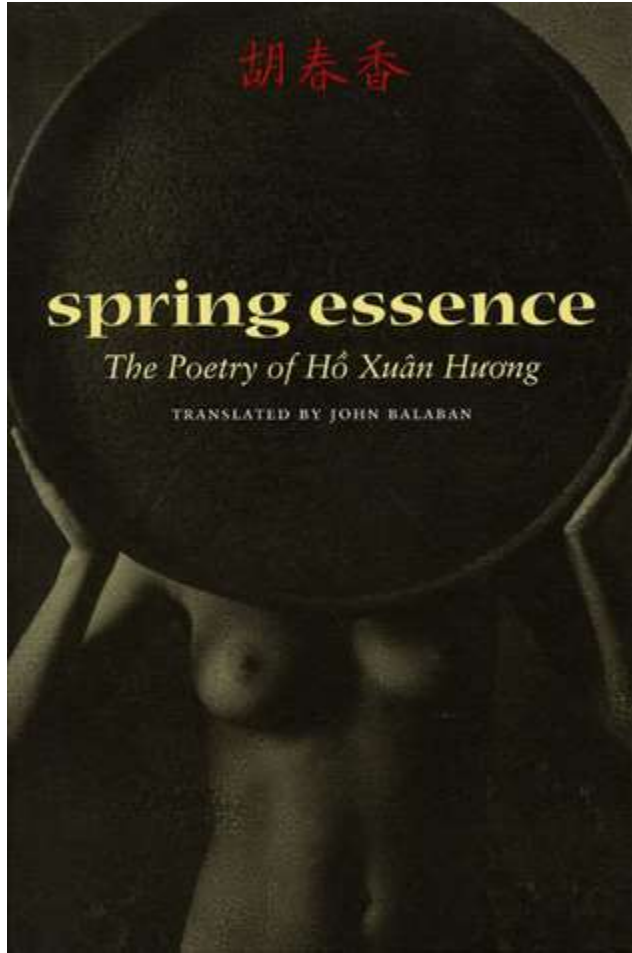March 15, 2022

Introduction

It is an unexpected pleasure that I have the honor of being invited to write an article on Hồ Xuân Hương for this collection. It is a *duyên nợ* "debt of fate" that I have always been in the footsteps of Hồ Xuân Hương, the Queen of Nôm poetry.

It was not by accident that I suggested a Nôm version of her poems upon hearing of an English translation of Hồ Xuân Hương poems by John Balaban. He asked me to help proof the *quốc ngữ* version. I accepted but said that, without a Nôm version, I was not sure how to proof the *quốc ngữ* poems. I was committed to this project, but I had no idea that Hồ Xuân Hương was much larger than life.

The task of presenting a Nôm version of this poetry book required background research in addition to some knowledge of the Nôm version closest to the original. I needed to determine, whether a selected poem was Hồ Xuân Hương's or written by an impersonator and the best reading of Nôm ideograms of the poem in *quốc ngữ* script. At the end, the

Figure 1: Cover page of *Spring Essence*

project could confidently present the poems in three scripts: Nôm, a *quốc ngữ* transliteration, and an English translation.

At the time, I had a list of about 5,220 Hán and Nôm proper ideograms in dot matrix images with associated *quốc ngữ* readings.[1] From 1992 to 1996, I and two colleagues, James Đỗ Bá Phước and Nguyễn Hoàng, had prepared Nôm ideograms to enter Unicode standards with Chinese, Japanese, and Korean scripts. I did not yet have a Nôm True Type font. It was my intention to use this chance to create a Nôm True Type font.

At the time, I only had the work of Maurice Durand, and the 春香遺稿 *Xuân Hương di cảo* [Loose manuscript of Xuân Hương] in 國音詩選 *Quốc âm thi tuyển* [Selected poetry in national language][2] as sources for Nôm ideograms. They were far from adequate, so was the ideogram repertoire I had prepared. All reference sources would be fully listed in the References section

---

[1] Institute of Linguistics, 1976, *Bảng tra chữ Nôm* [Nôm ideogram lookup table] became Tiêu chuẩn Việt Nam, TCVN 6056:1995. Nguyễn Quang Xỉ and Vũ Văn Kính, 1971, *Tự điển chữ Nôm* [A dictionary of Nom ideograms] became Tiêu chuẩn Việt Nam, TCVN 5773:1993.

[2] A copy can be found in National Library of Vietnam, index R.97, 18 sheets, 36 poems in Nôm and *quốc ngữ* (23 poems in *thất ngôn bát cú* and 13 poems in *tứ tuyệt*).

of the project to show our weaknesses. I painfully learned the limits of Nôm dictionary making, in both the north and south of Vietnam: they all professed to collect ideograms from formal, normative famous literary works, while lacking expressions of Vietnamese daily life. My preparations of ideograms in Unicode and *Tiêu chuẩn Việt Nam* "Vietnamese standards" were far from adequate. I then wrote a research paper on the nature of chữ Nôm, based on the 49 poems of the *Spring Essence* project.[3] But I trusted social sciences and the public electronic and web documentations to help gradually approach better versions when more sources were available.

