PACLIC-26
Bali, November 8-10, 2012

Idiomaticity and Classical Traditions in Some East Asian Languages[1]

Benjamin K. Tsou
Research Centre on Linguistics and Language Information Sciences,
The Hong Kong Institute of Education
btsou99@gmail.com

1. Introduction

A mark of erudition in verbal communication is the use of idiomatic language employing metaphors and figurative speech. Its study is important not only for linguistic research but also for the study of language, rhetorics, literacy studies and cultural history, and the relationship amongst them.

The rhetorical distinction between literal and metaphorical meanings and so semantic and discoursal opacity often associated with idioms is universal. But the format of idioms can stand out, and the means by which the expressions are formed, often drawing on the use of notable objects or events relevant to the native society concerned are often culture bound. These objects and events can often be drawn from relatively closed sets.

Idioms are commonly used in metaphors and figurative speech in all languages and in daily communication. They have not only attracted the attention of specialists interested in language, rhetorics and literary studies (Black 1962, Makkai 1972, Xiang 1979), but even visiting national leaders to China from USA and Japan in recent years have cited them in their speeches. In the last few decades, several major areas associated with idioms and metaphors have become noticeable: (a) Syntax and Semantics, e.g. Chafe's well-known 1968 paper on syntactic decomposability issues of frozen idioms; (Katz and Postal (1963) and Jackendoff (1995)); (b) Cognitive studies, e.g. Gibbs (1980, 1985, 1987), Nippold et al. (1989), Zuo (2006), Zhang (1984); and (c) Cultural studies, e.g. Lakoff (1987) [gender], Tang (2007) [food related items], Nall (2008) [numbers], Fontecha and Catalan (2003) [animals], Liu (1984), Fan (2007) [color terms], Mo (2001) [Chinese culture and idioms]. There are also notable anthologies on the relevant approaches, e.g. Everaert (1989, 1992, and 1995).

We note that when some salient linguistic features are found to be shared across two languages, the question often arises as to whether their origin might be due to: (a) shared genetic affinity, or (b) borrowing across language boundaries. Furthermore, they could be also (c) universal features if shared by all other languages, or (d) typological linguistic features if shared by structurally similar natural languages, as well as (e) areal or regional features if they are found only in a particular geographical region. Moreover, they are not mutually exclusive.

[1] This research is supported by the Research Grants Council Committee of the University Grants Council of Hong Kong ((1) General Research Fund (GRF) Project No. 844012 "Quadrasyllabic Idiomatic Expressions (QIEs) in Chinese and neighboring Languages: An Investigation into Linguistic and Cultural History" and (2) GRF Project No.148908 "A Quantitative and Qualitative Comparison of Word Formation in Modern Standard Chinese and Early Modern Chinese"). I am grateful for comments leading up to this paper from co-investigators in the two projects: Andy Chin, Hintat Cheung, and particularly Shin Kataoka who has drawn my attention to many of the examples in this paper.

26th Pacific Asia Conference on Language,Information and Computation pages 39–55

On the other hand, when two related languages have dissimilar terms to express similar objects or events, then the difference could well represent salient non-linguistic variations. For example, the word for government in Indonesia is *Pemerintah* and in Malay *Kerajaan*. In the latter case of Malay, the word reflects the structure of government involving constitutional monarchy (as indicated by "*Rajah*") whereas the case of Indonesian reflects an organization structure presided over by a leader. The form *Selamat* means "hello" in Indonesia and Malaysia, originating from Semitic languages: Arabic S*alam* "peace" e.g. *Salaam Alaikum* "peace be with you" and Hebrew *Shalom* (peace). But in the Philippine languages, it means "thank you". This shift of meaning may not be unreasonable if we consider the broader context of language contact interaction in which we find the universal and customary conversation opening and closing moves, which are the same in Islamic societies (*Salaam Alaikum*), in stark contrast to English (with *hello-hi* and *goodbye* respectively) and other languages. In the exchange of identical but multifunctional pragmatic expressions during the opening and closing communicative moves among participants, a possible semantic switching taking place could be understandable.

In Asia, two long standing major classical traditions have been recognized:

(I) Sanskrit base [Indosphere[2]]

Devanagari, on which the Sanskrit writing system is based, has influenced the writing systems of Indosphere languages of the South Asian subcontinent, Burmese, Thai, Lao, Tibetan etc, but not Indonesia and Malaysia in which once dominant Hindu Kingdoms in the Indonesian archipelago have given way to Islamic sultanates, with exceptions to be found in Bali, for example. In these languages, there has not been much evidence of the Indic past in non-materialistic terms, other than loan words, while Jawi, the script derived from Arbic, still survives.

(II) Sinitic base [Sinosphere]

Its emblematic logographic writing system has greatly influenced the historical development of Sinosphere writing systems in Japan, Korea, Vietnam, and among other ethnic groups like the Nasi etc, on which the associated classical traditions, including the Chinese classical language have had significant impact. Thus their students to this day are often exposed to literary classics of Chinese origin such as the *Chronicles of Three Kingdoms* (三國演義) and *Water Margin* or *All Men and Brothers* (水滸傳). This tradition bears interesting comparison with the lesser trend of students in Thailand, Laos, and Cambodia (but not Indonesia or Malaysia) studying the Indic epic Ramayama. One distinctive feature of languages associated with Sinosphere is the importance given to relatively unique idiomatic expressions such as 不三不四 [not-3-not-4] "improper", similar to English "neither fish nor fowl" but with stronger negative connotations. For example, civil servants in Japan, Korea and Vietnam, in order to gain promotion, have to take language examinations in which there are expectations on familiarity with such expressions. This is often seen as a difficult and arduous task because of the drastic typological linguistic differences between Japanese, Korean and Vietnamese on the one hand, and Chinese on the other hand. Thus, considerable efforts have to be made by the civil servant aspiring to promotion.

