

## A corpus-based comparative study of light verbs in three Chinese speech communities

**Benjamin K Tsou**

City University of Hong Kong,  
The Hong Kong University of Science and  
Technology / Hong Kong SAR  
btsou99@gmail.com

**Ka-Fai Yip**

Yale University / New Haven,  
Connecticut, United States  
kafai.yip@yale.edu

### Abstract

The universal category of light verbs has drawn considerable interest among researchers on Asian languages in recent years. For Chinese, recent research focused on several light verbs have included multiword expressions, some with parallels in English. This study begins with a common but little studied light verb DA 打 (“hit”) in Chinese, based on 22 years of data curated in the LIVAC Pan-Chinese corpus ([https://en.wikipedia.org/wiki/LIVAC\\_Synchronous\\_Corpus](https://en.wikipedia.org/wiki/LIVAC_Synchronous_Corpus)). The verb DA has differentially evolved from a regular transitive verb involving physical strike action to take on increasingly metaphorical extension and is well on the way to become a light verb. From 2175 lexical entries in LIVAC from 1995-2016, we are able to trace some longitudinal developments of this light verb across three Chinese speech communities (Beijing, Hong Kong, Taiwan) by comparing their differences in metaphorical extension. We also provide in detail how the three communities differ in the usage of this popular light verb, and the broader significance beyond linguistics, with Beijing and Hong Kong showing a higher degree of metaphorical shift in different ways. We also explore some means by which metaphorization may be compared with grammatical change. Moreover, we demonstrate how fruitful and new findings may be obtained with a rigorously curated corpus.

### 1 Background

Light verbs constitute a universal and unique class of words whose basic substantive meaning has been made opaque (or bleached) so that they could have become function words with little substantive

meaning. This universal linguistic class has interested linguists for sometime. In the case of English, Jespersen (1954) and Cattell (1984) have noticed them, such as *have* in *have a look*, *take* in *take a drive* etc. In the case of Chinese, as early as the Song dynasty Ouyang Xiu 歐陽修 (A.D. 1007-1072) had noted them in his book *Guitianlu* 歸田錄. In the 1980s modern linguists such as Lü Shuxiang, who used the term “dummy verbs” 形式動詞 (1999[1980]:294) and Wang Li, who referred to them as “markers for verbs” 動詞的記號 (1985:142) had also began to notice them.<sup>1</sup>

In 1985, the first major study of Chinese light verbs was published by Zhu Dexi. It focused on 6 major light verbs in Chinese (*jinxing* 進行 “proceed”, *zuo* 作 “make”, *jiayi* 加以 “add”, *geiyi* 給以 “provide”, *jiyu* 給予 “provide” and *yuyi* 予以 “provide”). This paper was followed by Zhou 1987, Li & Chai 1995, Mao 1997, Yan 1998, Chen 2003, Li 2003 and many others, which provided more descriptive coverage. Then came Diao (2004) who doubled the number of Chinese light verbs to cover 7 more of them (*congshi* 從事 “engage”, *zuo* 做 “do”, *guo* 搞 “make”, *gan* 幹 “do”, *nong* 弄 “make”, *jia* 加 “add” and *yu* 予 “provide”). His concern was with their grammatical usage in Mainland China. With data drawn from the Sinica Corpus, Wang (2004) adopted a corpus-based approach to study differences in the usage of three light verbs (*zuo* 做 “do”, *guo* 搞 “make”, *nong* 弄 “make”) in Taiwan. In 2014, Lin et al. and Huang et al. initiated comparison between light verb usage

<sup>1</sup> Light verbs are also widely found in other East and Southeast Asian languages, e.g. *suru* “do” in Japanese (Grimshaw & Mester 1988), *ha* “do” in Korean (Chae 1996), *lam* “do” in Vietnamese (Pham 1999), etc.

variations in Mainland and Taiwan Mandarin. They used data from the Annotated Chinese Gigaword Corpus which combines data from both Mainland and Taiwan in 1990-2002 to discuss the differential distribution of five light verbs (*jinxing* 進行 “proceed”, *congshi* 從事 “engage”, *zuo* 做 “do”, *guo* 搞 “make”, *jiayi* 加以 “add”). They pointed out that 進行 *jinxing* in Taiwan might take verb-object phrases as complements (e.g. *jinxing toupiao* 進行投票 “proceed to voting, (lit.) cast-ticket”), but not in Mainland. The general conclusion as reported in Jiang et al. (2016) was that light verbs in Taiwan were “more transitive” (i.e. more verbal) and less grammaticalized.<sup>2</sup>

It is interesting that one light verb stands out for escaping the attention of most linguists. This is *da* 打 “hit”, which had surprisingly interested Ouyang Xiu (1007-1072).<sup>3</sup> Most modern studies have focused on *da*'s diachronic development (J. Zhu 2004, Su 2009, Zhuang 2014 *i.a.*) but seldom on its light verb usage in Modern Chinese. Wang (1985) simply called it “a marker for verbs”. Ren (2013) studied how *da* was interpreted specifically in different grammaticalized contexts in Grounding Theory. A detailed empirical study is thus needed.

*Da* 打 is particularly remarkable in undergoing robust metaphorical extensions from denoting initially a physical strike action (e.g. *dasi* 打死 “beat to death”) to a complete light verb (e.g. *daya* 打壓 “suppress, (lit.) hit-press”). A systematic study on *da* could shed light on light verb developments and more generally on the process of metaphorization. Also, regional variations in metaphorization among Chinese communities may be fruitfully studied from a comparative perspective.