I will use font 喃那宋體筋 *Nôm Na Tổng thể mảnh* "light Song font" for this article, produced by the Nôm Na Group in Hà Nội starting in 2002. At present we have a repertoire of 20,592 unique ideograms and 32,315 *quốc ngữ* readings in current UniHan version 13.0.0 (March 2020, with a repertoire of 88,889 CJKV ideograms, and 8,301 unique Nôm ideograms).

Technical production of *Spring Essence*

James Đỗ Bá Phước offered me the use of his Vietnamese typeface *Vsibon*, redesigned from *Sabon*, for my *quốc ngữ* presentation to match with Copper Canyon's publisher's font for its poetry books. I could then just concentrate on developing a Nôm True Type font.

I created the handwriting using 49 sheets, size 4"x7" wherefrom the final project version is 1"x2⅜" or a tenth of the original size. I used PageMakeer 6.5 in several new Mac Powerbook G3's at NYU Computer Science Department to create fonts for the book. In practical terms, I turned my handwriting to SVG (*scalable vector graphic*) images, imported them into Fontographers and generated a True Type font. It withstood the Copper Canyon proofing process for the final version of the book. John Balaban's 49 poems from Maurice Durand, had 33 poems in *thơ Đường Thất ngôn bát cú,* "T'ang eight line, seven-syllable regulated poem," (i.e. 8x7x33=1,848 ideograms in total), and 16 in *thơ Đường Thất ngôn tứ tuyệt* "T'ang four line, seven-syllable regulated poem," (i.e. 4x7x16=448 ideograms). The poetry book has a total size of 2,296 ideograms, not counting the poem title ideograms. See Figure 2.

---

[3] Nhàn, Ngô Thanh. *A review of dictionary indexing and lookup methods for ideographic scripts*.1998.
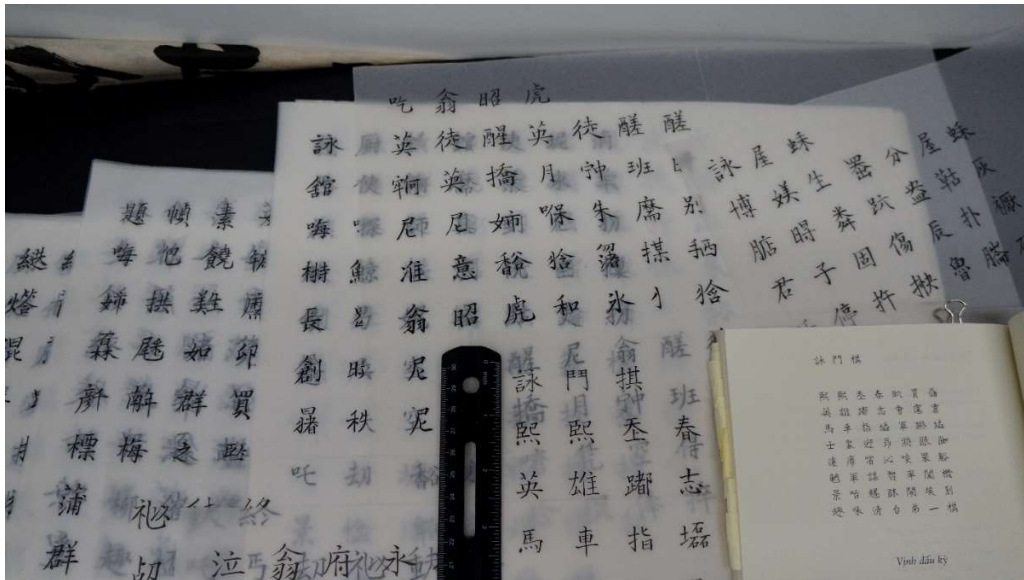
Figure 2: The draft handwriting: left, large handwriting originals, and, lower right, the final print in True Type font.