It is interesting to note that whereas Korea and Japan, for example, have adopted the Chinese logographic writing system, and have even incorporated it into basically at one time or another

[2] Matisoff (1990) proposed the terms *Sinosphere* and *Indosphere* to distinguish between two major and often superimposed cultural traditions within Asia.

bimodal writing systems. On the other hand, related languages such as Mongolians and Manchus switched to the Chinese language when they conquered all of China, rather than imposed their own language as the native language, with possible adaptation or adoption of the logographic script. There were some minor unsuccessful attempts such as that by the Kitan Kingdom (契丹) which developed a demotic script, and the use of Phags-Pa script of the Mongols, which though squarish in shape and written from right to left, was much more influenced by the writing system of the Tibetans who have shared Lamaism as a common religion.

2. Background on Quadrasyllabic Idiomatic Expressions (QIEs) of Chinese origin

Idioms have (a) relatively stable and unusual parallel phonological, syntactic and/or semantic patterns, (b) semantic sophistication (metonymy, hyponymy, locus classicus, etc.), requiring background knowledge and draws on (c) metalinguistic ability to differentiate between metaphorical literate versus literal meanings and projected positive or negative sentiments, as in the above English example of "neither fish or fowl" and 不三不四 [not-3-not-4] "improper", or logical deduction, such as "(as) poor as a church mouse" in English[3]. While similar structures are found in different idiomatic expressions, one unusual type of idiomatic expressions with origins in Sinosphere stands out from the others and they have pervasive presence in the region.

It would be rewarding to systematically explore: (a) The extent of spread of such similar idiomatic expressions in the region; (b) The sociolinguistic and historical status and extent of Chinese as a "High" or "Supreme" status language (Tsou and You 2007) in the relevant language communities, including the significance of the logographic writing systems or its absence; and (c) The degree of structural compatibility between the relevant regional languages and Chinese, and how it might influence horizontal transfer. There is considerable value to examining their emergence, alteration, innovation, or selection in the context of cultural equilibrium or punctuated equilibrium (Aikhenvald and Dixon 2001) and in terms of a hierarchy of borrowable elements (Curnow 2001) to shed light on the development and expansion of Sinosphere. More details on the structure of this type of Chinese idiomatic expressions are given below.

Even though the Chinese language has the tendency to be monosyllabic and its writing system morpho-syllabic, a large portion of its words consist of disyllables which can be aggregated as longer linguistic expressions.

The following table provides a comparison of very likely equivalent English and Chinese idiomatic expressions:

1. I'm all ears	洗耳恭聽 [wash-ear-polite-listen]
2. Strike while the iron is hot	打鐵趁熱 [strike-iron-during-heat]
3. Take the rough with the smooth	逆來順受 [negative-come-positive-take]
4. Walls have ears	隔牆有耳 [through-wall-have-ears]
5. Advice most needed is least heeded	忠言逆耳 [honest-words-negative to-ears]
6. After a storm comes a calm	否極泰來 [negative-extreme-calm-come]
7. An eye for an eye	以眼還眼 [take-eye-respond-eye]
8. Birds of a feather flock together	物以類聚 [thing-take-class-gather]

[3] This is because in puritanical times, churches would have been good examples of frugality and so there would not have been much leftover for the resident mice there.

9. Blood is thicker than water	血濃於水 [blood-thick(er)-than-water]
10. Do in Rome as the Romans do	入鄉隨俗 [enter-village-follow-custom]
11. Don't cry over spilt milk	覆水難收 [upset-water-hard-recover]
12. A man may dig his grave with his teeth	禍從口出 [calamity-from-mouth-come]

Table 1. Some Equivalent English and Chinese Idiomatic Expressions

It is quite clear from the above comparison that the English expressions are of uneven length but Chinese are quadrasyllabic (and quadra-logographic) expressions of even length.

The use of QIE in Chinese is pervasive in many domains of discourse and language use. For examples:

(13) *Greetings*: 好久不見 [very-long-no-see] "long time no see", 不見不散 [no-see-no-disperse] "wait until we meet"
(14) *Slogans*: 安全第一 [safe-whole-number one] "safety is top priority", 酒後勿駛 [drink-after-don't-drive] "don't drive if you drink"
(15) *Movie names*: 窈窕淑女 [slim-fit-gentle-lady] "My Fair Lady", 浩劫重生 [calamity-again-alive] "Cast Away"
(16) *Advertisement* (*Real Estate*): 全海靚裝 [all-sea-beautiful-renovation] "full seaview", 樓皇氣派 [building-emperor-air-atmosphere] "imperial bearing"

Chinese QIEs are relatively distinct linguistic structures, standing out from regular language, comparable to the use in English of Latin or Latinate expressions *Lacuna/ lacunae*; *Caveat emptor*. Specifically, some defining characteristics of QIEs may be summarized as follows:

a) Four syllables or logographs
b) Relatively fixed structure and patterns
c) Figurative meaning and semantic opacity

The quadrasyllabic structure draws on a basic disyllabic propensity in Chinese, reflecting, for example, a common reduplicative tendency in addressing close relatives:

媽 *ma* → 媽媽 *ma-ma* "mother"
爸 *ba* → 爸爸 *ba-ba* "father"
姐 *jie* → 姐姐 *jie-jie* "sister"

The quadrasyllabic propensity is further evidenced by contractions from pentasyllabic expressions, for examples:

