

SINO-VIETNAMESE MORPHEMES: THEIR GRAMMATICAL BEHAVIOUR AND IMPLICATIONS FOR SEMANTICS AND PHONETICS

ỨNG XỬ NGỮ PHÁP CỦA HÌNH VỊ HÁN VIỆT VÀ HỆ QUẢ CỦA NÓ VỀ NGỮ NGHĨA VÀ NGỮ ÂM

Hoang Dzung^{1,*}

DOI: <http://doi.org/10.57001/huivh5804.2025.033>

ABSTRACT

This paper shows that Sino-Vietnamese morphemes combine not only with other Sino-Vietnamese morphemes but also with non-Sino-Vietnamese morphemes, such as native Vietnamese and even Indo-European elements, particularly French loanwords. This phenomenon is not unique to Vietnamese: similar hybrid words can also be found in other Sino-xenic languages (Japanese and Korean) as well as Indo-European languages. In noun phrases where the Sino-Vietnamese morpheme serves as the head, the Chinese grammatical rule [modifier + head] tends to predominate, but in some cases, Vietnamese follows its native grammatical structure [head + modifier]. This grammatical behaviour has semantic consequences. Indeed, the criterion [\pm interchangeable in all contexts], used to classify synonyms into “absolute synonyms” and “non-absolute synonyms,” proves ineffective, because free Sino-Vietnamese morphemes can combine with both other Sino-Vietnamese morphemes and native Vietnamese ones, whereas their native Vietnamese synonyms generally cannot combine with Sino-Vietnamese morphemes. The differences in combinability of Sino-Vietnamese morphemes lead to semantic consequences in connection with phonology: Sino-Vietnamese morphemes with very low productivity tend to be assimilated into (near-)homophonous morphemes – whether Sino-Vietnamese or non-Sino-Vietnamese – with higher productivity.

Keywords: Grammatical behaviour, hybrid word, loanword, Sino-Vietnamese, Sino-xenic.

TÓM TẮT

Bài báo cho thấy hình vị Hán Việt không chỉ kết hợp với hình vị Hán Việt khác, mà còn với hình vị phi Hán Việt như thuần Việt và thậm chí cả các hình vị Ấn - Âu, đặc biệt là mượn tiếng Pháp. Đây không phải là hiện tượng riêng của tiếng Việt: những từ lai tương tự cũng có thể được tìm thấy trong các ngôn ngữ Sino-xenic khác (Nhật, Hàn) và các ngôn ngữ Ấn - Âu. Trong danh ngữ có hình vị Hán Việt đóng vai trò chính, chiếm ưu thế là quy tắc ngữ pháp tiếng Hán [phụ + chính] nhưng cũng có khi tiếng Việt lại áp dụng quy tắc ngữ pháp tiếng Việt [chính + phụ]. Đặc điểm ngữ pháp này đưa đến hệ quả về ngữ nghĩa. Quả vậy, tiêu chí [\pm có thể hoán đổi trong mọi ngữ cảnh] dùng để phân loại từ đồng nghĩa thành “đồng nghĩa tuyệt đối” và “đồng nghĩa không tuyệt đối” là không hiệu quả vì các hình vị Hán Việt tự do có thể kết hợp với cả hình vị Hán Việt khác lẫn hình vị thuần Việt, trong khi các từ đồng nghĩa thuần Việt của chúng thường không thể kết hợp với hình vị Hán Việt. Sự khác biệt về khả năng kết hợp của hình vị Hán Việt đưa đến hệ quả về phương diện ngữ nghĩa trong mối liên quan với ngữ âm: những hình vị Hán Việt có sức sản sinh quá yếu sẽ có xu hướng bị đồng nhất vào những hình vị (Hán Việt hay phi Hán Việt) đồng âm hay gần âm có sức sản sinh mạnh.

Từ khóa: Ứng xử ngữ pháp, từ lai, từ mượn, Hán Việt, Sino-xenic.

¹HCMC Linguists' Society, Vietnam

*Email: dunghoang07@gmail.com

Received: 18/01/2025

Revised: 22/02/2025

Accepted: 27/02/2025

1. INTRODUCTION

After Khúc Thừa Dụ's proclamation of autonomy (906 CE) and the victory of Ngô Quyền (938 CE), the Chinese

language in Vietnam began to separate from the Chinese language in China. Now following its own path, it gradually became phonetically, semantically and

grammatically distinctive, culminating in Sino-Vietnamese (see [13]) (*To be distinguished from Old Sino-Vietnamese and Vietnamized Sino-Vietnamese* [12]). Due to its significant role in the history of Vietnamese language, Sino-Vietnamese has long attracted the attention of experts [7, 12]. However, most studies on Sino-Vietnamese have focused on its phonetics or lexicon, and have been less interested in its grammatical behaviour.

This paper focuses on investigating the combinatory possibilities of Sino-Vietnamese morphemes, from which their semantic, and sometimes also phonetic, impacts can be examined. Our interest is thus in the syntax-semantics-phonetics interface of Sino-Vietnamese elements.

2. GRAMMAR OF SINO-MORPHEME COMBINATIONS

A common view among non-experts is that the only acceptable morphemic combination is [Sino-Vietnamese + Sino-Vietnamese] (*An example is the following claim: "Tặc [賊 literally 'thief, pirate, invader'] is chữ nho 'Chinese element(s)', hence it should be combined with another chữ nho (lâm tặc; lâm 林 meaning 'forest'); combining chữ nho with chữ nôm 'Vietnamese element(s)' (e.g., cát tặc 'sand thief') is not accurate (sa tặc is correct [sa 沙 'sand'])."* [3]. This is a typical manifestation of the tendency to venerate linguistic purity, while overlooking the existence of the considerable amount of data from modern Vietnamese that attests otherwise.

