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

Benjamin K. Tsou¹, Ka-Fai Yip²

¹City University of Hong Kong, ¹The Hong Kong University of Science and Technology; ²Yale University

¹btsou99@gmail.com, ²kafai.yip@yale.edu
Light verbs vs. lexical verbs

1. *give* a pull vs. *give* a book

   Others: *make*, *take*, *have* ...

2. 做研究 “to do research” vs. 做蛋糕 “make a cake”

3. 給予解釋 “give an explanation” vs. 給他一支筆 “give him a pen”

4. 打造光輝未來 “create a bright future” vs. 打造兵器 “forge a weapon”

Grammatical properties: abstract nouns vs. concrete nouns

Semantic properties: “bleached” meaning vs. lexical meaning
Research focus

1. An understudied yet very common light verb: da 打
   ◦ Only several da compound verbs are briefly discussed (Wang 1985, Ren 2013)

2. Metaphorization and semantic bleaching
   ◦ Key process in light verb development
   ◦ Mainly grammatical studies in the literature

3. New approach to measure Variations in Chinese language
   ◦ Mainly Mainland China and Taiwan in the literature
   ◦ Hong Kong?
Roadmap

- Corpus base LiVaC
- Metaphorization and word structure
- Variations in three Chinese speech communities
- Implications
Corpus base LiVaC
Corpus base: **LiVaC** *(Linguistic Variation in Chinese Speech Communities)*


More than 600 million characters of news media texts, rigorously curated with assisted verification, from Pan-Chinese communities:

Beijing, (Guangzhou,) Hong Kong, Macau, Shanghai, (Shenzhen,) Singapore, Taiwan, etc.
Corpus base: LiVaC da entries

LiVaC
2175 da 打 entries
(1995-2016)

Taiwan’s Guoyu Cidian
774 da entries

Xiandai Hanyu Cidian
現代漢語辭典
238 da entries

Verbs
Free-standing da;
Nouns
Loanword-entries...
Da+X entries: 1730
812
X+Da entries: 445
354
1166

198 Ovlp.

284 Ovlp.
Metaphorization and word structure
Metaphorization

- Major process in grammaticalization (e.g. *is going to*, Bybee & Pagliuca 1985)
- Involves an understanding of one conceptual domain in terms of another (Lakoff & Johnson 1980)

- Working definition for *da* 打:
  - **Literal** meaning: physical domain
    - *daji bangqiu* 打擊棒球 “hit a baseball”
  - **Metaphorical** meaning: other conceptual domains, e.g. emotion
    - *daji xinqing* 打擊心情 “(lit.)hit-heart status, emotionally striken”
Three types 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 II (Incipient Metaphorization):**
- 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”.

**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”
Distribution of 3 types by word structure

<table>
<thead>
<tr>
<th>entries</th>
<th>Type I</th>
<th>Type II</th>
<th>Type III</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-da 打</td>
<td>354</td>
<td>87%</td>
<td>8%</td>
</tr>
<tr>
<td>da 打-X</td>
<td>812</td>
<td>31%</td>
<td>24%</td>
</tr>
</tbody>
</table>

- "jida" 撃 打 “(lit.) strike-hit” vs. "daji" 打 撃 “(lit.) hit-strike”
- Jida (Type I): Only allows literal meaning, e.g. jida luogu 撃打鑼鼓 “strike gongs and drums”
- Daji (Type II): Allows metaphorical meaning, e.g. daji zuian 打擊罪案 “strike at crimes”.
Word structure

◦ X-da: Modifier-Verb 偏正/狀中 (e.g. jida 擊打 “(lit.) strike-hit”)
◦ Da-X:
  ◦ (i) Verb-Verb 並列 (e.g. dazao 打造 “(lit.) forge, fabricate”)
  ◦ (ii) Verb-Object 動賓 (e.g. daci 打氣 “(lit.) pump air”)
  ◦ (iii) Verb-Complement 動補 (e.g. dajin 打進 “compete to enter”)

◦ Take home message:
  ◦ Metaphorization is sensitive to word structure
Variations in three Chinese speech communities