(17) 傻人有傻福 → 傻有傻福
[Silly-person-has-silly-blessing] → [silly-has-silly-blessing]
"Innocence is blessing"

(18) 新瓶裝舊酒 → 新瓶舊酒
[New-bottle-contains-old-wine] → [new-bottle-old-wine]
"New wine in old bottle"

(19) 事後諸葛亮 → 事後孔明
[Event-after-Zhu-ge-liang[4]] → [event-after-Kong-ming]
"Wisdom in hindsight"

Quadrasyllabic expressions can result from systematic compression of well-known lines from the classics, as can be seen from examples derived through such compression of verse taken from *The Book of Odes* 詩經 (10^{th} – 7^{th} B.C.):

	A		B	+	C
(20)	夢寐以求	←	窈窕淑女，		寤寐求之
	[dream-sleep-to-seek] "desiring in dreams"				
(21)	愛莫能助		維中人甫舉，		之愛莫助之
	[love-cannot-able-help] "unable to help"	←			
(22)	人才濟濟		濟濟多士，		文王以寧
	[person-talent-crowd-crowd] "bountiful talents"	←			

It can be seen from the above examples that QIEs are pervasive and deeply entrenched within the *Chinese cultural tradition* since historical times.

QIEs contain relatively stable patterns of syntactic, semantic and phonetic parallelism, full or partial syllabic reduplication (i.e. phonetic parallelism) which are universal in language, such as *pera pera* meaning "fluent" in Japanese, and can cover alliteration, rhyming, and onomatopoeia e.g. *hanky-panky* in English, *xilihuala* 稀里啦 "noisy, messy" in Mandarin, *bingling-bamlam* "noisy" in Cantonese. However there can be more complex syntactic and semantic parallelism (e.g. synonymy) as well as antithetical parallelism (Tsou 1968) (e.g. contrasting or antonym pairs as in 天長地久 [sky-long-earth-lasting] "perpetual" or 水火不容 [water-fire-not-contain] "incompatible"). The rich and complex instances of parallelism are quite extensive.

QIE's complex semantic content is usually much greater than the aggregated meaning of the constituent morphemes and disyllabic words. They typically carry deeper connotations than their simple paraphrases, and can involve, if not project, awareness of shared cultural background and familiarity with Classical Chinese, for example: 三顧茅蘆 [three-glance-thatch-cottage], literally meaning "(paying) three visits to the thatched cottage". This QIE conveys an earnest invitation to someone to assume important responsibility, and is based on King Liu Bei's 劉備 three famous attempts to draw his chief strategist Kong Ming 孔明 (3rd Century AD) out of self-imposed isolation, as recorded in the *Chronicles of Three Kingdoms*.

[4] *Zhu-ge-liang* 諸葛亮 and *Kong-ming* 孔明 are names of the same minister whose wisdom is legendary from the *Chronicles of the Three Kingdoms*. In everyday language, quadrasyllabic, pentasyllabic expressions or expressions of other length may be found but the more frequent use of the former, especially in more formal discourse, would signify erudition.

QIEs involve discoursal opacity, which entails metalinguistic ability to differentiate between literal and metaphorical usage, which in turn can draw on logical deduction and can project positive or negative polar sentiments as rhetorical devices. For instance, the QIE 孤男寡女 [lonely-man-single-woman] "unmarried couple" has negative connotations arising from Confucian disdain for interaction among unmarried male and female. It is found among inappropriate sentences composed by secondary school students drawn from the author's previous fieldwork in China: "丈夫死後，他們娘倆孤男寡女，相依為命過著艱難的生活"，literally "*after the death of her husband, the widow and son, being "**lonely man and single woman**", relied on each other and lived a hard life*". In such an example, metalinguistic ability is absent to distinguish between literal and metaphorical meanings as well as the negative connotations, and there are hints of malapropism.

The traditional and extensive native Chinese literature on QIEs has been preoccupied with whether QIEs are words or set phrases, and with the proper classification of such expressions (Liu 1984; Zhou 1994, 1997; Xu 1997) into subcategories. For example:

- *Idioms* 成語*,* often involving Locus Classicus, e.g. No. (11) 覆水難收 [poured-water-hard to-recall] "irreversible case", which is based on a Han dynasty wife, who had left a poor husband, and who later could not reinstate herself as his wife after he passed the Imperial examination and became a high official. In this QIE, the conclusion of irreversibility could also be logically deduced without Locus Classicus;
- *Common sayings* 熟語*,* e.g. 不三不四 [not-three-not-four] "improper";
- *Colorful terms* 諺語*,* e.g. 你死我活 [you-die-I-live] "(fighting) fiercely", 混水摸魚 [muddy-water-catch-fish] "opportunistic"; and
- *Idiomatic riddles* 歇後語, e.g. 和尚打傘 [Buddhist-priest-hold-umbrella] implies 無法無天 [no-*hair (law)* (homophonic)-no-sky]. Here *hair* and *law* are homonyms in Chinese, and *sky*, the symbol of justice in Chinese culture, is blocked by the umbrella, therefore "a lawless society". Here, the first QIE is paired with a second, which is often unexpressed but appreciated after the puns are resolved.

3. QIEs in some East Asian languages

In comparison to tone and monosyllabicity, these QIEs are much more representative of a likely unique linguistic trait of the Chinese language and are much more emblematic of Sinitic civilization. Their use in Chinese has much more significant rhetorical and sociolinguistic status when compared with the parallel use for foreign expressions in English and other European languages. Their judicious use provides an indication of desirable erudition and cultured status of the user and, as maybe expected, they are commonly found in socio-culturally elevated speech registers. Such expressions have been imported and calqued in Japanese, Korean, and Vietnamese, etc (i.e. QIE-prone languages) with which Chinese has had intensive contact. Moreover they are found in great abundance among the non-Sinitic languages of Southwest China, such as the Zhuang-Dong and Loloish and there is overlap with Southern Chinese dialects, especially Cantonese.