Indeed, it is easy to find [non-Sino-Vietnamese + Sino-Vietnamese] noun phrases in which the head is final (*Học trò [literally '(school) pupil'] is an exception, as the pure Vietnamese morpheme is the phrasal head yet is word-final as if it were pure Sino-Vietnamese. The cause perhaps originated in history: trò and đồ 徒 'learner' are cognates, but the former is Old Sino-Vietnamese. In other words, học trò was used before the Tang dynasty. Note that Old Sino-Vietnamese and Vietnamized Sino-Vietnamese are both classified as 'pure Vietnamese' from the standpoint of synchronic etymology.*

The non-Sino-Vietnamese morpheme may be pure Vietnamese (henceforth simply Vietnamese), as seen in *bếp trưởng* 'head chef' (Vietnamese *bếp* 'cook' + Sino-Vietnamese *trưởng* 長 'chief'), *chuyển trưởng* 'line manager (in a factory)' (Vietnamese *chuyển* '[factory] line' + Sino-Vietnamese 'chief'), *cửa hàng trưởng* 'store manager' (Vietnamese *cửa hàng* 'store' + Sino-Vietnamese *trưởng* 'chief'), *máy trưởng* 'chief engineer' (Vietnamese *máy* 'machine' + Sino-Vietnamese *trưởng* 'chief'), *nhóm trưởng* 'group leader' (Vietnamese *nhóm*

'group' + Sino-Vietnamese *trưởng* 'chief'), *quầy trưởng* 'kiosk chief' (Vietnamese *quầy* 'kiosk' + Sino-Vietnamese *trưởng* 'chief'), *xóm trưởng* 'head of a neighborhood' (Vietnamese *xóm* 'neighborhood' + Sino-Vietnamese *trưởng* 'chief'),... *xoắn trùng / xoắn khuẩn* 'spirochaete bacteria' (Vietnamese *xoắn* 'spiral' + Sino-Vietnamese *trùng* 蟲 'strain' / *khuẩn* 菌 'bacteria'), *phẩy trùng / phẩy khuẩn* 'vibrio bacteria' (Vietnamese *phẩy* 'comma' + Sino-Vietnamese *trùng* 'strain' / *khuẩn* 'bacteria').

While some [Sino-Vietnamese + Sino-Vietnamese] noun phrases such as *hải tặc* 'pirate' (*hải* 海 'sea' + *tặc* 賊 'invader'), *không tặc* 'hijacker' (*không* 空 'air' + *tặc* 'invader'), *son tặc* 'mountain highwaymen' (*son* 山 'mountain' + *tặc* 'invader') and *thuỷ tặc* 'river pirate' (*thuỷ* 水 'water' + *tặc* 'invader') consist of the morpheme *tặc* ('invader') with the preceding element indicating the spatial domain in which the invader acts (*hải tặc* 'invader [acting at] sea', *không tặc* 'invader [acting in the] air (= in an aeroplane)', *son tặc* "invader [acting on a] mountain", *thuỷ tặc* "invader [acting on] water (= on a river)'), we have recently witnessed the emergence of [Vietnamese + Sino-Vietnamese] noun phrases, whose initial morphemes no longer depict a domain of activity, but instead an object of theft: (*Note that the combination [Sino-Vietnamese + tặc] can exist simultaneously with the [pure Vietnamese + tặc] variant: cf. cá tặc versus ngư tặc 'fish thief' (ngư 魚 'fish'); cát tặc versus sa tặc 'sand thief' (Người lao động 6/4/2013); chó tặc versus cẩu tặc 'dog thief' (cẩu 狗 'dog') (Đất Việt 20/09/2013); đất tặc versus thổ tặc 'land robber' (thổ 土 'land') (dantri.com 27/02/2011); kiểng tặc versus cảnh tặc 'bonsai plant thief' (cảnh 境 literally 'scenery') (Nông nghiệp Việt Nam, 24/4/2012), trâu tặc versus ngưu tặc 'buffalo thief' (ngưu 牛 'buffalo'); vàng tặc versus kim tặc 'gold thief' (kim 金 'gold') (Sài Gòn tiếp thị 8/12/2011). The usage [pure Vietnamese + tặc] is still not convincing for many people: "Why can [they] say cát tặc 'sand thief', but not giặc cát while we have giặc đói 'hunger crisis', giặc dốt 'illiterate crisis', giặc châu chấu 'grasshopper crisis', and so forth; why do [they] use đất tặc 'land robber' but not (bọn) cướp đất [...] I believe that we should not be tolerant here; we should instead provide guidance."* (An Chi, personal communication) *bò tặc* 'cow thief', *cà tặc* 'coffee bean thief' (*cà* is an abbreviation of *cà phê* 'coffee'), *chim tặc* 'bird thief' (Công an TP Đà Nẵng 15/8/2013), *dầu tặc* 'oil thief', *dế tặc* 'cellphone thief' (*dế* literally 'cricket', a slang for cellphone), *dưa tặc* 'melon thief', *dừa tặc* 'coconut thief' (Lao động 08/02/2013), *đá tặc* 'stone thief', *gà tặc* 'chicken thief', *gạch tặc* 'brick thief', *mai tặc* 'apricot blossom thief', *mèo tặc* 'cat thief', *ngheo tặc*

'clam thief' (*Tuổi trẻ* 25/4/2012), *rùa tặc* 'turtle thief', *sứa tặc* 'thief of *Dalbergia bouruana* *gagu* trees', *than tặc* 'coal thief', *thiếc tặc* 'tin thief' (*Người lao động*, 11/8/2013), *tôm tặc* 'shrimp thief', *xe tặc* 'vehicle thief', and so on (*It may be necessary to include a special case in which the first element depicts a body part under attack: mông tặc* (Vietnamese *mông* 'buttocks' + Sino-Vietnamese *tặc* 'invader') (*vnexpress.net* 25/5/2005) *is a person who sticks kim 'needle' (kim tặc, Cao Phi Yến [3] – homonymous to kim tặc, whose meaning is equivalent to vàng tặc 'gold robber/thief') into women's buttocks!*).