BEIJING, HONG KONG, TAIWAN
## Distribution of da-X entries

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Types (%)</th>
<th>Tokens (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIVAC</td>
<td>812 (100)</td>
<td>318,095 (100)</td>
</tr>
<tr>
<td>BJ, HK or TW</td>
<td>672 (83)</td>
<td>179,182 (56)</td>
</tr>
<tr>
<td>BJ, HK and TW</td>
<td>151 (19)</td>
<td>170,245 (54)</td>
</tr>
</tbody>
</table>

**Type II** entries: 48 (32%) types, 102,277 (60%) tokens → Predominant
Sampling Type II sentences

- Dual usage → must examine authentic sentences

- 1st period: 1995-2000
- 2nd period: 2011-2016

- At most 50 sentences per year for each word in each community
- Total 16955 sentences have been obtained and tagged
- Percentage of metaphorical usage
Metaphorization Index (MI)

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

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

- Aggregated for each community
- Enable quantitative comparison
Longitudinal variations: 
\textit{datong} 打通 “hit-through”

- BJ shows rapid metaphorization, followed by TW, with HK inert
Longitudinal variations: `daxiang` 打響 “hit-loud”

- Metaphorization is almost completed in HK and TW, but is still ongoing in BJ.
Longitudinal variations: *dazhang* 打仗 “hit-conflict”

- Metaphorization is initiated in HK and TW, but not in BJ
Incipient regional variations: Top 20 Type II da

- Top percentage of metaphorical usage among three communities

BJ (12/20):
dazao 打造 “hit-make”, dapo 打破 “hit-break”, dakai 打开 “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 打断 “hit-break”, dadiao 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”
Overall incipient regional variations
Metaphorical usage of top 20 Type II entries

12 years: 95-00,11-16
Variations in metaphorization: Aggregated MI for top 50 common da-X verbs

22 years: 95-16
## Predominant local entries

- Top most frequency (≥90%) local verbs with metaphorical usage, among top 200 entries in each community

<table>
<thead>
<tr>
<th>Community</th>
<th>Top 3 Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BJ (5.5%)</strong></td>
<td>① dafen 打分 “to score, (lit.) hit-points”, ② daxin 打新 “staging (to buy initial round of publicly offered stocks), (lit.) hit-new”, ③ danian 打蔫 “being tired, (lit.) hit-wither”, etc.</td>
</tr>
<tr>
<td><strong>HK (19.5%)</strong></td>
<td>① dajiu 打救 “to rescue, (lit.) hit-save”, ② dazhengqihao 打正旗號 “under the name of, (lit.) hit-right-flag”, ③ dashuding 打書釘 “to read at bookstores without purchase, (lit.) hit-book-nail”, etc.</td>
</tr>
<tr>
<td><strong>TW (10.5%)</strong></td>
<td>① dalian 打臉 “to offend or upset someone/ give a slap in the face, (lit.) hit-face”, ② dapi 打屁 “to chit-chat, (lit.) hit-fart”, ③ dashu 打書 “to promote new books, (lit.) hit-book”, etc.</td>
</tr>
</tbody>
</table>
Variations in three Chinese speech communities

- Take home message:
  - Beijing and Hong Kong show a higher degree of metaphorical shift in different ways
    - BJ: items common to the three communities
    - HK: items unique to their own community
      - could be a source of new metaphorical usage when the items enter other communities through internal language contact
  - Taiwan: the most conservative in terms of metaphorization
Implications
Correlation between semantic and grammatical properties

- For light verbs, the two seem to be correlated: semantic bleaching & subcategorization

- (1) Metaphorization is sensitive to word structure

- (2) Across Chinese speech communities, TW is the most conservative in terms of metaphorization, AND grammatical properties
  - Jiang et al. (2016): Taiwan light verbs are more transitive (“verbal”) than Mainland, and thus are less grammaticalized
Other light verbs:

*Jinxing* 進行 “proceed” & *Zuo* 做 “make”

- Show significant latitudinal and longitudinal variations in LiVaC (tokens)

![Graph showing variations in light verbs across regions]

- How about metaphorization?
  - E.g. 做研究 “to do research” vs. 做蛋糕 “make a cake”
Concluding remarks

◦ Filling a gap: Light verb *da*

◦ General mechanism of metaphorization and grammatical developments

◦ Dynamics of language variation in Chinese language
References


References

References


Thank you!

ANY COMMENTS ARE WELCOME!