<sup>2</sup> For a related discussion on “verbalness”, see Shen & Zhang (2013). Also see Her et al. (2016) and Tsai (2017) for the light verb developments in Taiwan, and see Jiang (2020) for a recent comprehensive study on the semantics of light verbs.

<sup>3</sup> Ouyang Xiu noted in his book *Guitianlu* several examples of *da* with bleached meaning, e.g. *dachuan* 打船 “to make ships”, *dache* 打車 “to make cars”, *dayu* 打魚 “to fish”, *dashui* 打水 “to get water”, *dafan* 打飯 “to buy a meal”, *dayiliang* 打衣糧 “to distribute clothes and food to soldiers”, *dasan* 打傘 “to hold an umbrella”, *dalian* 打黏 “to stick papers”, *daliang* 打量 “to measure”, *dashi* 打試 “to examine vision”. While some of them are no longer in use (e.g. *dalian*, *dashi*), some have undergone semantic changes in Modern Chinese, e.g. *dache* means “to take a taxi” instead of manufacturing cars.

This paper focuses on metaphorization of *da* with reference to both latitudinal and longitudinal variations in three Chinese speech communities Beijing (BJ), Hong Kong (HK) and Taiwan (TW). From 2175 lexical entries in LIVAC drawn from 1995-2016, we trace the developments of this light verb across a 22-year long time span and also across three Chinese speech communities (BJ, HK and TW) by initiating comparison of their differential metaphorical extensions. We show in some details how the three communities differ in their usage of this popular light verb, and the broader significance beyond linguistics, with BJ and HK showing a higher degree of metaphorical shift in different ways. We also explore how metaphorical extension might come about and demonstrate the valuable use which can be made of a rigorously cultivated corpus.

The paper is organized as follows. Section 2 introduces the corpus base LIVAC and the methodology for analysis. Section 3 overviews *da*'s literal usage and metaphorical usage with regard to word structures. Section 4 investigates the latitudinal and longitudinal variations among three Chinese speech communities BJ, HK, and TW in metaphorization. Section 5 explores theoretical implications for metaphorization and corpus-based studies on variations, and concludes the paper.

## 2 Corpus base and methodology

This paper draws on the Pan-Chinese synchronous database LIVAC (<http://www.livac.org/>). Since 1995, LIVAC has processed and filtered representative media texts from Pan-Chinese communities including Beijing, Guangzhou, Hong Kong, Macau, Shanghai, Shenzhen, Singapore, Taiwan. By 2016, more than 600 million characters of news media texts have been rigorously curated (Tsou & Kwong 2015).

A total of 2175 lexical entries with *da* 打 in word initial (1730) and final (445) positions are found in LIVAC (1995-2016). These form the basis for the current study.<sup>4</sup> After filtering out singleton *da*, nouns and loanword entries, we have 812 remaining entries with *da* in initial position

<sup>4</sup> Compare with 774 entries in Taiwan's *Guoyu Cidian* 國語辭典 (<http://dict.revised.moe.edu.tw/cbdic/>), among which 284 overlap with LIVAC and 238 entries in *Xiandai Hanyu Cidian* 現代漢語辭典, among which 198 overlap with LIVAC.

and 354 in final position.<sup>5</sup> They are analyzed into three types in terms of metaphorization in usage:

**Type I (Literal):** Only literal meaning, e.g. *dasi* 打死 “beat to death, (lit.) hit-die”, *daquan* 打拳 “fist boxing, (lit.) hit-fist”

**Type III (Metaphorical):** Only metaphorical meaning, e.g. *daya* 打壓 “suppress, (lit.) hit-press”, *dajia* 打假 “crack down on counterfeit, (lit.) hit-falsehood”

**Type II (Incipient Metaphorization):** An intermediate type whose words could be used both literally (Type I) and metaphorically (Type III), e.g. *dazao* 打造 “(lit.) fabricate (furniture)” vs. “(metaph.) forge (bright future)”, *daci* 打氣 “(lit.) pump air” as in *wei luntai daqi* 為輪胎打氣 “pump air into tires” vs. “(metaph.) cheer on” as in *wei qiuyuan daqi* 為球員打氣 “cheer the players on”.

The extent of metaphorization in different speech communities could be represented by the ratio between tokens in metaphorical usage and total usage. A higher ratio indicates a greater degree of metaphorization. To capture metaphorical variations across the three communities, we further analyze Type II entries.

BJ, HK and TW have in total 672 entries in use among them. However, only 151 entries are in common use among them with minimum frequency of two. In Table 1, we can see that the small ratio of common usage indicates great regional diversity where the common denominator is smaller than the peripheral variations, while the reverse would indicate a greater uniformity. This is indicative of the richness of variety in metaphorization in the Chinese language. It is noteworthy that these 151 entries take up 95% of the total usage frequency in the three communities and thus are sufficiently representative.