There are also cases in which the first element indicates the instrument used by the actor: *bùn tặc* (Vietnamese *bùn* 'mud' + Sino-Vietnamese *tặc* 'invader'), *rác tặc* (Vietnamese *rác* 'garbage' + Sino-Vietnamese *tặc* 'invader'), *bụi tặc* (Vietnamese *bụi* 'dust' + Sino-Vietnamese *tặc* 'invader') are not those who steal mud, garbage, or dust, but rather who pollute the environment by spilling mud, garbage, or producing dust; *khoan tặc* (Vietnamese *khoan* 'drill' + Sino-Vietnamese *tặc* 'invader') is a person who damages the aesthetic of public spaces by illegally advertising with the slogan *Khoan, cắt bê tông* ('Concrete cutting and drilling'); *cào tặc* (Vietnamese *cào* 'rake' + Sino-Vietnamese *tặc* 'invader') is a person who illegally dredges fish; *câu tặc* (Vietnamese *câu* 'angling' + Sino-Vietnamese *tặc* 'invader') is a fisherman who uses a rod to steal fish owned by others; *lửa tặc* (*Hoả tặc* 'fire breakouts', the [Sino-Vietnamese + *tặc*] counterpart of *lửa tặc*, does not have the same meaning. In *hoả tặc* [Sino-Vietnamese *hoả* 火 'fire' + Sino-Vietnamese *tặc* 'invader'], the morpheme *tặc* is used metaphorically to mean a danger: "68% of the warehouses have been completely equipped with lightning rod sensors which readily cope with *hoả tặc* in the dry season." (*Quân đội Nhân dân* 13/10/2008)(*Giáo dục*, 4/3/2011) is a person who assaults others by burning them (Examples of [X + *tặc*] combinations here are from Lê Minh Hoàng [8] unless indicated otherwise).

Yet another case of [Vietnamese + Sino-Vietnamese] noun phrases is the [Vietnamese + *quán* 館 literally 'shop'] combination, which appears more widespread every day across Vietnam in restaurant names: *Ba Miền quán* (Vietnamese *ba* 'three' + Vietnamese *miền* 'region' + Sino-Vietnamese *quán* 'shop'), *Cây Sung quán* (Vietnamese *cây* 'tree' + Vietnamese *sung* '*Ficus racemosa*' + Sino-Vietnamese *quán* 'shop'), *Cây Đa quán* (Vietnamese *cây* 'tree' + Vietnamese *đa* '*Ficus bengalensis*' + Sino-Vietnamese *quán* 'shop'), *Lò Đất quán* (Vietnamese *lò* 'stove' + Vietnamese *đất* 'earth, soil, land' + Sino-

Vietnamese *quán* 'shop'), *Miền Trung quán* (Vietnamese *miền* 'region' + Sino-Vietnamese *trung* 中 'central [Vietnam]' + Sino-Vietnamese *quán* 'shop'), *Sông Trăng quán* (Vietnamese *sông* 'river' + Vietnamese *trăng* 'moon' + Sino-Vietnamese *quán* 'shop'). This construction adds nuance to restaurants' names and makes an impression on the customers.

The non-Sino-Vietnamese morpheme may also be Indo-European. Many of these combinations have become popular: *băng trưởng* 'gangleader' (*dantri.com.vn* 29/3/2012) (*băng* < French *bande* 'gang'), *ca trưởng* 'shift manager' (*ca* < French *quart* 'quarter'), *kíp trưởng* 'team manager' (*kíp* < French *équipe* 'team') (*Quán* 'shop' and *trưởng* 'chief' are both free morphemes. As such, in addition to the [non-Sino-Vietnamese + Vietnamese] combination, they can form [Sino-Vietnamese + non-Sino-Vietnamese] constructions (e.g. *nhóm trưởng / trưởng nhóm* 'head of a group', *ca trưởng / trưởng ca* 'shift manager'; *Cây Sung Quán / Quán Cây Sung* 'a restaurant name'); *cồn kế* 'alcoholmeter' (*cồn* < French *alcool* 'alcohol'; *kế* 計, literally 'count') (The [Vietnamese + Sino-Vietnamese] equivalent here is *rượu kế* (Vietnamese *rượu* 'wine' + Sino-Vietnamese *kế* 'meter')), *vôn kế* 'voltmeter', *ampe kế* 'Ampere meter (ammeter)', *ohm kế* 'ohmmeter'; *cà phê tặc* 'coffee bean thief' (*cà phê* < French *café*), *cao su tặc* 'rubber thief' (*cao su* < French *caoutchouc*); *game thủ* 'gamer' (*thủ* 手, literally 'hand'), *cơ thủ* 'billiards player' (*cơ* < French *queue* 'cue')(Monosyllabic Indo-European words are perceived by Vietnamese speakers as (pure) Vietnamese, as long as they do not include exotic phonemes (or written characters). For example, while Vietnamese usually consider *a-xít* (< *acide*), *a-mi-ăng* (< *amiante*), *a-míp* (< *amibe*), *bê-rê* (< *beret*), *bu-gi* (< *bougie*), *rô-nê-ô* (< *ronéo*) and *ô-liu* (< *olive*) as 'Western words', *tem* (< *timbre* 'stamp'), *xăng* (< *essence* 'gasoline'), *kem* (< *crème* 'cream'), *xi* (< *cire* 'wax'), *săm* (< *chambre* (à air) 'tire tube', *lốp* (< *enveloppe*) are easily thought of as Vietnamese, though all were borrowed from French. The case of *cà phê* 'coffee' is exceptional: due to homonymy, numerous Vietnamese speakers believe the *cà* in this word to be the same as the *cà* of *cà chua* 'tomato', i.e. coffee is a type of *cà*. Evidence is that people living in the Vietnamese Central Highlands (a region known for coffee plantations) often use *cà* to refer to coffee in a clear context); *logic học* 'logic' (*học* 學 'study'), *robot học* 'robotics', *topo học* 'topology', *virus học* 'virology' (In these listed [modifier + head] noun phrases, *học* is a nominal bound morpheme. It