Examples given below are taken from other Asian languages, constituting distinctive and often autonomous linguistic expressions, which stand out from the usual language but which are integrated with the full discourse structure, much as the Latin expressions in English, as mentioned earlier.

(23) QIE examples from Japanese:

a) **山紫水明** (さんしすいめい)
b) **人事不省** (じんじふせい)
c) **解衣推食** (かいいすいしょく)
d) **闌摧玉折** (らんさいぎょくせつ)
e) **広大無辺** (こうだいむへん)
f) **以夷制夷** (いいせいい)
g) **前人未踏** (ぜんじんみとう)
h) **遠慮会釈** (えんりょえしゃく)

(24) QIE examples from Korean:

a) 각골명심 (刻骨铭心)
b) 간담상조 (肝胆相照)
c) 객반위주 (客反为主)
d) 가담항설 (街谈巷说)
e) 거안사위 (居安思危)
f) 견리망의 (见利忘义)
g) 거안제미 (举案齐眉)
h) 격물치지 (格物致知)

(25) QIE examples from Vietnamese

a) đồng bệnh tương lân (同病相憐)
b) ngư ông đắc lợi (漁翁得利)
c) tụ tinh hội thần (聚精會神)
d) thủy trung lao nguyệt (水中撈月)
e) hữu danh vô thực (有名無實)
f) phu xướng phụ tùy (夫唱婦隨)
g) nhập gia tùy tục (入家隨俗)
h) Đả thảo kinh xà (打草驚蛇)

(26) QIE examples from Zhuang

a) Dem gyaeuj dem rieng (添枝加葉)
b) Dub gu fong rek (挖肉補瘡)
c) Duh caeg sim diuq (做賊心虛)
d) Bae naj yawj laeng (瞻前顧後)
e) Nyaeb sip haeuj rwz (自討苦吃)
f) Sam sim song hoz (三心兩意)
g) Langh bit roengz raemx (正中下懷)
h) Ep meuz gwn meiq (強人所難)

(27) QIE examples from Cantonese

a) 九牛一毛 [9-ox-1-hair] "a drop in the ocean"
b) 人山人海 [people-mountain-people-sea] "a large crowd"
c) 人頭豬腦 [human-head-pig-brain] "a stupid person"
d) 九唔搭八 [9-not-match-8] "completely nonsensical"
e) 朝行晚拆 [morning-set-night-demount] "industrious"
f) 秤不離舵 [libra-not-leave-rudder] "inseparable"
g) 逆來順受 [negative-come-positive-take] "take the rough with the smooth"

It can be seen from the above examples that these languages are part of the logographic cultural circle in Sinosphere with varying degrees of overlapping cultural traits, and with the presence of QIEs.

According to Shibatani (1990), about 60% of entries in a modern Japanese dictionary are estimated to be Sino-Japanese. QIEs (yojijukugo 四字熟語) are also an integral part of Sino-Japanese, reflecting a millennium of contact since the adaptation of the Chinese logographic writing system. They are part of the syllabus for the national language Kokugo 国語 and even for high school and university entrance exams as well as civil-service exams. Interestingly, as early as 1007, Minamoto Tamenori had already compiled a book of idioms *Sezoku Genbun* 世俗諺文 for Japanese students. Korean and Vietnamese also have many QIEs of Chinese origin, which are called 사자성어 四字成語 and thành ngữ Hán 成語漢 respectively.

It is not surprising that Japanese, Korean, and Vietnamese speakers would encounter significant challenge to comprehend Chinese QIEs because of typological differences from their own languages, e.g. opposite order of [Object + Verb] and [Attribute + Head] Therefore the common adaptation of QIEs in Japanese and Korean present an unusual opportunity to study how and, more importantly, why typologically different languages might overcome such severe linguistic barriers. Given such linguistic handicap, there is a need to consider the sociolinguistic history and nature of language contact China has had with Japan and Korea.

Structural accommodation is necessary in the indigenization of some Chinese QIEs in Japanese, Korean and Vietnamese and their calques. We could note below 3 kinds of processes: (a) Manipulation of word order: Japanese and Korean are SOV languages. Some QIEs with SVO order have become SOV in Japanese and Korean: e.g. Chinese 不省人事 [not-recognize-people-matter] "fully unconscious" (VO) becomes 人事不省 [people-matter-not-recognize] (OV) in Japanese. Also Chinese 露出馬腳 [expose-out-horse-leg] "betray oneself" (VO) becomes □□□□ (馬腳露出) [horse-leg- expose-out] (OV) in Korean; (b) Paraphrase: 俟河之清 [wait- Yellow river-attrib-clarity] "wait for something that never happens" in Minamoto's 1009 book appears now in contemporary Japanese only after syntactic accommodation (reversal): 河清を俟つ [river-clarity+acc. marker+wait]. Vietnamese has [HEAD+ATT.] whereas Chinese has the reverse order. Chinese QIE 井底之蛙 [well-bottom-attrib-frog] "a person with limited vision" [ATT.+HEAD] has two manifestations in Vietnamese: (i) tỉnh ñế chi oa (井底之蛙) (original Chinese), but also (ii) ếch ngồi ñáy giếng (蛙坐底井) [frog-sit-bottom-well] "indigenized"; and (c) Innovation: Original

extensions of QIEs are found in Japanese, Korean, Vietnamese languages: e.g. ichigoichie 一期一会 [one-cycle-one- meeting] “an encounter with someone only occurs once in life” (Japanese), 문전옥답 (門前沃畓) [gate-front-abundant-field] “well-off family” (Korean), or tình nhập lý (入情入理) [enter-feeling-enter-logic] “reasonable” (Vietnamese). These examples suggest a hypothesis that *structural incompatibility may be accommodated in purposeful indigenization by restructuring.