Distribution	Types (%)	Tokens (%)
LIVAC	812 (100)	318,095 (100)
HK, BJ or TW	672 (83)	179,182 (56)
HK, BJ and TW	151 (19)	170,245 (54)

Table 1. Distribution of *da-X* entries

Among the 151 *common* entries, 48 of these have both literal and metaphorical usage. They

<sup>5</sup> Three uses are excluded: (i) singleton *da* 打, (ii) noun uses such as *dahuoji* 打火機 “lighter”, as well as proper names, e.g. *Daguling* 打鼓嶺 “Ta Kwu Ling (Hong Kong place name)”, (iii) loanwords, e.g. *dabi* 打吡 (from English *derby*).

take up 60% (102,277) of the *common* usage (170,245).<sup>6</sup> To analyze and compare the relative frequencies of literal vs. metaphorical usage, we extract sample sentences containing them from two time periods: (a) 1995-2000, (b) 2011-2016 and obtain 16955 sentences from our database.<sup>7</sup> We then manually tag each case as either literal usage or metaphorical usage.<sup>8</sup> We can quantify the extent of metaphorization for *each* Type II entry by comparing these two kinds of usage.

An overall Metaphorization Index (MI) may be proposed to provide an objective measure:

$$MI = \frac{C + (B_1 + B_2 \dots + B_n)}{A + B + C} \times 100\%$$

(where A, B, C = total tokens of Types I, II & III respectively, and  $B_n$  = tokens of metaphorical usage of a Type II entry)

Following analysis, the incipient type (II) may be redistributed as either literal (I) or metaphorical usage (III) in a bipolar division among the 16955 sample sentences.

The current approach differs from previous studies in at least five aspects: (a) **Choice of light verbs.** While the previous studies were concerned with 13 light verbs mentioned above such as *jinxing* 進行 and rarely studied *da* 打, this study focuses on 2175 lexical entries of *da*. (b) **Metaphorization.** While the previous studies emphasized grammatical differences of light verbs, e.g. compatibility with aspectual markers, argument structures and eventualities of the complements, etc., the current study, on the other hand, centers on metaphorization, a crucial process in light verb development and in language. One of the defining properties of light verbs is the reduction of substantial meaning as the shift from literal meaning to metaphorical meaning. (c) **Regional variation.** In contrast to the previous studies focused on BJ and TW, this study also includes HK to provide a fuller picture of the Pan-Chinese language situation. (d) **Longitudinal variation.** Huang et al. (2014), as a major corpus-based study, focuses on only synchronic variations. The present study samples sentences through a

<sup>6</sup> The literal Type I takes up 12% (20179) and metaphorical Type III takes up 28% (47792).

<sup>7</sup> For each Type II word at each community each year, a maximum of 50 sentences were extracted (i.e. at most 600 sentences for each word).

<sup>8</sup> Every sentence was tagged by two annotators. Sentences with disagreement were adjudicated by a third annotator.

time window of 22 years to explore longitudinal comparison. (e) **Sampling size.** Huang et al. (2014) extracted 2000 sentences from the Chinese Gigaword Corpus in total, 200 sentences for each light verb in each region. This study extracts 16955 sentences in LIVAC, with over 170,000 tokens of 打 *da* from two periods.

### 3 Metaphorization of *da*

#### 3.1 Literal vs. Metaphorical

The primary meaning of *da* 打 is ‘hitting by hands or with an additional instrument’, and often refers to intentional actions and involves patients (e.g. *daren* 打人 “to beat someone”). Its meaning is extended to hitting in general (e.g. *damen* 打門 “to score goal”, “hit-door” in soccer) and actions involving physical contact (e.g. *dajing* 打井 “to dig a well”), as well as fighting (e.g. *dadou* 打鬥 “to fight”). All these senses involve physical actions though the function of the hand is no longer foregrounded in some cases. We consider them to convey the basic and literal meaning.

*Da* may also be used metaphorically as in *daji* 打氣 “hit-air”, which could mean cheering on somebody, in addition to literally “pumping air”. Metaphor involves an understanding of one conceptual domain in terms of another (Lakoff & Johnson 1980). Thus, metaphorical meaning of *da* could be characterized as the extension from physical actions to other conceptual domains, such emotion and culture, i.e. for physical pumping of air to the injection of emotional support to encouragement. The same extension can also be observed in *daji* 打擊 “(lit.) hit-strike”. Literally, *daji* means physical striking (e.g. *daji bangqiu* 打擊棒球 “hit a baseball”), but it could undergo metaphorical extension to mean affecting someone emotionally (e.g. *daji xinqing* 打擊心情 “hit-heart status” “affect one’s feeling”).

#### 3.2 Word structure

*Da* may be in initial position (*da-X*) or final position (*X-da*) in compound words. A fundamental difference between them is in their degree of metaphorization. It is noteworthy that most (87%) of the *X-da* cases retain literal meaning among the 354 entries, e.g. *ouda* 毆打 “to

beat up, (lit.) beat-hit”, *duda* 毒打 “to beat cruelly, (lit.) poison-hit”, with rare exception of metaphorical usage (only 13%) like *zhida* 主打 “promote, (lit.) main-hit”, *yanda* 嚴打 “to severely crack down, (lit.) severe-hit”. By comparison, few *da-X* cases do. Among the 812 *da-X* entries only 31% retain literal meaning (Type I) like *dadou* 打鬥 “to fight, (lit.) hit-fight”, *dagu* 打鼓 (“to drum, (lit.) hit-drum”). 45% are fully metaphorized (Type III) such as *daya* 打壓 “suppress, (lit.) hit-press”, and 24% may be used both literally and metaphorically (Type II), i.e. *daci* 打氣 “(lit.) pump air” and “(metaph.) cheer on”.<sup>9</sup> The asymmetry is summarized in Table 2:

	entries	Type I	Type II	Type III
<i>X-da</i>	354	87%	8%	5%
<i>da-X</i>	812	31%	24%	45%

Table 2. Distribution of 3 types by word structure

This asymmetry can also be appreciated from a minimal pair: *jida* 擊打 “(lit.) strike-hit” and *daji* 打擊 “(lit.) hit-strike”. *Jida* can only mean literally physical striking, e.g. *jida luogu* 擊打鑼鼓 “strike gongs and drums”, whereas *daji* may additionally have metaphorical meaning of affecting someone emotionally or cracking down on something, e.g. *daji zuian* 打擊罪案 “strike at crimes”.

