The Polysemy of Da3: An ontology-based study

Jia-Fei Hong¹, Kathleen Ahrens¹, Chu-Ren Huang²
¹ Graduate Institute of Linguistics, National Taiwan University
² Institute of Linguistics, Academia Sinica
¹ No. 1, Sec. 4, Roosevelt Road, Taipei, Taiwan, R.O.C, 106
² No. 128, Section 2, Academia Road, Taipei, Taiwan, R.O.C, 115
jiafei@gate.sinica.edu.tw; kathleenahrens@yahoo.com; chruen@gate.sinica.edu.tw;

0. Abstract
In this study, we explore all possible concepts for physical activity senses and metaphorical event senses of da3 through the SUMO ontological concept system. We use SUMO to find out their concepts in order to group and cluster these same concept words. For that reason, we can get different concept groups for the physical activity senses or metaphorical event senses of da3. We also compare our analysis with Gao’s study (2001) and find our results are very similar. This leads us to propose that a concept-base approach is a viable one when exploring the sense of polysemous verbs in Chinese.

Keyword: Concept, SUMO, Da3, Polysemy

1. Introduction
In this study, we explore all possible concepts for physical activity and metaphorical event senses (excluding other idiomatic senses and extension senses) of da3 through the Suggested Upper Merged Ontology (SUMO) ontological concept system (Huang et al. 2004). According to previous work (Gao 2001), da3 is a basic verb in the domain of physical action verbs in Chinese, as 1) it refers to the most basic action of the hand; and 2) at the same time it can refer to a wide-range of actions or events that involve physical
contact of one kind or another. Also da3 is polysemous to a very high degree. Its prototypical meaning reaches a high frequency of occurrence in the empirical corpus data. In this paper, we will compare Gao’s (2001) analysis with the analysis we provide based on SUMO.

First, we will collect data from Sinica Corpus and check senses from Chinese Wordnet. Next, we analyze these physical activity and metaphorical event senses of da3 in SUMO (Huang et al. 2004) to find all possible concepts and categorize them. Finally, we analyze these concepts for semantic features in order to compare our analysis with the analysis in Gao’s study (2001).

2. Previous research

Regarding verb studies, previous research has focused on VV compound verbs in Modern Chinese (Hong and Huang, 2004; Li, 1990), or on near synonyms in Modern Chinese (Chief et al, 2000; Huang et al. 2000; Liu 2002; Tsai, 2002; Huang and Hong, 2005). These studies focused on the semantic relations between lemmas. However, some studies were focused on different senses or relations in the same lemma. Also, some scholars have worked on da3 polysemy analyses. Da3 is one of the most frequently used verbs, as it is ranked 16 in the list of most frequently used verbs in Chinese (Bei and Zhang, 1988). Gao (2001) explored the semantic properties of da3 and its prototypical meaning and categorized its semantic representations to show the systematic patterning of its meaning extensions.

3. Motivation and Goals

Language knowledge representation is a manifestation of the systematic contrasts found in human communication, which defies conventional description. Take modal verbs as an example. Modal verbs have similar semantic functions and cannot be easily distinguished in terms of their lexical senses. Therefore, they are considered to be interchangeable. Nevertheless, it is not uncommon that this kind of polysemy always has contrasts in usage, as we can see the contrast between hui4 (know) and hui4 (can) below:
The Polysemy of Da3: An ontology-based study

(1a) 台灣廠商到德國開展時，非常需要會德語和中文的人，做為溝通橋樑。
Taiwan factory to Germany exhibit time,

When Taiwan factories exhibit in Germany, they really need people who could speak German and Chinese to communicate with other persons.

(1b) 在上課的時候，他不會講德語或中文和學生溝通，因為他怕自己沒辦法完整表達意思。
In class time, he will not speak German or Chinese with students communicate, because he worry himself no way complete express meaning.

“In class, he can’t communicate with his students in German or Chinese, because he worries that he can’t completely express his meaning (in these languages).”

Considering the lexical sense of modal verb polysemy and its natural language use, hui4 means both “know” or “can”. As defined, we notice that they differ from each other, even though they share similar concepts.

This paper will investigate the lexical semantic relations between each sense of da3
polysemy (excluding idiomatic senses and extended senses) by studying its sense distinctions, word formation collocation, and distribution pattern.

4. SUMO

In this study, we use Suggested Upper Merged Ontology (SUMO) to analyze all concepts for da3. We find out all possible concepts and divide them into different categories.

WordNet is inspired by current psycholinguistic and computational theories of human lexical memory (Fellbaum (1998), Miller et al. (1993)). English nouns, verbs, adjectives, and adverbs are organized into synonym sets, each representing one underlying lexicalized concept. Different semantic relations link the synonym sets (synsets). The version of WordNet that Sinica BOW implemented is version 1.6, with nearly 100,000 synsets.

In Sinica BOW, each English synset was given up to the 3 most appropriate Chinese translation equivalents. In cases where the translation pairs are not synonyms, their semantic relations are marked (Huang et al. 2003). The bilingual WordNet is further linked to the SUMO ontology. We use the semantic relations in bilingual resource to expand and predict domain classification when it cannot be judged directly from a lexical lemma.

5. Data collection

From Sinica Corpus and Gigaword Corpus, we find several patterns for da3. According to the Chinese Wordnet Group analysis (Huang et al., 2003), there are 114 senses which include physical activity senses, metaphorical events, idiomatic senses and extended senses. In this study, we would like focus on physical activity senses, and metaphorical event senses, but not on idiomatic senses and extened senses. The sense number distribution of da3 was shown as below Table 1.