On the other hand, there is relatively low adoption of QIEs among typologically similar, if not genetically related, Mongolian (e.g. Tanaka 2005), Manchu and Uyghur, which shows great contrast with QIE-prone Japanese and Korean and invites explanation. It is noteworthy that these QIE-resistant languages had made short-lived attempts to develop different writing systems, ranging from the Tibetan inspired Mongolian ‘Phags-pa’ script (Coblin 2006) and Uyghur inspired Jurchen Script (Kane 2009), which showed Chinese influence mostly by being written vertically down and from right to left, with essentially mono-syllabic symbols. The reasons for the demise of these scripts deserve extensive studies in the context of this project.

Furthermore, in the south and as noted, there are many QIE-prone non-Sinitic languages which have not seriously adopted Chinese logographic writing system, or any sustained writing tradition (e.g. Li in Hainan, Bai in Yunnan and Zhuang in Guangxi). We note that QIE-prone Zhuang and related languages have internal rhyme and show evidence of related rhyming metathesis which bear interesting comparison with Cantonese lexical metathesis not found in northern dialects. This complex and unusual feature allows us to consider whether QIEs may not be a readily borrowed feature but could be a possible shared genetic linguistic feature between Cantonese-Yue and Zhuang, which will need to be fully examined and tested. Spoken Cantonese lexicon contains many native QIEs, in addition to those shared with Mandarin. Of special interest would be constituent switching or lexical metathesis found in Cantonese QIEs.

(28) A1 A2+ B1 B2 => A1 B1+ A2 B2
揀擇飲食[choose-select-eat-drink] =>
揀飲擇食[choose-eat-select-drink]
“picky on food”

(29) A1 A2+ B1 B2 => A1 B2+ B1 A2
朝拆晚行 [a.m.-dissemble-p.m.-assemble] =>
朝行晚拆 [a.m.-assemble-p.m.-dissemble]
“for convenience”

(30) A1 A2 => A1 + XY + A2
事實 [fact] =>
事不離實 [matter-NOT-LEAVE-substance]
“factually speaking”

In No. (28), near-synonyms or hyponyms (*drink*, *eat*) have been juxtaposed and a play on the normal Cantonese phrase 揀食 [choosy-food] “picky on food” by switching to the unusual 揀飲 [choosy-drink] “picky on drink”. No. (29) shows the interesting result of clear metathesis, which

would be illogical to the discerning hearer because in cramped living quarters (as in Hong Kong), a collapsible bed should be dissembled in the morning and reassembled at night (and not the reverse order indicated by the surface structure). In No. (30), a disyllabic word 事實 [matter-substance] "truth" has been paraphrased quadrasyllabically with infixing morphemes 事不離實 [**matter**-not-leave-**substance**] thereby leading to the semi-productive creation of a new QIE. It is also an analogic derivation from a traditional Cantonese rhyming paired QIEs drawing on the similie: 公不離婆 [husband-not-leave-wife] (**like**) 秤不離砣 [scale-not-leave-weight] "the husband and wife being together like the scale and its weight" i.e. "showing a close and intimate relationship" where 婆 (po) and 砣 (to) are rhymes.

4. The internal structure of QIEs

Chafe (1968) draws on the famous example of English idiom: ***kick the bucket*** and shows that it shares the same part of speech as its idiomatic counterpart '***to die***'. Thus the sentence "the bucket was kicked by him" can only have the literal meaning but not the metaphorical meaning of dying because 'to die' is intransitive just as *waterloo* would be a mother noun like its literal counterpart *defeat*. Similarly, Chinese QIE can also assume different parts of speech accordingly. For examples,

(31) as noun:

你們都是+ABCD
[you-are-all-ABCD] (ABCD = 烏合之眾 [dirty-group's-gang] "motley crew")

就像+ABCD+一樣
[just-like-ABCD] (ABCD = 井底之蛙 [well-bottom's-frog] "frog under the well")

(32) as adjective:

V 得+ABCD
[V-until-ABCD] (ABCD = 落花流水 [fall-flower-flow-water] "like fallen flowers")

這麼+ABCD
[so-ABCD] (ABCD = 粗心大意 [thick-heart-big-meaning] "careless")

(33) as verb:

一定+ABCD
[definitely-ABCD] (ABCD = 盡力而為 [all-effort-to-do] "with all (his) might")

你應該+ABCD
[you-should-ABCD] (ABCD = 再接再厲 [re-take-re-sharpen] "continue on and on")

They are finite possibilities for the internal morphological and syntactic structures of QIE.

(34)

a) ABCD = ABC+D / AB+CD = NP
b) ABCD = A+B+CD / AB+C+D = SV
c) ABCD = AB+CD = VP sequence
d) ABCD = AB+CD = coordination
e) ABCD = AB+CD = subordination

It follows from the above that 3 kinds of linguistic knowledge are evident in QIEs: i.e. (a) *structural parallelism*; (b) *semantic saliency*; (c) *discoursal opacity*.