### 4 Longitudinal and regional variations in metaphorization of *da*

#### 4.1 Incipient longitudinal variations

Developments of incipient metaphorization (Type II) verbs could be traced by comparing their metaphorical usage in the two periods 1995-2000 and 2011-2016. Some entries have increased metaphorical usage over time, indicating a shift to metaphorical type (III).<sup>10</sup> As an example, *datong*

<sup>9</sup> *Da-X* may be further classified into mainly three groups in terms of the nature of X: X could be objects (*dagu* 打鼓 “to drum”, *daci* 打氣 “(lit.) pump air”), resultative or directional complements (*dasi* 打死 “beat to death”, *dajin* 打進 “compete to enter”) or verbs (*dadou* 打鬥 “to fight”, *dazao* 打造 “(lit.) forge, fabricate”), symbolized as VO, VC and VV respectively. While VO (42%) and VV (64%) have the highest proportion of Type III, VC is distinct from them in having the largest proportion of Type II (65%).

<sup>10</sup> About half have undergone no significant change, e.g. *dapo* 打破 “(lit.) break (a vase) vs. (metaph.) break (a deadlock)”. A



打通“(lit.) hit-through” could mean literally “to open up passage by hitting or digging” as in (1) below, and could also mean metaphorically “to connect” as in (2). The overall metaphorical usage of *datong* has been doubled from 29% to 60% in the two periods. Notably, there are latitudinal variations which have resulted from longitudinal developments. As shown in Figure 1, BJ has the greatest increase in metaphorical usage (+58%), followed by TW (+19%), with HK having a slight decrease (-6%).

- (1) *Zhongyu datong suidong jiuchu san-ming-gongren* 終於打通隧洞救出三名工人 “Finally dug through a tunnel and rescued three workers.” (1995)
- (2) *Datong yu shehui dazhong de lianxi* 打通與社會大眾的聯繫 “Punched through the connection to the public in society.” (2016)

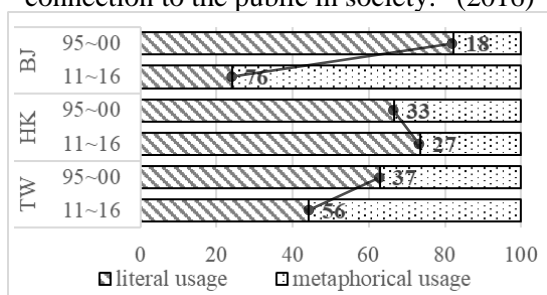


Figure 1. Change in usage of *datong* 打通

Another example of a shift to full metaphorical usage is *daxiang* 打響“(lit.) hit-loud”. Literally, it could mean “to make noise by gun firing” as in (3). Metaphorically, it could mean “to increase one’s popularity” as in (4). Different from *datong* 打通, *daxiang* 打響 has almost shifted to full metaphorical usage. The overall percentage of metaphorical usage has increased from 82% to 92% in the two periods. The results of regional variations are shown in Figure 2. In BJ, the metaphorical usage of *daxiang* has increased from 72% to 85% in the two periods. In contrast, HK and TW have maintained around 95% metaphorical usage in the two time periods. This shows that while *daxiang* in BJ is still on its way to gaining metaphorical usage, HK and TW have almost completed the final stage of metaphorization.

few have gained literal usage, e.g. *daji* 打擊“(lit.) hit (a baseball) vs. (metaph.) affect (one’s feeling)”, which could be attributed to common popularity of baseball in Taiwan.

- (3) *Nanchang Qiyi daxiang di-yi-qiang de zhandou didian* 南昌起義打響第一槍的戰鬥地點 “The place where they fired the first shot in the Nanchang uprising” (1997)
- (4) *Bing xunsu zai meishiquan-zhong daxiang zhimingdu* 並迅速在美食圈中打響知名度 “And rapidly struck up popularity in the restaurant trade.” (2015)

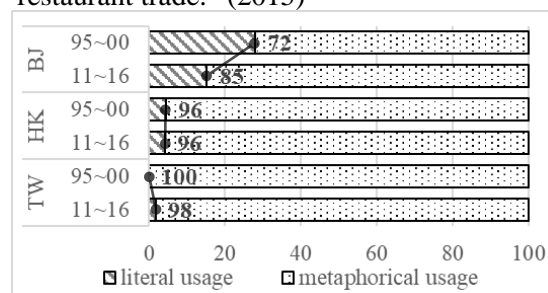


Figure 2. Change in usage of *daxiang* 打響

*Dazhang* 打仗 “hit-conflict”,<sup>11</sup> on the other hand, shows the initial stage of metaphorization. It could literally mean “to fight a war” as in (5), or metaphorically mean “to compete” as in (6). *Dazhang*’s metaphorical usage has increased from 5% to 12% in the two periods. Figure 3 shows that while *dazhang* retains full literal usage in BJ (0% metaphorical usage), it has increasing metaphorical usage in HK (+15%) and TW (+14%). This could point to the initiation of metaphorization in HK and TW but not in BJ.