can be a free morpheme only in the role of a verb. Hence, it is impossible for [Sino-Vietnamese + non-Sino-Vietnamese] synonyms to exist, unlike the case of quán or trường).

Thus far, [non-Sino-Vietnamese + Sino-Vietnamese] noun phrases are limited to the following few morphemes: *kế* 'meter', *quán* 'shop', *học* 'study', *khuẩn* 'bacteria' / *trùng* 'strain', *tặc* 'invader', *thủ* 'hand', *trưởng* 'chief'.

It can be observed that in all examples above, the phrasal head is always Sino-Vietnamese and comes last, following the structure of Chinese noun phrases, while the preceding modifier may be non-Sino-Vietnamese. Hence, here, Vietnamese use Chinese noun phrase structure rather than Chinese words. This is a case of borrowing grammatical structure and has so far been little noticed (Discussing grammatical borrowing of Sino-Vietnamese, M. J. Alves [1, 2], for instance, focuses only on functional words with Chinese origins, numerals, and comparatives, rather than on structural borrowing itself). Notably, some have pushed this usage much further, as seen in *nữ nhà văn* 'female prose writer' (Sino-Vietnamese *nữ* 女 'female' + Vietnamese *nhà* 'person' + Sino-Vietnamese *văn* 文 'literature') (*Thanh niên*, 17/3/2013) or *nữ cây bút* 'female writer' (Sino-Vietnamese *nữ* 'female' + Vietnamese *cây* [classifier for tree-like objects] + Sino-Vietnamese *bút* 筆 'pen') (*Công an*, 24/3/2013). These examples show a structural mixture between Vietnamese noun phrases ([head + modifier]: *nhà* + *văn*, *cây* + *bút*) and Chinese noun phrases ([modifier + head]: *nữ* + [*nhà văn*], *nữ* + [*cây bút*]). This combination has thus far been viewed as unacceptable.

Nevertheless, not all noun phrases containing Sino-Vietnamese elements show a [modifier + head] structure in which the head is occupied by the Sino-Vietnamese morpheme. We also find cases with [head + modifier] word order (the typical noun phrase structure in Vietnamese), where the modifier position is filled by a bound Sino-Vietnamese morpheme while the head is Vietnamese: (If they are free Sino-Vietnamese morphemes (like *học* in *bàn học* 'desk'), they are in fact counted as pure Vietnamese) *thuỷ* 'water' in *nước thuỷ* (Vietnamese *nước* 'water' + Sino-Vietnamese *thuỷ* 'water') (the thin layer of mercury in a mirror), *lính thuỷ* 'marine' (Vietnamese *lính* 'soldier' + Sino-Vietnamese *thuỷ* 'water'), *máy thuỷ* 'motor (for boats)' (Vietnamese *máy* 'machine' + Sino-Vietnamese *thuỷ* 'water'), *tàu thuỷ* 'ship' (Vietnamese *tàu* 'vessel' + Sino-Vietnamese *thuỷ* 'water'); *hoả* in *dầu hoả* 'kerosene' (Vietnamese *dầu* 'oil' + Sino-Vietnamese *hoả*

火 'fire'), *tàu hoả* 'train' (Vietnamese *tàu* 'vessel' + Sino-Vietnamese *hoả* 'fire').

The phenomenon is not without a trace in the Vietnamese verb phrase. The [X + *hoá* 化 'to transform'] construction which conforms to Chinese grammatical structure, but in which X is non-Sino-Vietnamese, has become popular: *cứng hoá* 'to harden' (Vietnamese *cứng* 'hard' + Sino-Vietnamese *hoá* 'to transform'), *lành mạnh hoá* 'to purify (metaphorically)' (Vietnamese *lành mạnh* 'wholesome, sound' + Sino-Vietnamese *hoá* 'to transform'), *mềm hóa* 'to soften' (Vietnamese *mềm* 'soft' + Sino-Vietnamese *hoá* 'to transform'), *môi hoá* 'to labialize' (Vietnamese *môi* 'lip' + Sino-Vietnamese *hoá* 'to transform'), *mũi hoá* 'to nasalize' (Vietnamese *mũi* 'nose' + Sino-Vietnamese *hoá* 'to transform'), *ngói hoá* 'to tile [a roof, etc]' (Vietnamese *ngói* 'tile' + Sino-Vietnamese *hoá* 'to transform'), *vôi hoá* 'to calcify' (Vietnamese *vôi* 'lime (material)' + Sino-Vietnamese *hoá* 'to transform'), *xát hoá* 'to fricativize' (Vietnamese *xát* 'to rub' + Sino-Vietnamese *hoá* 'to transform'). Multi-element examples include *a xít hoá* (*a xít* < French *acide* 'acid'), *bê tông hoá* (*bê tông* < French *béton* 'concrete'), *xi măng hoá* (*xi măng* < French *ciment* 'cement').