Table 2 below provides some examples of structural parallelism:

(35)**千 山 萬 水** '1K-mountain-10K-waters'		千-萬、山-水	
(36)不**明**不**白** 'not-bright-not-clear'	不-不	明-白	
(37)**如霜似雪** 'like-frost-like-snow'	如-似	霜-雪	
(38)先**苦**後**甜** 'first-bitter-later-sweet'			先-後、苦-甜
(39)無**拘**無**束** 'no-arrest-no-restrict'	無-無、拘-束		
	Synonymy	Hypernymy	Antonymy

Table 2. Examples of structural parallelism

It can be seen that 如-似 [like-similar] "similar to" and 拘-束 [arrest-restrict] "control" are synonymous and 不-不 [no-no] and 無-無 [without-without], being reduplications, are extreme cases of synonymy. By comparison, 山-水 [mountain-water] share the hypernym "terrestrial objects", 明-白 [bright-clear] "clarify" share the hypernym "cognition", 霜-雪 [frost-snow] share the hypernym "weather". Furthermore, 先-後 [precede-follow] "sequence" and 甜-苦 [bitter-sweet] "life's extremes" are antonymous. It can be seen that the rhetorical devices used involve synonymy, hyponymy and antonymy and are commonly deployed in the projection of discoursal opacity.

More specifically, the relevant internal linguistic features may be further analyzed as in the following:

a. Hypernymy

(40) 三五成群 [3-5-become-crowd] "in small groups"
(41) 三六九等 [3-6-9-etc] "in different groups"
(42) 三教九流 [3-religion-9-branch] "the riff raff"
(43) 三心兩意 [3-heart-2-mind] "undecided"
(44) 張三李四 [Zhang-3-Li-4] "any Tom, Dick or Henry"

b. Classical language usage

(45) 三年五載 [3-year-5-year] "in-a-few-years"
(46) 三思而行 [3-think-then-act] "think before acting"
(47) 三差五錯 [3-error-5-mistake] "any deviation"

c. Culture bound

(48) 三生有幸 [3-incarnation-have-luck] "forever indebted"
(49) 三從四德 [3-obedience-4-virtue] "traditional loyalty (for women)"

d. Locus Classicus (50) 三過其門 [3-pass-his-door] “devoted to duty” (51) 朝三暮四 [morning-3-evening-4] “indecision” (52) 舉一反三 [propose-1-reply-3] “good logical deduction” (53) 孟母三遷 [Mencius-mother-3-move] “moving to better environment”
e. Synonymy (54) 三回四次 [3-times-4-occasions] “many times” (55) 說三道四 [say-3-call-4] “mumbling insignificant things”
f. Word Morphology (56) 三差五錯 [3-error-5-mistake] “any deviation” (差-錯) (57) 三災八難 [3-calamity-8-difficulty] “disaster” (災-難) (58) 三長兩短 [3-long-2-short] “accident” (長-短)
g. Homonymy (phonetic/semantic replication or rhyme) (59) 三三五五 [3-3-5-5] “in small groups” (cf. 不三不四)
h. Antonymy (60) 三長兩短 [3-long-2-short] “accident” (61) 朝三暮四 [morning-3-evening-4] “Indecision” (62) 三好兩歉 [3-good terms-2-apologies] “inconsistent relationship”

Table 3. Eight major linguistic features associated with QIE

Table 3 singles out eight special features of QIEs drawn from LIVAC[5]. From more than 30K entries found there, 130 entries, involving the numeral 3, are used as examples:

a) The *hypernymic* relation is by far most commonly drawn on to convey metaphorical meaning. Thus, No. (42) 三教九流 [3-religon-9-branches] signifying too many diversified sects is used to project the image of disorganized ‘riff raff’. In No. (44) *Zhang* 張 and *Li* 李, being common manifestations of the hypernym *surname*, alternate with the hypothetical given names: sequential numbers 3 and 4, which belong to the hypernym of *number*.

[5] The LIVAC (Linguistic Variations in Chinese Speech Communities) [http://livac.org] synchronous corpus has been based at the Research Centre on Linguistics and Language Information Sciences of The Hong Kong Institute of Education since 2010. It continuously draws on the analysis of texts from representative Chinese newspapers and electronic media of major Chinese communities in Beijing, Hong Kong, Shanghai, Singapore, Taipei from 1995. By 2012, 450 million characters of texts have been analyzed and 1.5M words have been culled from them in the corpus.

b) *Classical Chinese* knowledge is needed. For examples, No. (45) 三年五載 (MSC), No. (47) 三差五錯 (MSC), where 載 and 差 are semi-bound nouns in Modern Standard Chinese, but free morpheme in Classical Chinese.

c) *Culture bound*. No. (48) 三生有幸 refers to multiple sequential reincarnations and so extended duration of gratitude. No. (49) 三從四德 refers to traditional obedience for women toward her father, her husband, and her son, a reflection of customary culture of loyalty of the past.

d) *Locus Classicus*. In addition to Chinese cultural tradition, some items are drawn from historical events (compared to *Achilles' heel*, *Waterloo* (defeat) etc). In No. (53) 孟母三遷, the mother of the sage Mencius 孟子 moved three times in order to ensure her son kept good company. No. (50) 三過其門 refers to Xiayu 夏禹 who was Minister in charge of flood control and who was so devoted to duty that he did not stop by even when passing by his own home.

e) *Synonymy* – terms with equivalent meaning are used as a way to reinforce the thrust of the semantic content, e.g. 說三道四 [say-3-call-4] "mumbling insignificant things"

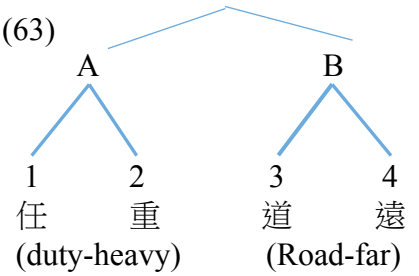
f) *Morphological structure* of Modern Standard Chinese where the distinction between free and semi-bound morphemes exists, e.g. 三差五錯, 三災八難 where 差 and 災 are semi-bound morphemes in MSC.

g) *Homophony* - Identity in terms of phonological and semantic content is a simplistic reinforcement of the parallelism in structure.

h) *Antonymy* - Ability to binary opposite distinction (in addition to lateral similarity as in synonymy, and hierarchical similarity (in most cases of hyponymy) is important to complement the linguistic, cultural, and cognitive skills.