There were already 1,775 Nôm proper and 3,455 Hán ideograms encoded in Unicode before 1996, which I thought would be enough to use for all the Nôm poems of Hồ Xuân Hương for public use. It turned out to be far from enough. I had to study further, how to deal with the shortcomings, which turned into a paper on building ideogram dictionary entries and searching for the First International Conference on Vietnamese Studies.[4] Later, to increase the Nôm repertoire, in the following year, 2000, I accepted the job of turning Father Anthony Trần Văn Kiệm's life-long work to a fourth edition of 觧讀喃吧漢越 *Giúp đọc Nôm và Hán Việt* "An aid to reading Nôm and Hán Việt [ideograms]" for a new open source. I had to find a Vietnamese font for chữ Nôm. I was fortunate that Prof. Dr. of Science Nguyễn Quang Hồng from *Viện Nghiên cứu Hán Nôm* "Hán Nôm Research Institute" provided me a copy of 安子山陳朝禪宗指南傳心國語行 *Yên tử Sơn Trần Triều thiền tông chỉ nam truyền tâm quốc ngữ hạnh*, "Mount Yên Tử, Trần Dynasty, Essence of Zen Heart Sutra in National Script[5]." It is a Vietnamese *Trúc lâm* "Bamboo Forest" Zen collection by Venerable Chân Nguyên in 1745, carved on jackfruit woodblocks, and printed on *giấy dó* "mulberry paper". This helped us to design a Vietnamese Nôm Song light font called *Nôm Na Tống*.

In doing this work, I have learned several lessons:

(1) **To aim for a complete Nôm repertoire is impossible.** Unicode misconstrues a 'codepoint' which is a 'character' [like the Latin letter 'a'] to equal an 'ideogram.' An East Asian ideogram (Chinese, Japanese, Korean, Nôm), in fact, is a linear combination of 'characters' stacks neatly in an imaginary square. In Vietnamese Nôm spelling, this becomes 八刀分米粉 *bát đao phân mễ phấn* "eight knives chop the rice into powder," as a young woman named 粉 *Phấn* introduced herself to her future husband. It is a linear recursive binary combination, that is, two ideograms form a new ideogram with a

---

[4] Ref. Nhàn, *op.cit.* 1998.
[5] Chữ Nôm then was the national script.

combinatory rule, *trái trước, phải sau, trên trước, dưới sau, ngoài trước, trong sau* [the left part first, then the right; the top part first, then the bottom; the outside part first, then the inside], and stacking to fit inside a uniform imaginary square. This rule, simple and straightforward, generates infinite ideograms. In 1992, we defined an *orthographic unit*, which is a 'character' and a 'codepoint', like "a" and "â" of the Vietnamese *quốc ngữ* alphabet, and an *orthographic element*, which is a *chữ*, like "bấm," formed by orthographic units in spelling order: linearly from left to right for *quốc ngữ*,[6] and linear recursive binary composition for *chữ Nôm*. As such, 八 *bát* "eight" over 刀 *đao* "knife" makes 分 *phân* "divide," and again, the second recursive combination, 米 *mễ* "rice" before 分 *phân* "divide" makes 粉 *phấn* "powder." This recursive binary combination of the orthographic units: 八, 刀, 米,… while simple, is limitless. Unicode has yet to exhaust the UniHan repertoire after more than 30 years. Hồ Xuân Hương showed us this vernacular way of thinking while creating new ideograms for her poetry.

Figure 3: The first page of 禪宗本行 *Thiền tông bản hạnh* "Essential sutra of the Zen Sect".

(2) **Nôm has never been standardized**. To me, this is a blessing. Nôm ideograms reflected the writer's dialect and ideolect. For examples, the two ideograms 得, 𠊚 are both pronounced *người* with exactly the same meaning (ideogram 人/radical 亻 *nhân* + 导 *ngại*); four ideograms 字,𡨸, 𡦂, 苧 are pronounced *chữ* with the same meaning "script."

---

[6] Đỗ, James Bá Phước, Ngô, Thanh Nhàn, and Nguyễn, Hoàng, "A proposal for standard Vietnamese character encodings in a unified text processing framework" 1/1992.

Thus, we are still far from a complete repertoire if we have consulted centralized libraries and known archives alone.

(3) **Dictionaries are incomplete**. This is a corollary of (1) above. Even more so, ideogram dictionary compilers considered themselves Confucian teachers. The dictionaries were compiled from fine literary works, not from people's daily life. Therefore, dictionaries have been incomplete because their sources of ideograms are limited.

(4) **The goal is Nôm Studies**, not a complete list of ideograms. A complete Nôm repertoire is important (we were still barely at the very start), and the goal is to "understand" the documents, which requires a multi-disciplinary field of Nôm Studies, such as philosophy, literature, arts, library science, mathematics, sociology, history, geography, entomology, etc.

(5) **Hồ Xuân Hương existed.** She "was well known, as a talented woman in literature and politics."[7] She was beautiful, was a talented poet in *thơ đường* "T'ang Dynasty" regulated poetry. She later tried and bested a new style of Nôm poetry, a second poem hidden in the first, featuring the banality of men's power over women.