The examples mentioned above are all in subordinate compound structures. In coordinate compound structures, we can also observe bound Sino-Vietnamese morphemes combining with their Vietnamese synonyms or near-synonyms in noun phrases and verb/adjective ones: *binh lính* (兵) *lính* 'soldiers', *ẩm thấp* (濕) *thấp* = *ẩm* 'damp', *bao* (包) *gồm* 'to consist of', *bao* (包) *bọc* 'to wrap/cover', *biến đổi* (變) *đổi* 'to change', *bồi đắp* (培) *đắp* 'to accrete', *kỳ lạ* (奇) *lạ* 'mysterious/strange', *mũ mào* (帽) *mào* or *mạo* [4] and [17] give the *mạo* reading, while Phan Văn Các [15] accepts both *mạo* and *mào* = *mũ* 'hat'), *nuôi dưỡng* (養) 'to raise', *sinh* (生) *sống* 'to live', among others.

3. HYBRID WORDS IN SINO-XENIC LANGUAGES AND BEYOND

Hybrid words are not confined to Vietnamese; they can also be found in other Sino-xenic languages (*The term is due to Martin [9] to refer to languages with a lexicon showing widespread, systematic Chinese borrowing; -xenic originates from Greek *xénos* 'foreign'*) such as Japanese or Korean.

In Japanese, (pure) Japanese morphemes may combine with non-Japanese elements. Sino-Japanese (*on'yomi*) examples include [pure Japanese + Sino-Japanese] 場所 *basho* 'place', 湯桶 *yutō* 'hot water

bucket'; [Sino-Japanese + pure Japanese] 金 色 *kin'iro* "yellow", 重 箱 *jūbako* 'food box'; [pure Japanese + Sino-Japanese + Sino-Japanese] 合 気 道 *aikidō* 'aikido'. Indo-European examples include [Japanese + Indo-European] カラオケ *karaoke* (Japanese *kara* 'empty' + English *oke* < *orche(stra)*); [Indo-European + Japanese] シャボン玉 *shabondama* 'soap bubble' (*shabon* < Portuguese *sabão* 'soap' + Japanese *dama* < *tama* 'bubble'). There are also cases of [Indo-European + Sino-Japanese] as in ハンチング帽 *hanchingu-bō* 'hunting cap' (*hanchingu* < English *hunting* + Sino-Japanese *bō* 'cap').

Korean is similar. Sino-Korean + (pure) Korean constructions are seen, for instance, in 공부하다 *kongpuhata* 'to learn' (Sino-Korean *kongu* 'to learn' + Korean *hata* 'to do'), 행복하다 *hengpokhata* 'happy' (Sino-Korean *hengpok* 'happy' + Korean *hata* 'to do'), 다행이다 *tahengita* 'lucky' (Sino-Korean *taheng* 'relief' + Korean *ita* 'to be'). Examples of [Indo-European + pure Korean] include 메리추석 *Merry Chuseok* 'Happy Mid-autumn' (English *merry* + Korean *Chuseok* 'Mid-autumn'), 굿밤 *gutbam* 'goodnight' (*gut* < English *good* + Korean *bam* 'night'), 방울토마토 *bangultomato* 'cherry tomato' (Korean *bangul* 'drop (of water)' + English *tomato*), 비닐봉투 *binilbongtu* 'plastic bag' (*binil* < English *vinyl* + Korean *bongtu* 'bag').

Hybrid words are also common in Indo-European languages. English, for instance, has many words of Greek and Latin origin, which may sometimes show hybrid forms: cf. *homosexual* (< Greek *homos* + Latin *sexus*), *hypercorrection* (< Greek *hyper* + Latin *correction*), *metadata* (< Greek *meta* + Latin *data*); *sociology* (< Latin *socius* + Greek *logos*), *television* (< Greek *tēle* + Latin *visio*). As in Vietnamese, English allows the parallel existence of both hybrid and non-hybrid synonyms, cf. *aquaphobia* (< Latin *aqua* + Greek *phobos*) versus *hydrophobia* (< Greek *hydro* + Greek *phobos*), *monolingual* (< Greek *monos* + Latin *lingua*) versus *unilingual* (< Latin *uni* + Latin *lingua*), and *divalent* (< Greek *di-* + Latin *valentem*) versus *bivalent* (< Latin *bi-* + Latin *valentem*).

Thus, hybrid words are a universal phenomenon. It is fair to say that where there is language contact, we find hybridization; it is as hard to imagine a language without hybrid words as it is to think of a language without contact with other languages.

4. TWO TYPES OF GRAMMATICAL BEHAVIOUR IN VIETNAMESE HYBRID WORDS

Why do hybrid words assert themselves so strongly puristic criticisms, no matter how furious, fail to stop their

advance? With respect to structure, in order to explain their existence, it is crucial to keep in mind the following well-known quote from Ferdinand de Saussure: "La synchronie ne connaît qu'une perspective, celle des sujets parlants [...]" [16].