- (5) *Budui xingjun dazhang* 部隊行軍打仗 “The troops are transferred to fight in war.” (1998)
- (6) *Gai-ju bei zhi nalai gen Gangshi kaitaiju duihan dazhang* 該劇被指拿來跟港視開台劇對撼打仗 “This show was said to be taken to compete against HKTV’s show.” (2014)

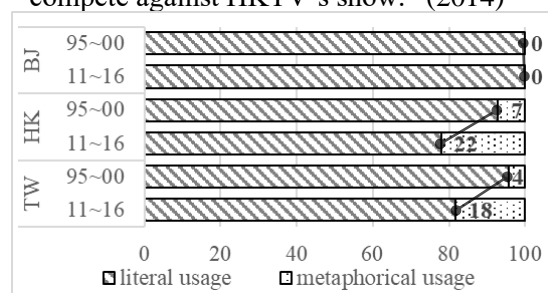


Figure 3. Change in usage of *dazhang* 打仗

<sup>11</sup> Note that *zhang* 仗 means “conflict” only in collocation with *da* 打 “hit”. For example, *da* in *dabizhang* 打筆仗 “to fight a pen battle” cannot be replaced by *fadong* 發動 “launch”, in contrast with *zhan* 戰, which means “war/battle” by itself, e.g. *fadongbizhan* 發動筆戰 “to wage a pen battle”.

## 4.2 Incipient regional variations

The overall latitudinal variations among the three communities are also significant. The most frequent entry *daji* 打擊 “(lit.) hit-strike” may be used literally to refer to physical hitting in (7) and metaphorically as affecting someone emotionally as in Sent. (8). Remarkably, BJ, HK and TW differ in the overall percentage of metaphorical usage of *daji* in the two time periods. Almost all usage in HK are metaphorical (99%), while by comparison, there are about 88% in BJ and about 73% in TW.

- (7) *Junfang yi jue ding zhe kuan diduidi daji de duoguan huojian jiang bu-zai qianjin bushu waidao* 軍方已決定這款地對地打擊的多管火箭將不再前進部署外島 “The military had decided that this type of land-to-land multi-tube strike rocket would no longer be deployed to the outer islands.” (2011)
- (8) *Gai ming pengyou zai jijin chushi hou jingshen da shou daji* 該名朋友在基金出事後精神大受打擊 “Following the incident with the Foundation, this friend was emotionally stricken.” (2012)

Another prominent example is *dadio* 打掉 “(lit.) hit-drop”. It literally means “to let something drop through physical striking”, as in (9). It may also be used metaphorically to mean “wipe out some (illegal) parties” as in (10). It displays great regional variations with BJ having almost full metaphorical usage (98%) while 79% for HK and 39% for TW.

- (9) *que bei hui quan dadao zuolian, lian shouji ye bei dadio* 卻被揮拳打到左臉，連手機也被打掉 “Rather his left face was struck by the swinging fist, even his handphone was struck off.” (2011)
- (10) *gongan bumen dadio fanzui tuanhuo 1230 ge* 公安部門打掉犯罪團夥 1230 個 “The public security bureau struck down 1230 criminal gangs.” (2012)

Among the top 20 most frequent incipient words (covering 95% of all tokens of common incipient words), BJ stands out by having the largest number of words (12) with top percentage of metaphorical usage. HK, by comparison, only has 6, and TW has even fewer, only 4:<sup>12</sup>

**BJ (12/20):** *dazao* 打造 “hit-make”, *dapo* 打破 “hit-break”, *dakai* 打開 “hit-open”, *dachu* 打出 “hit-out”, *daxia* 打下 “hit-down”, *datong* 打通 “hit-through”, *dajia* 打架 “hit-fight”, *dajiaodao* 打交道 “hit-contact-road”, *dasao* 打掃 “hit-sweep”, *daduan* 打斷 “hit-break”, *dadio* 打掉 “hit-drop”, *dazhong* 打中 “hit-at”

**HK (6/20):** *daji* 打擊 “hit-strike”, *daru* 打入 “hit-in”, *daqi* 打氣 “hit-air”, *dacheng* 打成 “hit-become”, *dazhang* 打仗 “hit-conflict”, *dadao* 打倒 “hit-collapse”

**TW (4/20):** *dajin* 打進 “hit-enter”, *daqi* 打氣 “hit-air”, *daxiang* 打響 “hit-loud”, *dazhong* 打中 “hit-at”

These findings on regional variations could be confirmed by the metaphorical usage of the top 20 incipient words in each community. Figure 4 shows that BJ again has the highest percentage (71.8%), followed by HK (69.6%) and TW has the smallest (66.8%). These findings point to BJ having a greater tendency than either HK or TW to use common words metaphorically.