(6) **She asserted the view of a normal Vietnamese woman in literature.** Literature was the realm of and dominated by men. Almost by accident, three famous poetry epics appeared about the same time were about women, 征婦吟曲 *Chinh phụ ngâm khúc* "Lament of the soldier's wife," 宮怨吟曲 *Cung oán ngâm khúc* "Complaint of a Palace Maid," and 傳翹 *Truyện Kiều* "The Tale of Kiều." They were written by men. She, like other poetesses, Bà Huyện Thanh Quan, Đoàn Thị Điểm,... brought daily life using Vietnamese *láy* reduplications, loose compounds, and *nói lái* rhyme switching word game[8], ... the much richer vernacular expressions, into Vietnamese literature.

The Hán poetry of Hồ Xuân Hương

Hồ Xuân Hương showed her talents in writing in *Thơ Đường thất ngôn bát cú* "T'ang poem of eight seven-syllable verses" and *thơ Đường thất ngôn tứ tuyệt* "T'ang poem of four seven-syllable verses." They are written in the Chinese strictly regulated poetry style, favored by Vietnamese literary scholars and mandarins. The first couplet is called *đề* "exposition." The second couplet is called *thực* "description," the third, *luận* "comment", and the fourth, *kết* "conclusion". Two verses in each couplet must be of the same syntactic structure. Each syllable must be of the same part of speech with its counterpart in the next verse. They must contrast in meaning, syllable by syllable. The first verse sets the *vần* rhyme for the entire poem, whose second, fourth, sixth and eighth syllables must follow.

The Vietnamese differentiates 㗂,諎 *tiếng* and 𡦂,𡨸,�字 *chữ*. 㗂,諎 *tiếng* is a spoken syllable. 𡦂,𡨸,�字 *chữ* is a written syllable. Each syllable has three elements: initial consonant, a 韻,韵 *vần* "rhyme" and a tone. In the following poem, *vần* shows *iên/yên* the last syllable of verses 1,2,4,6, and 8 in the transliterated version: *tiền, nguyên, yên, thiên,* and *duyên,* but they do not seem to

---

[7] Balaban, p. 7.
[8] Linguistic game of children, and secret language of traders, … like pig Latin.

show in corresponding Hán poem (i.e. 前,源,煙,天,緣 one needs to look up for them in a rhyme ideogram dictionary).

I again started reading 瑠香記 *Lưu hương ký* "Memories of the lingering scent". Hồ Xuân Hương asked a young talented Phan Tốn Phong, who had a crush on her, to write the preface for it. From the 31 poems in *đường luật*, we learn that Hồ Xuân Hương, alias Hồ Thị Mai, was a young beautiful woman. Watch the language of Hồ Xuân Hương used in a *xướng* dare-to-respond to her betrothed lover Mai Sơn Phủ when he was still alive. (He later died young and unmarried):

Hán original	***Phiên âm* Transliteration** (all in lower case)
南浦相知十載**前**	nam phố tương tri thập tải **tiền**
桃花猶似隔仙**源**	đào hoa do tự cách tiên **nguyên**
北城相見今年景	bắc thành tương kiến kim niên cảnh
柳絮還應帶野**煙**	liễu nhứ hoàn ưng đới dã **yên**
半點春情梅欲雪	bán điểm xuân tình mai dục tuyết
萬重離思水如**天**	vạn trùng ly tứ thuỷ như **thiên**
有時蘭室挑銀燭	hữu thời lan thất khiêu ngân chúc
共話江湖訂宿**緣**	cộng thoại giang hồ đính túc **duyên**
Cf. Hoàng[9], p.41.	

English translation:[10]

Ten years ago, Southern Gate, that **spree**

Under the blossoms of a cherry **tree**.

Today, we met again at Northern Citadel,

Willows, like cloaking mist, kept us **free**.

Apricot blooms fluttering in Spring's snowy breeze,

Like ten thousand separations in Heaven's **sea**.

Lighting silver lamps at my Orchid House,

We shared stories and promised predestined **intimacy**.

She expressed her love to Mai Sơn Phủ. Note her sharp and romantic *thực* "description" couplet: "Today, we met again at Northern Citadel, willows, like cloaking mist, kept us free."

---

[9] Hoàng, Bích Ngọc, 2003.

[10] By a stroke of pure luck, I received this translation, and others below, from Lady Borton, who co-edited the bi-lingual anthology of 1,000 years of Vietnamese women's poetry. I am truly grateful. Personal communication. Lady Borton has received three honorary doctorates for her fifty-five years of work with all sides during and after the American War in Việt Nam. She has published three books about Việt Nam, including the only Boat People account written from inside the exodus and the first foreigner's account of post-war life among ordinary Vietnamese. She has also published six annotated book-length translations from Vietnamese to English and twenty bi-lingual booklets on Vietnamese culture. Lady organized the first two co-editions between Vietnamese and American publishers.