Indeed, native speakers do not need to know the history of their language. For them, the only reality is synchronic. From the perspective of synchronic etymology, Vietnamese perceive free Sino-Vietnamese morphemes in [head + modifier] noun phrases as if they were pure Vietnamese morphemes, therefore there is no difference between [Vietnamese + Sino-Vietnamese] and [Vietnamese + Vietnamese] combinations. Examples include bàn học 'desk' (pure Vietnamese bàn 'table' + Sino-Vietnamese học 學 'to learn, to study') versus bàn ăn 'dining table' (pure Vietnamese bàn 'table' + Vietnamese ăn 'eat'); hòn ngọc 'gem' (Vietnamese hòn 'roundish (object)' + Sino-Vietnamese ngọc 玉 'gem') versus hòn đất 'a ball of soil' (Vietnamese hòn 'roundish (object)' + Vietnamese đất 'earth, soil, land'); máy ảnh 'camera' (Vietnamese máy 'machine' + Sino-Vietnamese ảnh 影 'image') versus máy may 'sewing machine' (Vietnamese máy 'machine' + Vietnamese may 'to sew'); người bệnh 'patient' (Vietnamese người 'human' + Sino-Vietnamese bệnh 病 'illness') versus người lành 'healthy person' (Vietnamese người 'human' + Vietnamese lành 'healthy / recovered'); nỗi khổ 'misery' (Vietnamese nỗi 'feeling' + Sino-Vietnamese khổ 苦 'misery') versus nỗi buồn 'grief' (Vietnamese nỗi 'feeling' + Vietnamese buồn 'sad'); tàu điện 'electric train' (pure Vietnamese tàu 'vessel' + Sino-Vietnamese điện 電 'electric') or tàu hỏa (pure Vietnamese tàu 'vessel' + Sino-Vietnamese hỏa 火 'fire') versus tàu lửa 'train' (Vietnamese tàu 'vessel' + pure Vietnamese lửa 'fire'); thuốc bổ 'tonic' (Vietnamese thuốc 'medicine' + Sino-Vietnamese bổ 補 'tonic') versus thuốc ho 'cough medicine' (pure Vietnamese thuốc 'medicine' + Vietnamese ho 'cough'); vòng đồng 'copper bracelet' (Vietnamese vòng 'bracelet' + Sino-Vietnamese đồng 銅 'copper') versus vòng đá 'stone bracelet' (Vietnamese vòng 'bracelet' + Vietnamese đá 'stone'); vết thương 'wound' (Vietnamese vết 'mark' + Vietnamese thương 傷 'hurt') versus vết mổ 'surgical scar' (Vietnamese vết 'mark' + Vietnamese mổ 'to operate (surgically)') (The following claim by Jae Jung Song [6] regarding Sino-Korean is also applicable to these Sino-Vietnamese words: "Arguing that Sino-Korean words are loanwords is like arguing that English words based on Latin and Greek elements are loanwords. Native English speakers, unless they are

etymologists or have specialist knowledge, don't realize the Latin or Greek origin of words like *client*, *library*, *essence*, *dissolve*, *idiosyncrasy*, *democracy* and *history*; even if they were aware of their origins, they would hesitate to call them loanwords, simply because they have been part of the English vocabulary for so long, and thus entrenched so firmly in the English language that it will be purely academic whether to call them loanwords or not. Similarly, Sino-Korean words are so firmly established in Korean that it may not be appropriate to view them as loanwords, albeit based on Chinese characters").

The combination between a free Vietnamese morpheme and a bound Vietnamese morpheme is not rare in Vietnamese, e.g., *cá chình* 'lamprey' (free *cá* 'fish'), *cá ngát* 'a type of catfish', *chó má* 'villain, rogue' (free *chó* 'dog'), *chợ búa* 'markets' (free *chợ* 'market'), *diều hâu* 'hawk' (free *diều* 'hawk'), *dưa hấu* 'watermelon' (free *dưa* 'melon'), *đậu nành* 'soybean' (free *đậu* 'bean'), *lúa hẻo* (free *lúa* 'paddy'), and so on. The [free Vietnamese + bound Sino-Vietnamese] construction like *thuỷ* 'water' in *tàu thuỷ* 'ship' is thus unexceptional.

It is worth mentioning those noun phrases which obey the Chinese structural configuration of [modifier + head], but in which the modifier is non-Sino-Vietnamese. Contrary to the other two types, this is a case of Vietnamese following Chinese grammatical rules; in other words, borrowing Chinese syntactic structures. This behavior enriches the Vietnamese, offering additional means of different rhetorical nuance (cf. *Sông Trăng Quán* versus *Quán Sông Trăng*, whose different nuances can be imperfectly approximated by 'Café Moon River' versus 'Moon River Café' in that order (marked versus unmarked, with the French head-modifier order giving the more exotic feel in English)). In other cases, this combination yields a phrase that can be better understood: for 'viscometer', Vietnamese uses *nhớt kế* (Vietnamese *nhớt* 'viscous' + Sino-Vietnamese *kế* 'meter'), rather than Chinese 黏滯計; for 'spirochaete bacteria', *xoắn khuẩn* (Vietnamese *xoắn* 'spiral' + Sino-Vietnamese *khuẩn* 菌 'bacteria'), rather than 螺旋菌; for 'vibrio bacteria', *phẩy khuẩn* (Vietnamese *phẩy* 'comma' + Sino-Vietnamese *khuẩn* 菌 'bacteria'), rather than 弧菌. Alternatively, it helps avoid potential misunderstandings caused by the combination of exclusively Sino-Vietnamese elements; cf. *ẩm kế* versus *thấp kế* 濕計 'hygrometer', the latter easily mistaken for a tool to measure height. This is because in

Vietnamese, *thấp* with the meaning 'damp' appears only in the word *ẩm thấp*, while as a free morpheme, it means 'short' (While it is possible to use *máy đo độ nhớt* (literally 'machine measure degree oil') 'viscometer' and *máy đo độ ẩm* ('machine measure degree damp') 'hygrometer', it is not coincidental that *nhớt kế*, *ẩm kế* are far popular (and appear in many dictionaries, such as Hoàng Phê's *Từ điển tiếng Việt* [5]). The advantage of the latter forms is that their combination is 'tighter' due to the [modifier + head] construction; they are hence more suitable as scientific terminology).