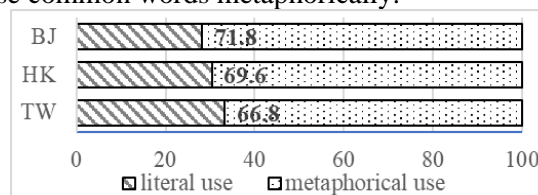


Fig 4. Metaphorical usage of top 20 Type II entries

## 4.3 The extent of metaphorization in three Chinese communities

The important incipient Type II verbs with variable entries have provided the basis for the study on the variations in metaphorization across the three communities. The Metaphorization Index (MI) we have proposed characterizes the extent of metaphorization. When aggregated for the top 50 common entries in each community as shown in Figure 5, BJ has the highest MI (81.6%), closely followed by HK (80.2%), with TW the lowest (73.3%). It shows that both BJ and HK have a greater tendency to use words metaphorically than TW. This supports the finding in the previous section and underscores even more striking overall differences among the three communities.

<sup>12</sup> *Daji* 打氣 has the same percentage in HK and TW while *dazhong* 打中 has the same percentage in BJ and TW.

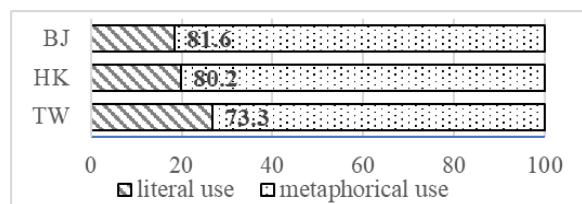


Figure 5. MI for top 50 common *da*-X verbs

#### 4.4 Predominant local entries in three Chinese communities

In contrast to comparing the MI of the common incipient entries in the three communities, we also examine verbs found only in a single community, i.e. top most frequency ( $\geq 90\%$ ) local verbs with metaphorical usage in each community. Among the top 200 entries in each community, HK stands out with 19.5% predominant metaphorical entries, which is even larger than the sum of BJ (5.5%) and TW (10.5%):

**BJ (5.5%):** e.g. *dafen* 打分 “to score, (lit.) hit-points”, *daxin* 打新 “staging (to buy initial round of publicly offered stocks), (lit.) hit-new”, *danian* 打蔫 “being tired, (lit.) hit-wither”, etc.

**HK (19.5%):** e.g. *dajiu* 打救 “to rescue, (lit.) hit-save”, *dazhengqihao* 打正旗號 “under the name of, (lit.) hit-right-flag”, *dashuding* 打書釘 “to read at bookstores without purchase, (lit.) hit-book-nail”, etc.

**TW (10.5%):** e.g. *dalian* 打臉 “to offend or upset someone/ give a slap in the face, (lit.) hit-face”, *dapi* 打屁 “to chit-chat, (lit.) hit-fart”, *dashu* 打書 “to promote new books, (lit.) hit-book”, etc.

It shows that HK has a higher degree of metaphorization when it comes to items that are unique to their own community. This could be a source of new metaphorical *da* 打 in the Chinese language as a basis of internal language contact.

## 5 Concluding remarks

### 5.1 Metaphorization and light verb development

Light verbs are distinct from regular verbs in terms of both *grammatical* and *semantic* properties. For instance, English *give* could only take concrete nouns in its regular ditransitive usage with the meaning of “offering something to someone”. In terms of light verb usage (e.g. *give a pull*, *give a wink*), however, *give* may take eventive nouns with

bleached meaning of performing actions. In Chinese, previous studies had often focused on the *grammatical* properties of light verbs (D. Zhu 1985, Diao 2004, *i.a.*). However, the process of *semantic* bleaching has been rarely studied, especially in the context of metaphorization.

Based on 2175 entries with *da* 打, the current study shows a striking asymmetry in metaphorization on the basis of word structure. *Da* in word-final position (*X-da*) tends to be used literally such as *jida* (*luogu*) 擊打(鑼鼓) “strike gongs and drums”, whereas in word-initial position (*da-X*) it tends to be used metaphorically such as *daji* (*zuian*) 打擊(罪案) “strike at crimes”. We observe that *grammatical* structure is correlated with *semantic* bleaching in terms of metaphorization. We note that *X-da* is often manifested as a modifier-head structure where the modifier specifies the manner of *da* “hit”. Here, *ji* 擊 “strike” is a hyponym of *da* 打 “hit” and it specifies that the main action of hitting is accompanied by forceful striking, as opposed to other hitting actions such as *guoda* 擰打 “to slap” or *chuida* 捶打 “to pound”. As noted, in *jida* “strike-hit” *da* is the main action modified by *ji* “strike” which denotes a subset of hitting actions. This hierarchical relationship preempts metaphorization which involves an opposite relaxation of meaning. In contrast, *da* in *daji* “hit-strike” is not modified by *ji* but forms with it a *coordinate* VV structure of near-synonyms. Since *da*’s meaning is not hierarchically constrained but forms an equal partnership in VV structure, it could undergo semantic bleaching and metaphorization by becoming associated with other kinds of objects when there is an appropriate triggering agent.<sup>13</sup>

The correlation of grammatical properties and metaphorization is also reflected in regional variations. Our current study arrives at a conclusion similar to that of Jiang et al. (2016)

<sup>13</sup> Metaphorization may be triggered by quasi-concrete objects, such as (*dazao*) *qicheye de hangkongmujian* (打造) 汽車業的航空母艦 “(forge) an aircraft carrier of the automotive business” (Tsou, Chin & Kwong 2011). Notably, the metaphorization of *dazao* is achieved by relaxing the selectional restriction on concrete nouns to include abstract nouns through pseudo objects. This again shows an interaction of grammatical properties and metaphorization in light verb development and deserves further exploration.

which studied the grammatical variations of light verbs in BJ and TW. They noted that *jinxing* 進行 in TW may take VO complements, but not BJ. *Jinxing* in TW was thus regarded as “more transitive” and more verbal than in BJ. On the other hand, the current study has found that the MI for TW is lower than BJ and HK. This echoes the conclusion of Jiang et al. (2016) from a different perspective: while they found TW light verbs to be more conservative *grammatically*, this study has found them to be more conservative in *metaphorization* of *da*, which is a *semantic* process. The correlation of grammatical properties and metaphorization in light verb development seems to hold across Chinese speech communities.