Pre-publication draft…

Another poem in her 香亭古月詩集 *Hương đình Cổ Nguyệt thi tập* "Poetry collection of scent pavilion of the Old Moon," [i.e. "Old Moon": 古 *cổ* "old" and 月 *nguyệt* "moon" spells 胡 *hồ*, her last name] and her eight poems in 涂山八詠 *Đồ Sơn bát vịnh* "Eight chants of Đồ Sơn". The poem 題鎮國寺 *Đề Trấn Quốc Tự* "On Trấn Quốc Temple," has totally different tone from her later Nôm poem.

Hán original	***Phiên âm* Transliteration** (all lower case)
莊臨誰是景中人	trang lâm thuỳ thị cảnh trung nhân
細細清風扇扇熏	tế tế thanh phong phiến phiến huân
水月波籠蓮掛茁	thuỷ nguyệt ba lung liên quải chuất
霜煙寶座露連雲	sương yên bảo toạ lộ liên vân
洗空塵慮花含彩	tẩy không trần lự hoa hàm thái
喚醒迷途草亦春	hoán tỉnh mê đồ thảo diệc xuân
到景令人回首問	đáo cảnh linh nhân hồi thủ vấn
東南拂袖雁成群	đông nam phất tụ nhạn thành quần.

Cf. Hoàng, *op.cit.* p. 197.

English translation:[11]
Anyone here, do honor this pagoda's **wellspring**,
The fragrant breath of gentle breezes **blowing.**
The moon, water, waves, the lotus blossoms,
The spiced smoke of lit incense **wafting**.

Let the shy flowers cleanse your worries,
Revive your affection for herons in **spring**.
This scene reminds everyone: Turn and watch
The Southeast beckoning swallows on the **wing**.

Again, note her signature vivid *thực* "descriptive" couplet, "The moon, water, waves, the lotus blossoms, the spiced smoke of lit incense wafting." ... reveal her respect for the country founders' temple...

In contrast to the above a later Nôm poem of hers in *Spring Essence*, entitled 昂廚鎮國 *Vịnh chùa Trấn Quốc* "Trấn Quốc Temple" she visited later. The two *thực* couplet again shows her sharp and imaginative observation. It shows the place being neglected:

Nôm original	***Phiên âm* Transliteration** (all lower case)
鎮國行宮鞊𣰹油 、	trấn quốc hành cung cỏ dãi dầu
憫念故國擬𪮏疴	chạnh niềm cố quốc nghĩ mà đau
沒座蓮鎖哹香御 、	một toà sen toả hơi hương ngự,
舾式靈封攝襖朝	năm thức mây phong nếp áo chầu
浺泣廢興瞜沕嗹	sóng lớp phế hưng coi vẫn dộn
鐘回今古潮彊毛	chuông hồi kim cổ lắng càng mau

---
[11] Lady Borton, *op. cit.* personal communication.

得罟景廟兜兜佐 、   người xưa cảnh cũ đâu đâu tá,
窖謹魚台屢濁頭   khéo ngẩn ngơ thay lũ trọc đầu.
Cf. Balaban, *op.cit.* p.92.

English translation:[12]
Trấn Quốc Pagoda torments, its grass bedraggled,
Its melancholy reminiscent of our ancient country.
A lotus bloom emanates royalty's lingering scent
Of robes showing five faded, embroidered clouds.

Generations of dethroned kings seem noisily busy,
Their high-pitched bell's sound dissipates quickly.
Where might we discover elders from antiquity
Cleverly superseded by stupid, bald-headed men?

Nôm poetry of Hồ Xuân Hương

Let's analyze 詠螺蚌 *Vịnh Ốc Nhồi* "The snail." *Ốc nhồi* is the black apple snail, found everywhere in Vietnam, and is one of the most favorite snacks of Vietnamese, rich or poor. She writes that, like the fate of all women in her time, they lived every day and night among the *cỏ hôi* "smelly grass", which is, in popular term, *cỏ cứt lợn* "pig's shit grass." Still they were great enjoyments for children, women, men, scholars, kings and street bumps, alike.

The poem describes exactly how a black apple snail is born and lives. And to eat it properly, people must use a needle or a toothpick, to first take off the operculum. If this fails, one pokes open the tail, on the other end, and sucks the meat out.

It also meant to describe how men treated women in her time. The men would never realize their distasteful behaviors towards women. She smartly revealed their shortcomings as gentleman.

Nôm original	***Phiên âm* Transliteration** (all lower case)
博媄生𠺙分螺蚌	bác mẹ sinh ra phận ốc nhồi
腅晬璘跙盎鮯灰	đêm ngày lăn lóc đám cỏ hôi
君子固傷辰扑襜	quân tử có thương thì bóc yếm
吀停扴揪魯𦢞碎	xin đừng ngó ngoáy lỗ trôn tôi.

Cf. Balaban, *op.cit.* p. 38.

English translation:[13]
My parents birthed me, a voluptuous snail,
Rolling about, smelly grass buffs my wales.
Yes, remove my halter: Delight, tasty tales!
But please, gentlemen, don't fondle my tail.