5. THE COMBINATORY ABILITY OF SINO-VIETNAMESE ELEMENTS FROM A SEMANTIC PERSPECTIVE CONCERNING PHONETICS

Does the combinatorial ability of Sino-Vietnamese elements analyzed above relate to semantics and/or phonetics?

Some have attempted to resolve semantic issues using the combinatorial behaviour of these morphemes. For example, the textbook *Vietnamese* for grade 5 includes the distinction between "absolute synonyms" and "non-absolute synonyms" based on their [\pm interchangeable] ability; examples for the former are *hổ*, *cọp*, *hùm* 'tiger' [11]. It is easy to show that these examples are unconvincing: In fact, it is possible to say *con hổ*, *con cọp*, or *con hùm* 'a/the tiger' (Vietnamese *con* [classifier used for animals] + Sino-Vietnamese *hổ* / Vietnamese *cọp* or *hùm*), but only *hổ phụ sinh hổ tử* 'like father, like son' (literally 'a tiger dad begets a tiger son'), and not **cọp/hùm phụ sinh cọp/hùm tử*, is acceptable. The reason is simple: *hổ* is the only Sino-Vietnamese morpheme among the three, therefore it is the only one that can combine with other Sino-Vietnamese morphemes in noun phrases having Chinese phrase structure. The classification criterion is thus not supported with *hổ*, *cọp*, and *hùm* (*Theoretically, a classification criterion as such is also unacceptable. Language is a large but parsimonious system: if two synonyms can replace each other in all contexts, one of them is redundant and will be eliminated. In other words, a combinatory difference between two synonyms is a reason for their parallel existence*).

The difference in the combinatory ability of Sino-Vietnamese morphemes leads to some semantic consequences, constrained by phonetics: Vietnamese speakers tend to merge Sino-Vietnamese morphemes with lower productivity into (near-) homophonous morphemes with higher productivity.

Sometimes this involves the merger of two Sino-Vietnamese morphemes. For example, *trữ* 貯 as a free morpheme means 'to contain', and all of its combinations in Vietnamese (*dự trữ* 'to save, reserve', *tàng trữ* 'to store', *tích trữ* 'to stockpile, to stash away', *tồn trữ* 'to store, storage', *trữ kim* 'savings', *trữ lượng* 'backlog, stockpile', and so on) involve this meaning except for *trữ tình* 'lyric(al)' 抒情, in which *trữ* 抒 means 'to express'; thus *trữ tình* is often understood as 'to be full of sentiment' [10]. Another example is the following line of the poet Bà Huyện Thanh Quan: "*Gác mái, ngư ông về viễn phố*". Here *phố* 浦 has the meaning 'riverbank, shore' (*viễn phố* 遠浦 means 'distant riverbank, shore'). Yet Vietnamese has only the free morpheme *phố* 鋪, originally meaning 'shop', developing the sense 'a street with houses and shops along its sides', as seen in *phố cổ* 'ancient street (and its surroundings)' and *phố Hàng Đào*, a popular Hanoi shopping street. It is, therefore, easy for *phố* 浦 to be mistaken for *phố* 鋪 [10].

Elsewhere, we observe the merger of a Sino-Vietnamese and a pure Vietnamese morpheme. For example, the Sino-Vietnamese morpheme *yếu* 要 'important' (as seen in *hiểm yếu* 'dangerous and strategic' (= difficult to be accessed), *trích yếu* 'summarize', *trọng yếu* 'important', *yếu nhân* 'important person') is homonymous with the Vietnamese morpheme *yếu* 'weak' (the antonym of *mạnh* 'strong'). As the former is a bound morpheme while the latter is a free one, with higher combinatory ability, *yếu điểm* is often misunderstood as 'a weak point', rather than 'an important point'. As a further example, *thăm quan* 'sightseeing' replaces *tham quan*. *Tham* 參 (in *tham quan*) meaning 'to attend' appears only as a bound morpheme (*tham chiến* 'to join a war', *tham chính* 'to take part in politics', *tham dự* 'to attend', *tham gia* 'to participate', *tham kiến* 'to meet') while the free morpheme *tham* 貪 means 'greed(y), to covet'. As it does not make any sense to assign *tham* in *tham quan* 'sightseeing' the meaning of the free morpheme, a near-homonym which yields a more reasonable meaning - *thăm* 'to visit' - is used (*Besides, tham quan is easy to be confused phonetically with thăm quan because both tham and thăm are not stressed, so that a and ă are unclearly pronounced*). Another case is *xán lạn* 燦爛 'bright and splendid', in which both *xán* and *lạn* mean 'bright, glowing'. Because *xán* is found only in this compound, it is merged with *sáng* 'bright', which is its near homonym and also its synonym. In this case, the two prototypical prerequisites for folk-etymological

reanalysis - sound and meaning - have acted to generate the more common expression *sáng lạn*.