## 5.2 Latitudinal and longitudinal variations based on a synchronous corpus, LIVAC

The current study is based on the LIVAC corpus which differs from other corpora through the adoption of a rigorous sampling (“Windows”) approach in curating data from six Chinese speech communities since 1995 (Tsou & Kwong 2015). Thus LIVAC could contribute to the study of longitudinal variations by addressing latitudinal variations across the 22-year time span. In particular, it offers an opportunity to view some longitudinal variations in light verb developments. In the case of *da*, it is observed that incipient metaphorization has developed at different rates in the three communities from 1995 to 2016. As an illustration, while metaphorization of *daxiang* 打響 “(lit.) hit-loud” may be completed in HK and TW with steady metaphorical usage (>95% over time), it is still ongoing in BJ as indicated by an increase in metaphorical usage from 72% to 85% over the same period. This kind of comparison has to rely on a rigorously curated corpus.

The efficacy of a rigorously cultivated corpus could also be demonstrated by comparing two other top light verbs in Chinese speech communities such as *jinxing* 進行 “proceed” and *zuo* 做 “make”, which have been studied in the literature. Under normal circumstance of natural societal equilibrium, it could be assumed that there would be no significant variation in usage, and their relative percentages within the two periods 1995-2000 and 2011-2016 should be about equal. However, the distribution of these two light verbs has varied drastically with different extent in BJ,

HK and TW. As shown in figure 6, the relative percentage of *jinxing* has undergone decrease in the three communities and HK’s drop is the sharpest (-43%), followed by TW (-32%) and Beijing the modest (-11%), accompanied by a corresponding increase of *zuo*. The change in relative percentage of *jinxing* and *zuo* is striking in terms of both the sharp rates and the regional variations. While the cause for this shift awaits careful exploration, it may be suggested that the formal register marked by *jinxing* is on the decline more prominently in HK and TW than BJ with a possible compensatory shift to the most colloquial *zuo*. A systematic study in terms of register and genre markers in LIVAC could provide fuller and better answers.

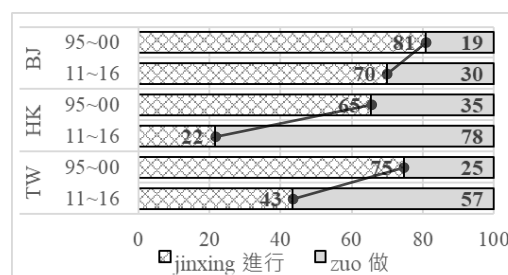


Figure 6. Relative percentage changes of *jinxing* 進行 and *zuo* 做

## 5.3 Towards the future

This paper examines a very common but little studied light verb *da* 打 in three Chinese speech communities: Beijing, Hong Kong and Taiwan, by monitoring a synchronous corpus. The findings offer some glimpses into the differential metaphorization of *da* in the three communities and also their longitudinal developments as well as into how and why metaphorization has come about, particularly the general interaction between grammatical structure and semantic bleaching. We expect this approach would facilitate further explorations of the underlying mechanisms of metaphorization by examining more Chinese dialects in which it is an ongoing process and other languages as well.

## Acknowledgements

For data preparation and annotation, we wish to thank Wing Fu Tsoi, Janice Chong, and especially Renee Ji. We also wish to thank the two anonymous reviewers for their valuable comments. All errors remain the authors’ own responsibilities.