---

[12] Lady Borton, *op. cit.* Personal communication.
[13] Lady Borton, *op. cit.* Personal communication.

The poem has 28 syllables, among them 4 are Sino-Vietnamese, *bác*, *sinh*, *quân* and *tử*, and nativization, *phân>phận*, *cố>có*, *thìn>thì*, *đình>đừng*. The rest are proper Nôm ideograms, representing unique Vietnamese sounds.

Let's see how each Nôm ideogram is contructed. Interestingly they are generally composed of two ideograms. In the Position column, each line has the first number being the line number, and the second number, the ideogram position from left to right on such line (in old Nôm documents, from top down, right to left, each verse usually occupies one column). For example, 4.2 in the Position column represents line 4 ideogram 2:

Position	Ideogram	Ideogram Construction	Quốc ngữ reading
1.1	博	十 thập + 尃 phô	**bác**, mác, vác, bạc
1.2	媄	女 nữ + 美 mỹ	**mẹ**, mệ
1.3	生	生 sinh	**sinh**, sanh, siêng, xênh, xinh
1.4	𡦂	罗 la + 出 xuất	**ra**
1.5	分	八 bát + 刀 đao	phân, **phận**, phần
1.6	蝺	虫 trùng + 屋 ốc	**ốc**
1.7	蟍	虫 trùng + 耒 lỗi	giòi, dòi, ruồi, **nhồi**, ròi
2.1	𣎀	月 nguyệt + 店 điếm	**đêm**
2.2	𣈙	日 nhật + 寻 ngại	**ngày**
2.3	𡑩	土 thổ + 舛 lân	**lăn**, sân
2.4	跞	⻊ túc + 立 lập	lộp, rập, sụp, **lóc**
2.5	盎	央 ương + 皿 mãnh	ang, áng, **đám**, ăng
2.6	𦹳	草 thảo + 古 cổ	**cỏ**
2.7	灰	厂 hán + 火 hoả	**hôi**, hoi, khói, khôi, vôi, hui
3.1	君	尹 doãn + 口 khẩu	**quân**, vua
3.2	子	了 liễu + 一	**tử**, tí, tờ
3.3	固	囗 vi + 古 cổ	**có**, cố, cúa
3.4	慯	忄 tâm + 𠂯昜 *thang	**thương**
3.5	辰	辰 thần	thần, thìn, **thì**
3.6	扑	扌 thủ + 卜 bốc	bốc, buốt, phốc, vộc, vục, buộc, phác, **bóc**
3.7	襂	衤 y + 厭 yếm	**yếm**, ướm
4.1	吁	口 khẩu + 千 thiên	xiên, **xin**
4.2	停	亻 nhân + 亭 đình	đình, dành, đành, **đừng**, dừng, rành
4.3	扝	扌 thủ + 午 ngọ	**ngó**, ngỏ
4.4	揆	扌 thủ + 快 khoái	**ngoáy**, khuấy
4.5	魯	魚 ngư + 日 nhật	**lỗ**, lõ, nhỗ, sỗ
4.6	膦	月 nhục + 崙 lôn	lồn, **trôn**
4.7	碎	石 thạch + 卒 tốt	toái, thôi, thỏi, **tôi**, toả, tui, tủi

There are several observations from Nôm table above:

— Nôm-to-*quốc ngữ* is one-to-many (and vice versa), shown by the *quốc ngữ reading* column. Each Nôm ideogram corresponds to one *quốc ngữ* reading, and vice versa. Transliterators today will have to select their own preferred syllable of each ideogram in the poem. Thus, what we call *phiên âm* "transliteration", is actually a translation, personal and preferential.

— The readings are, nevertheless, definitely phonetic. That is, Vietnamese proper, and Vietnamized, in the form of Hán-Việt, not 'pictogram'. Indeed, one element of the binary composition represents the sound. If we look at Table 1. 19 of the 28 ideograms have this binary structure <classifier>+<sound> while the ordering, front-to-back, top-down, or outside-to-inside, is optional. This structure occupites 67.85%, a third of the ideograms in the poem: 媄 mỹ>mẹ [*mỹ* in column **Construction** > *mẹ* in column **Quốc ngữ Reading**], 羅 la>ra, 蝸 ốc>ốc, 蛛 lỗi>nhồi, 脂 điếm>đêm, 晦 ngại>ngày, 獜 lân>lăn, 踃 lập>lóc, 鈷 cổ>cỏ, 固 cố>có, 慯 thang>thương, 扑 bốc>bóc, 襺 yếm>yếm, 吁 thiên>xin, 停 đình>đừng, 扜 ngọ>ngó, 揆 khoái>ngoáy, 膈 lôn>lồn, 碎 tốt>tôi. This structure is useful for Vietnamese etymologists to use the sound elements of Nôm ideograms to verify the reconstruction method in historical linguistics.