Sometimes, we find a combination of the above types. For example, *cứu* with the meaning 'final' appears only in one Vietnamese compound, *cứu cánh* 究竟 '(last) refuge/recourse', while *cứu* 救 meaning 'help' appears not only in [Sino-Vietnamese + Sino-Vietnamese] combinations such as *cứu hộ* 'to rescue', *cứu tế* 'salvation', *cứu thế* 'redemptive', *cứu tinh* 'savior') but also exists as a free morpheme. Similarly, *cánh* 竟, which means 'final, at last', exists only in *cứu cánh*, while the pure Vietnamese morpheme *cánh* 'wing, petal, side', is a free morpheme. As a result, *cứu cánh* is frequently understood as 'assistant at one's side', even though its correct meaning is 'the final goal'.

6. CONCLUSION

The grammatical behaviour of Sino-Vietnamese morphemes nowadays does not correspond to their historical behaviour ([modifier + head] noun phrases following the pattern [non-Sino-Vietnamese + Sino-Vietnamese] did, in fact, appear in legends such as *Cây Gỗ Đại vương*, *Bến Đò Đại vương*, and *Cửa Sau Đại vương* [14]. However, most of these noun phrases are limited to proper nouns).

Investigation of the behaviour of Sino-Vietnamese morphemes helps us to understand better their combinatorial mechanisms in modern Vietnamese. Based on this investigation, we can make predictions about the development of the language. This also provides a basis on which to elucidate the mechanism of incorrect usage or misinterpretation of Sino-Vietnamese elements, rather than just to list individual errors.

The structural mixture between Vietnamese and Chinese and the semantic change of Sino-Vietnamese words under the pressure of pure Vietnamese should be accepted. From the viewpoint of synchronic etymology, it is necessary to count Sino-Vietnamese elements as part of the Vietnamese language despite their historical status as loanwords, similar to Latin and Greek etyma retained in modern Indo-European languages. This viewpoint will enable us to respect the linguistic facts of Vietnamese, as opposed to denying reality with the excuse of protecting the 'purity' of Vietnamese.

ACKNOWLEDGEMENTS

Many thanks are sent to An Chi, Masaaki Shimizu (Osaka University), Trần Đình Sử, and Trần Trọng Dương for their insightful input.

REFERENCES

- [1]. Alves M. J., "Sino-Vietnamese Grammatical Borrowing: An Overview," in: Matras, Yaron & Sakel, Jeanette, eds. *Grammatical Borrowing in Cross-Linguistic Perspective* (pp. 343-361). Berlin: Walter de Gruyter, 2007.
- [2]. Alves M. J., "Sino-Vietnamese Grammatical Vocabulary and Sociolinguistic Conditions for Borrowing," *Journal of the Southeast Asian Linguistics Society* 1, 1-9, 2009.
- [3]. Cao Phi Yến, "Lạm dụng các từ 'tặc', 'sạch' và 'xanh' [The overusage of 'tặc', 'sạch' and 'xanh'],", *Nhân dân*, 2010.
- [4]. Đào Duy Anh, *Hán Việt từ điển* [Sino-Vietnamese Dictionary]. Saigon: Trường Thi, 1957.
- [5]. Hoàng Phê (Ed), *Từ điển tiếng Việt*. NXB Đà Nẵng, 2003.
- [6]. Jae Jung Song, *The Korean Language: Structure, Use and Context*. New York: Routledge, 2005.
- [7]. Lê Đình Khẩn, *Từ vựng gốc Hán trong tiếng Việt* [Chinese lexicon in Vietnamese]. HCMC: Đại học Quốc gia TP HCM, 2002.
- [8]. Lê Minh Hoàng, "Phiếm luận về "... tặc" [Satire on tặc],", *Tuổi trẻ cười*, 423, 2011.
- [9]. Martin, S. E., *The Phonemes of Ancient Chinese*. Baltimore: American Oriental Society, 1953.
- [10]. Nguyễn Lân, *Từ điển từ và ngữ Hán Việt* [Dictionary on Sino-Vietnamese words and phrases]. Hanoi: Từ điển Bách khoa [Encyclopedic Dictionaries], 2002.
- [11]. Nguyễn Minh Thuyết (Ed), *Tiếng Việt 5* [Vietnamese Grade 5], vol.1. Hanoi: Giáo dục [Educational Press], 2006.
- [12]. Nguyễn Tài Cẩn, *Nguồn gốc và quá trình hình thành cách đọc Hán Việt* [The origins and the formation process of Sino-Vietnamese readings]. Hanoi: Khoa học Xã hội [Social Sciences Press], 1979.
- [13]. Nguyễn Tài Cẩn, "Thử phân kì lịch sử 12 thế kỉ của tiếng Việt [An attempt to dichotomize twelve centuries of Vietnamese linguistic theory],", *Tạp chí Ngôn ngữ* [Journal of Linguistics], 7-12, 1998.
- [14]. Nguyễn Thị Trang, "Bước đầu tìm hiểu hệ thống chữ Nôm ghi tên hiệu các vị thần thành hoàng làng [First steps in exploring the chữ Nôm writing system used in the names of village gods],", *Nghiên cứu chữ Nôm* (kỷ yếu Hội nghị chữ Nôm) [*Chữ Nôm Studies* (Proceedings of the Conference on Chữ Nôm)]. NXB Khoa học Xã hội [Social Sciences Press], 2006.
- [15]. Phan Văn Các, *Từ điển từ Hán Việt* [Sino-Vietnamese dictionary]. HCMC: NXB TP HCM [HCMC Press], 2003.
- [16]. Saussure F. de., *Cours de linguistique générale*. Paris: Payot, 1971.
- [17]. Thiều Châu, *Hán Việt từ điển* [Sino-Vietnamese dictionary]. HCMC: NXB TP HCM [HCMC Press], 2002.

THÔNG TIN CÁ NHÂN
Hoàng Dũng

Hội Ngôn ngữ học TP. HCM