## References

- Cattell, Ray. 1984. *Composite Predicates in English*. North Ryde, New South Wales: Academic Press Australia.
- Chae, Hee-Rahk. 1996. Light verb constructions and structural ambiguity. *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation (PACLIC 11)*: 99-107.
- Chen, Yongli. 2003. Xingshi dongci de fanwei, cilei ji tezheng [The scope, subcategories and features of dummy verbs]. *Academic Journal of Jinyang* 3: 92-94.
- Diao, Yanbin. 2004. *Xiandai Hanyu Xuyi Dongci Yanjiu* [A Study on Delexical Verb in Modern Chinese]. Dalian: Liaoning Normal University Press.
- Grimshaw, Jane, and Armin Mester. 1988. Light verbs and  $\theta$ -marking. *Linguistic Inquiry* 19: 205-232.
- Her, One-Soon, Wei-Tien Dylan Tsai, Jung-Hsing Chang, Kawai Chui, Jennifer M. Wei, and D. Victoria Rau. 2016. *Yuyan Ai-bu-ai? Yuyanxuejia de Kanfa* [Language Cancer: From the Perspective of Linguists]. Taipei: Linking.
- Huang, Chu-Ren, Jingxia Lin, Menghan Jiang, and Hongzhi Xu. 2014. Corpus-based study and identification of Mandarin Chinese light verb variations. In Marcos Zampieri, Liling Tan, Nikola Ljubešić and Jörg Tiedemann, eds., *Proceedings of the First Workshop on Applying NLP Tools to Similar Languages, Varieties and Dialects*, 1-10. Dublin: Association for Computational Linguistics and Dublin City University.
- Jespersen, Otto. 1954. *A Modern English Grammar on Historical Principles*. London: George Allen & Unwin & Copenhagen: Ejnar Munksgaard.
- Jiang, Haiyan. 2020. Chouxian dongzuo shijian yuyi fanchou yu Hanyu xingshi dongci [Abstract action event semantic category and Chinese dummy verbs]. Doctoral dissertation, Jilin University.
- Jiang, Menghan, Dingxu Shi, and Chu-Ren Huang. 2016. Transitivity in light verb variations in Mandarin Chinese – a comparable corpus-based statistical approach. *Proceedings of the 30th Pacific Asia Conference on Language, Information and Computation (PACLIC 30)*: 459-468.
- Lakoff, George, and Mark Johnson. 1980. *Metaphor We Live By*. Chicago and London: University of Chicago.
- Li, Feng, and Yun Chai. 1995. “V+yi” lei xuhua dongci de binyu [The objects of “V+yi” dummy verbs]. *Journal of Xinjiang Education Institute* 4: 68-71.
- Li, Hanlei. 2003. “Jiayi” de yuyong gongneng [The pragmatic functions of *jiayi*]. *Journal of Suzhou Education College* 1: 11-16.
- Lin, Jingxia, Hongzhi Xu, Menghan Jiang, and Chu-Ren Huang. 2014. Annotation and classification of light verbs and light verb variations in Mandarin Chinese. In Takuya Nakamura, ed., *Proceedings of Workshop on Lexical and Grammatical Resources for Language Processing*, 75-82. Dublin: Association for Computational Linguistics and Dublin City University.
- Lü, Shuxiang. 1980. *Xiandai Hanyu Babaici* [Modern Chinese 800 Words]. Beijing: Commercial Press.
- Lü, Shuxiang. 1999. *Xiandai Hanyu Babaici (Zengdingben)* [Modern Chinese 800 Words (Enlarged Edition)]. Beijing: Commercial Press.
- Mao, Hongyuan. 1997. Hua xingshi dongci he xingshihua dongci [On formal verbs and formalized verbs]. *Journal of Kashgar Teachers College* 4: 78-79.
- Pham, Hoa. 1999. Cognate objects in Vietnamese transitive verbs. *Toronto Working Papers in Linguistics* 27: 227-246.
- Ren, Fengmei. 2013. A grounding approach to the semantic meaning of the light verb *da*. In Pengyuan Liu and Qi Su, eds., *Chinese Lexical Semantics: 14th Workshop, CLSW 2013, Zhengzhou, China, May 10-12, 2013. Revised Selected Papers*, 88-96. Berlin & Heidelberg: Springer-Verlag.
- Shen, Jiaxuan, and Jiangzhi Zhang. 2013. Ye tan xingshi dongci de gongneng [On the grammatical function of dummy verbs in Chinese]. *TCSOL Studies* 2: 8-17.
- Su, Lei. 2009. Lun “da” de yufahua [On the grammaticized process of the word “fight”]. *Journal of Hubei University of Education* 5: 30-31.
- Tsai, Wei-Tien Dylan. 2017. Jiwuhua, shiyong jieyou yu qingdongci fenxi [Transitivity, applicative construction and light verb analysis]. *Contemporary Research in Modern Chinese* 19: 1-13.
- Tsou, Benjamin K., and Oi Yee Kwong. 2015. LIVAC as a monitoring corpus for tracking trends beyond linguistics. *Journal of Chinese Linguistics Monograph Series* 25: 447-471.
- Tsou, Benjamin K., Andy C. Chin, and Oi Yee Kwong. 2011. From synchronous corpus to monitoring

- corpus, LIVAC: The Chinese case. In Friedrich Laux and Lena Strömbäck, eds., *The Third International Conference on Advances in Databases, Knowledge, and Data Applications (DBKDA 2011)*, 175-180. St. Maarten: IARIA XPS Press.
- Wang, Leslie Fu-mei. 2004. A corpus-based study of mandarin verbs of doing. *Concentric: Studies in Linguistics* 30.1: 65-85.
- Wang, Li. 1985. *Zhongguo Xiandai Yufa* [Modern Chinese Grammar]. Beijing: Commercial Press.
- Yan, Zhongsheng. 1998. Shuo “houxu dongcixing binyu dongci” [On “verbs taking verbal objects”]. *Journal of Hebei Normal University (Natural Science)* 2: 91-93.
- Zhou, Gang. 1987. Xingshi dongci de cifenlei [Subdivision of dummy verbs]. *Chinese Language Learning* 1: 11-14.
- Zhu, Dexi. 1985. Xiandai shumian hanyu li de xuhua dongci he mingdongci [Dummy verbs and nominal verbs in Modern Written Chinese]. *Journal of Peking University (Humanities and Social Sciences)* 5: 1-6.
- Zhu, Jian-jun. 2004. “Da-V” zhi “da” de yufahua tanxi [The grammaticalization of *da* in the structure of *da-V*]. *Research in Ancient Chinese Language* 3: 38-44.
- Zhuang, Huibin. 2014. Xiandai Hanyu Qingdongci “da” de lai yuan chuyi [On the etymology of the light verb *da* in Chinese]. *Language Teaching and Linguistic Studies* 3: 67-74.