— One approach is to read the ideograms with the best *đường luật* rhyme rules, based on the tone features, *bằng* vs *trắc* "level vs contour", and *cao* vs *thấp* "high vs low register"[14]. This helps to select a transliterated syllable that adheres to the T'ang poetic regulations of rhymes and tone registers. For example, the syllable 4.7 碎 by T'ang poetic rule should be level, one should reject contour syllables, *toái, thỏi, toả, tủi,* and *tui* and should choose *thôi* or *tôi* with rhyme *-ôi* of the poem. We may not arrive at a single appropriate syllable in the end: such as ideograms 1.7 蛛, 3.6 扑, 4.4 揆 và 4.6 膈.

— As a result, we may still have more than one ideogram that matches the rhyme rules, for example, ideograms 1.7 蛛, 3.6 扑, 4.4 揆 and 4.6 膈.

— Another approach to select the best match is to apply Vietnamese grammatical rules, consistency and parts of speech, to hopefully further reduce the readings down to one.

— One may add a rule of literary "decency" reading, which may not be the intention of Hồ Xuân Hương: This removes 膈 *lồn* reading and leaves 膈 *trôn* reading in 4.6, where the meanings are the same. We could choose *trôn* over *lồn* because *trôn* is the popular name of the tip apex of a black apple snail.

— Ideogram 3.2 子 *tử* is a Chinese radical, but not to Hồ Xuân Hương. She sees it as a woman with a horizontal growth in her belly. In the poem 詠得薦荒 *Vịnh người chửa hoang* "The pregnant maiden,"[15] line 4, 分了𥘷它抎涅昂 *phận liễu sao đà đẩy nét ngang* "a maiden, how could you grow sideways." She equates visually the ideogram 子 *tử* "baby, child" being a 了 *liễu* "maiden" with a visual 一 horizontal bar, not *nhất* "one".

---

[14] Cf. Nhàn, 1984, op.cit. Appendices, Charts, p.2, and http://mlp.cs.nyu.edu/folk.arts/linguistics/syllables.php.

[15] Cf. Balaban, *op. cit.*, p. 52.

The poem shows her view of a normal woman towards men and their behaviors. She describes a woman's view of a popular enjoyment of snail snacks, while creating new Nôm ideograms to represent the normal daily activities of the Vietnamese people. Note that she did not cite any *điển tích Trung hoa* Chinese classic reference, a literary showoff by the men literati. She also hinted at the preferred binary structure of the language of the Vietnamese people.

A great demonstration of her efforts is shown in her famous 詠豁割據 *Vịnh hang Cắc cớ* "Ill-fated Cave" where she used vernacular reduplications, word plays and compounds, bold and italicized below.

Nôm original	***Phiên âm* Transliteration** (all lower case)
歪坦生𡀭矴没篸	trời đất sinh ra đá một chòm
�föld台𦟐吼函歆	nứt làm hai mảnh **hỏm hòm hom**
技培蓑木諸喧呬	kẽ hầm rêu mọc **trơ toen miệng**
篭𨤼椿嘹撫乏𠀧	luồng gió thông reo vỗ **phập phòm**
喫渃有情涑境澟	giọt nước hữu tình rơi **lõm bõm**
昆塘無岸瞵焙焙	con đường vô ngạn tối **om om**
唏埃攔矴才穿鑿	khen ai **đẽo đá** tài **xuyên tạc**
窖吘馨器𤭪几旺	khéo **hở hang** ra lắm kẻ nhòm.

Cf. Balaban *op.cit.* p. 88.

English translation:[16]
Heaven and Earth birthed this sheltering grove,
Splitting granite to create a sunken cove.
Moist, comely moss conceals its narrow gate
Although gusts of pulsing wind still penetrate.
Moisture seeps from this dark, enticing vale,
No guardrail accompanies its deep, dim trail.
Praise whoever chiseled an entrance so shy,
If indecently exposed, many men would pry!

This poem exposes the brilliant use of vernacular *trơ toen miệng*, reduplications *hõm hòm hom, phập phòm, lõm bõm, om om*, and *hở hang*, ... and *nói lái* "rhyme switching word play," *đẽo đá > đã đéo*, as well as the natural ambiguity of compounds, creating two meanings at once, e.g. *xuyên tạc* means both "carve through" and "distort," and *hở hang* means "to leave the cave open wide" and "revealing [dress]."

Even if I believed I was prepared with a large Nôm and Hán ideogram repertoire from the two most famous dictionaries of *Bảng tra chữ Nôm* from Hà Nội and *Tự điển chữ Nôm* from Sàigòn, (both became Vietnamese Nôm standards, TCVN 5773 and TCVN 6056), I still sorely lacked vocabulary to deal with Nôm poetry by Hồ Xuân Hương. Thus, I found the handwriting option for this project necessary.