The internal morphology of QIE can be represented as a coordinate and parallel structure.

(63)



"important and arduous mission"

The following table provides a breakdown of the different internal grammatical patterns in QIEs.

Types	%
Coordinative	35.0
Attributive	21.5
Subject-predicate	17.5

Verb-object	15.0
Other	11.0
Total	100.0

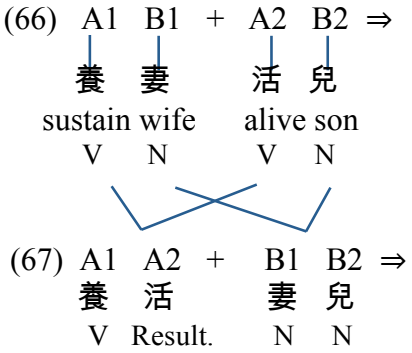
Table 3. Distribution of structural types

Following the common preference of structural parallelism, some likely and interesting structural variations between underlying and surface forms are noted.

Thus, variations in terms of permutation or metathesis could involve different comparable syntactic units and revisions in argument structure from a base structure, consider:

(64) 養妻活兒 [sustain-wife-alive-child]
"to maintain family"
(65) 養活妻兒 [sustain-alive-wife-child]

The structural ambiguities in (64) and (65) can be structurally represented as (66) and (67) below:



In (66), the static verb 活 "alive" has apparently become a causative verb "to cause to be alive" with 兒 "son" as object, in parallel with verb-object 養妻 [sustain-wife] because of structural parallelism, and poetic license, but in actual fact it could be also the simple metathesis between resultative verb 活 "alive" in the disyllabic verb 養活 [feed-alive] "sustain" with the first object 妻 "wife" of the disyllabic compound 妻兒 [wife-son] "family" in the underlying No. (67). Such a case invites the hypothesis that the path of production of the QIE may be different from the path of cognition. Preliminary investigation shows that Southerners like Cantonese quite readily accept categorial shift between *stative verb* and *transitive verb* for 活 "alive, cause to be alive" and so they readily accept No. (66). But Northerners tend to see exceptional poetic license in No. (66), which they would normally not accept.

Another relevant pair of examples can be seen in No. (68) and No. (69).

(68) 魚沉雁落 [fish-sink-duck-down]
"unusual beauty"
(69) 沉魚落雁 [sink-fish-down-duck]

No. (68) and No. (69) refer to the understood exposure to unrivaled beauty which could cause fish to sink (to hide out of shame) and likewise wild geese to descend from flight (to hide). This situation has been rendered more graphic and dynamic with the normally intransitive static verbs 沉

"sink" and 落 "fall" in No. (68) projecting dynamic development as transitive verbs before the objects 魚 "fish" and 雁 "wild geese" as objects respectively in No. (69), where rhetorical if not poetic license has been exercised.

Our preliminary analysis from the above common Chinese QIEs involving numerals indicate that a gradation exists amongst five top cognitive skills associated with the eight features discussed earlier.

(1) Hyponymic relation
(2) Classical language usage
(3) Culture bound
(4) Locus classicus
(5) Similarity relationship (synonym and homonymy)

It would be useful to compare language acquisition among children with language attrition among language handicapped adults, such as those who suffer from Alzheimer's disease in relation to the attributes noted here, especially to see if there are complementary trends between the two situations.

5. Conclusion[6]

The eight linguistic and rhetorical features of QIEs and the indulgence in syntactic ambiguities and rhetorical niceties encouraged by poetic license are related to those often employed in traditional Chinese verse and prosaic discourse. The parallel relationship between similar linguistic elements, and the binary opposition of linguistic elements as well as their manipulation in QIEs are fundamental in Chinese literary traditions, culminating in the famous *Regulated Verse* 律詩 form and in rhetoric discourse, as in *The Literary Mind and Carving of Dragons* 文心雕龍 (5th Century AD). As such, they are an integral part of poetics: It is noted that "The poetic resources concealed in the morphological and syntactic structure of language, briefly the poetry of grammar, and its literary product, the grammar of poetry, have been seldom known to critics and mostly disregarded by linguists but skillfully mastered by creative writers" (Jacobson, 1961). Given the popularity of original and derived QIEs in the region (even for native Chinese speakers), but the immense complexity in structure and consequently the efforts needed to overcome linguistic hurdles by peoples within Sinosphere, a natural question can be readily posed: why should such cognitive handicaps be retained, even after the traditional cultures in Sinosphere have been challenged if not partially replaced by Western ones?

References

Aikhenvald, Alexandra Y. and R. M. W. Dixon (eds.) 2001. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics.* Oxford: Oxford University Press.

Chafe, Wallace L. 1968. "Idiomaticity as an Anomaly in the Chomskyan Paradigm." *Foundations of Language*, 4, 109-127.

Coblin, W. South. 2006. *A Handbook of 'Phags-pa Chinese*. Honolulu: University of Hawai'i Press.

Curnow, T.J. 2001. "What Language Features Can Be 'Borrowed'? In Aikhenvald and Dixon (eds.), 412-436.

[6] On the basis of the analysis of QIEs, an unprecedented *Chinese QIE Crossword Puzzle Games* 成語填字坊 has been developed and available through The Research Centre of Linguistics and Language Information Sciences of The Hong Kong Institute of Education and other platforms: (1) web: http://www.rclis.ied.edu.hk/crossword/, (2) Android: http://chilin.no-ip.org/android/; (3) iOS: http://chilin.no-ip.org/iphone/.

Everaert, Martin, E.-J. van der Linden, A. Schenk,. and R. Schreuder. (eds.) 1995. *Idioms: Structural and Psychological Perspectives*. Hillsdale, NJ: Erlbaum.

Everaert, Martin, et al. 1989. *Proceedings of the First Tilburg Workshop on Idioms*. Tilburg: ITK.

Everaert, Martin, et al. 1992. *Proceedings of IDIOMS.* Tilburg: ITK.

Fan, Shiyan. 2007. *Colors in Idioms*. MA thesis. Liaoning Normal University.

Fillmore, Charles, Paul Kay and Catherine O'Connor (1988). Regularity and Idiomaticity in Grammatical Constructions: The Case of let alone. *Language* 64: 501–38.

Fontecha, Almudena F. & Rosa M. J. Catalán. 2003. Semantic Derogation in Animal Metaphor: A Contrastive-Cognitive Analysis of Two Male / Female Examples in English and Spanish. *Journal of Pragmatics*, 35, 5: 771-797.

Gibbs, R. 1980. "Spilling the beans on understanding and memory for idioms in conversation." *Memory and Cognition*, 8: 449-456.

Gibbs, R. 1985. "On the process of understanding idioms." *Journal of Psycholinguistic Research*, 14; 465-472.

Gibbs, R. 1987. "Linguistic factors in children's understanding of idioms." *Journal of Child Language*, 14: 569-586.

Goldberg, A. & Suttle, L. (2010). *Construction Grammar*. Wiley.

Jackendoff, R. 1995. *Semantics and Cognition*. Cambridge, MA: MIT Press.

Jakobson, R. 1961. Closing Statement: Linguistics and Poetics. In Sebeok (ed), *Style in Language*. Cambridge: Mass., The M. J. T. Press.

Jakobson, R. 1981. *Selected Writings, Volume III: Poetry of Grammar and Grammar of Poetry*. The Hague: Mouton, 18-51.

Jakobson, R. 1990. "Two Aspects of Language and Two Types of Disturbances." In Linda Waugh and Monique Monville-Burston. *On Language.* Cambridge, MA: Harvard University Press.

Kane, Daniel. 2009. *The Kitan Language and Script.* Leiden; Boston: Brill.

Katz, Jerrold J. and Paul M. Postal. 1963. "Semantic Interpretation of Idioms and Sentences Containing Them." M.I.T. R.L.E., *Quarterly Progress Report*, 70, 275-82.

Lakoff, G. 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: Chicago University Press.

Langacker, Ronald (1987, 1991). *Foundations of Cognitive Grammar*. 2 vols. Stanford: Stanford University Press

Liu, Shuxin. 1984. "Fixed Expressions and Its Classification." *Yuyan Yanjiu Luncong*, 2.

Matisoff, James A. (1990). On Megalocomparison. *Language* 66.1, p. 113.

Mo, Pengling. 2001. *Chinese Idioms and Chinese Culture.* Nanjing: Jiangsu Jiaoyu Chubanshe.

Nall, Timothy M. 2008. *Analysis of Chinese Four-character Idioms Containing Numbers: Structural Patterns and Cultural Significance*. Unpublished PhD thesis. Ball State University.

Nippold, M.A. & S.T. Martin. 1989. "Idiom interpretation in isolation versus context: A developmental study with adolescents." *Journal of Speech and Hearing Research*, 32: 59-66.

Shibatani, Masayoshi. 1990. *The Languages of Japan*. Cambridge: Cambridge University Press.

Tang, Chihsia. 2007. "A Comparative Study of English and Chinese Idioms with Food Names." *UST Working Papers in Linguistics*, 3: 83-93.

Tsou, B.K. 1968. Some Aspects of Linguistic Parallelism and Chinese Versification. In Charles E.Gribble (ed.) *Studies Presented to Professor Roman Jakobson By His Students*. Cambridge, Mass:Slavica Publishers.

Tsou, B.K., T. Lee, H. Cheung, and P. Tung. 2006. *HKCOLAS: Hong Kong Cantonese Oral Language Assessment Scale*. Hong Kong: Language Information Sciences Research Center, City University of Hong Kong and Department of Health, HKSAR.

Tsou, B.K. and R.J. You. 2007. *A Course in Sociolinguistics*. Taipei: Wunan.

Xiang, Guangzhong. 1979. “Idioms and Ethnic and Cultural Tradition.” *Zhongguo Yuwen*, 2.
Xu, Yaomin 1997. “Definitions and Classification of Idioms.” *Zhongguo Yuwen*, 1.
Zhang, Zonghua 1984. “Semantics of Idioms.” *Cichu Yanjiu*, 4.
Zhou, Jian 1994. “Typicality and Atypicality of Idioms.” *Yuwen Yanjiu*, 3.
Zhou, Jian 1997. “On Typicality of Idioms.” *Nankai Journal*, 2.
Zuo, Zhijun 2006. *Acquisition of Chinese Idioms: From the Perspective of Cognition*. MA thesis. Ocean University of China.