I am truly humble to be working on Hồ Xuân Hương's poems! They help me to understand the necessity, the scope and importance of Nôm Study. Most importantly, the necessity to focus on

---

[16] Lady Borton, *op. cit.* Personal communication.

language of normal daily lives of the people. While writing poetry, Hồ Xuân Hương showed the poverty of learned men, while expanding Nôm ideogram vocabulary of vernacular language and elevated it into literature. Indeed! The Vietnamese have lived and spoken their language for millenia. They created and expanded Nôm ideograms for at least one millenium in their daily activities. Their use of Nôm to reflect the spoken language has always been extensive. On my part, this project meant for me to focus on further expansion of Nôm ideogram repertoire for public use. It also meant the expansion of sources beyond libraries and archives inside and outside of Vietnam, but also from all record keepings within Việt Nam of families, clans, villages, pagodas, temples, religious institutions, … and folk arts. The poetry of Hồ Xuân Hương had changed the directions of my research.

References

Balaban, John. Ed. and tr. 2000. *Spring Essence: The poetry of Hồ Xuân Hương*. Port Townsend, Washington: Copper Canyon Press. 138pp.

Durant, Maurice. 1968. *L'oeuvre de la poétesse vietnamienne Hô-Xuân-Huong*. École française d'Extrême-Orient. Paris: Adrien-Maisonneuve. 208pp.

Hoàng, Thị Bích Ngọc. 2003. *Hồ Xuân Hương: Con người – Tư tưởng – Tác phẩm* [Hồ Xuân Hương: The author – The thought – The poetry]. Nhà xuất bản Văn hóa Thông tin. 750pp.

Đỗ, James Bá Phước, Ngô, Thanh Nhàn, and Nguyễn, Hoàng, "A proposal for standard Vietnamese character encodings in a unified text processing framework", *Computer Standards & Interfaces 14* (1/1992):3-12. Cf. https://cs.nyu.edu/~nhan/csni_14_1992.pdf.

Ngô, Thanh Nhàn. 1984. *The syllabeme and word formation in Vietnamese*. Ph.D. Thesis. New York University.

Ngô, Thanh Nhàn, Đỗ, James, và Nguyễn, Hoàng. 1992. Introduction to Nôm – A Vietnamese script for computer character encoding, presented at the Han Subcommittee, Unicode Consortium: Xerox Corporation, Palo Alto, California, May 7, 1992. 68 pp. Cf. https://cs.nyu.edu/~nhan/draft_npct_19920322.pdf.

Ngô, Thanh Nhàn. 1998. *A review of dictionary indexing and lookup methods for ideographic scripts*. Presented at the First International Conference on Vietnamese Studies, by Hanoi National University and National Center for Social Sciences and Humanity. Hà Nội. Việt Nam. July 14-17, 1998. Cf. https://cs.nyu.edu/~nhan/vsic98.pdf.

Ngô, Thanh Nhàn. 2001. "Đơn vị chính tả và các đặc điểm của tiếng Việt: chữ quốc ngữ, hệ latinh, chữ Nôm, hệ biểu ý và Unicode/ISO IEC 10646," [Orthographic unit, characteristics of Vietnamese language: romanized script, ideographic Nôm and Unicode/ISO IEC 10646] for *The Vietnamese Unicode/ISO 10646 Committee (VUIC)*, July 1, 2001. Xem https://www.cs.nyu.edu/~nhan/Orthographic_unit.pdf.

Viện Ngôn ngữ học [Institute of Linguistics]. 1976. *Bảng tra chữ Nôm* [Nôm ideogram lookup table]. Hà Nội: The Social Sciences Publishing House. 426pp.

Nguyễn Quang Xỉ and Vũ Văn Kính. 1971. *Tự điển chữ Nôm* [A dictionary of Nôm ideograms]. Sàigòn: Trung tâm Học liệu.

Tiêu chuẩn Việt Nam. 1993. TCVN 5773: *Information Technology – The Nôm 16-bit character standard code set for information interchange – Chữ Nôm Việt* (1.775 ideograms). Hanoi. Vietnam. 59 pp.

Tiêu chuẩn Việt Nam. 1995. TCVN 6056: *Information Technology – The Nôm 16-bit character standard code set for information interchange – Chữ Nôm Hán* (3.455 ideograms). Hanoi. Vietnam. 62 pp.

Trần, Văn Kiệm, F. Anthony. 2004. *Giúp đọc Nôm và Hán Việt* 𡨸讀喃吧漢越 "An aid to the pronunciation of Nôm and Hán Ideograms". Đà Nẵng: Đà Nẵng City and the Vietnamese Nôm Preservation Foundation. xiii pp. Introductions and Legends. 906 pp.