<table>
<thead>
<tr>
<th>Category</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical sense</td>
<td></td>
</tr>
<tr>
<td>打桌子 41 (35.96%)</td>
<td>114</td>
</tr>
<tr>
<td>Metaphorical sense</td>
<td></td>
</tr>
<tr>
<td>打交道 33 (28.95%)</td>
<td></td>
</tr>
<tr>
<td>Other meaning extensions</td>
<td></td>
</tr>
<tr>
<td>打失 40 (35.09%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The sense number distribution of da3 in Chinese Wordnet
Also, we take the analysis result of da3 from Chinese Wordnet to check manually the sense distribution in Sinica Corpus and GigaWord Corpus. Therefore, we can observe the empirical presentations of da3 in the context. Table 2 shows the sense distribution in Sinica Corpus and GigaWord Corpus.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Physical Sense</th>
<th>Metaphorical Sense</th>
<th>Other meaning extensions</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sinica Corpus</td>
<td>941 (63.97%)</td>
<td>205 (13.94%)</td>
<td>325 (22.09%)</td>
<td>1,471</td>
</tr>
<tr>
<td>GigaWord Corpus</td>
<td>37,098 (48%)</td>
<td>12,366 (16%)</td>
<td>27,824 (36%)</td>
<td>77,288</td>
</tr>
</tbody>
</table>

Table 2: The sense distribution in Sinica Corpus and GigaWord Corpus

6. Data analysis
The analysis, given in Chinese Wordnet (Table 3 and Table 4), is based on the Sinica Corpus, Gigaword Corpus and the criteria proposed by Huang et al. (2003) to differentiate the lexical meanings, presents several different senses for da3.

<table>
<thead>
<tr>
<th>打</th>
<th>da3</th>
</tr>
</thead>
<tbody>
<tr>
<td>詞義 1:【及物動詞，VC】以手施力使手或手持物撞擊特定對象。{tap, 01496422V}</td>
<td></td>
</tr>
</tbody>
</table>
| 例句：男子說完用力的<打>一下桌子，整個桌子馬上變成木屑一片。
    例句：他用扇骨用力<打>門板好幾下，然後說：我是如此的人嗎？專程來抓你的小辮子？ |
| 詞義 2:【名詞，nom】用手或手持物打後述對象，使其感到疼痛或受到傷害。{hit, 00960484V} |
| 義面 1:【及物動詞，VC】用手或手持物打攻擊後述對象，使其感到疼痛或受到傷害。
    例句：因為他<打>人也是為了完成國家任務，說清楚群眾是會諒解的。
    例句：母親忽然沉下臉<打>他一下手背，並告誡他不能指月亮娘娘，會爛耳朵的。 |
| 義面 2:【名詞，nom】用手或手持物打攻擊後述對象，使其感到疼痛或受到傷害。
    {hit, 00960484V} |
例句：說謊一臉狡猾，輕輕的對我耳語道：主人，你如果認錯，至少要挨一頓打。

例句：這兒的老師也常在學生不乖或考試成績不好時祭出「班法」，給學生一頓好打，這一點也令我難以接受。

詞義 3：【及物動詞，VC】易碎物品因受到撞擊使破碎。
例句：在高速公路不知哪裡天上飛來一石，把擋風玻璃打成一個小凹洞。
例句：廟裏的人把碗打了，他心裡都不好過，因為一旦它被破壞掉的時候，那個生命體要釋放出來，它沒有完成它的生命里程，就無處可去。

Table 3: The analysis of physical activity senses of da3 from Chinese Wornet

詞義 11：【及物動詞，VC】比喻建立基礎，常與「基礎」、「基石」連用。同義詞「打下（0600）」。
例句：英國球隊在這方面因為基礎打得很好，後面的訓練比賽也就有了更好的體能保証。
例句：這一群功夫底子很深的訓練員，在正式表演的時候，都會有很好的表現，不需要擔心。

詞義 12：【及物動詞，VC】比喻建立，常與「名聲」、「成果」連用。同義詞「打下（0700）」。
例句：全市警察機關和全體民警要充分認識這次活動的重大意義，統一思想，全警動員，打出一番成果，打出一片聲威。
例句：南方網訊綜藝大哥大胡瓜又要出唱片了，不過這回卻是鎖定要在大陸發行，胡瓜笑說，唱片賣得好不好不重要，最重要是藉機大打在大陸的知名度。

詞義 13：【及物動詞，VC】比喻在商業上針對特定對象或管道經營。同義詞「建立」、「立（1210）」、「樹（0210）」。
例句：Vitria 的氣液應用系統整合軟體在業界已經是領導品牌，張若玫首先在歐美打市場，近半年才積極進軍亞洲。
例句：明基電通認為未來市場須多元化，在開發中國家打年輕人市場，因為年輕人口充沛，在已開發國家也要打銀髮族。
Among these 41 physical activity senses of da3, we divide them into two main physical activity senses for da3: 1) hit and 2) pump such as below:

(2) a. 母親忽然沉下臉 <打>他一下手背，並告誡他不能指月亮娘娘，會爛耳朵的。
   
   統 qin1  hui1  ran2  chen2  xia4  lian3  da3  ta1  yi2  xia4  shou3  bei4,
   
   Bi1  gao4  jie4  ta1  bu4  neng2  zhi3  yue4  liang4  nian2  nian5,
   
   and  warn  he  can’t  point  moon  queen,
   
   bing4  gao4  jie4  ta1  shou3  bei4,
   
   and  warn  him,
   
   “His mother suddenly becomes long-faced and hits him on the back of his hand, then warned him that he will get rotting ears if he points to the moon.”
   
   b. 這時每一鞭都如 <打> 在她的身上一般痛楚。
   
   Zhe4  shi2  mei3  yi1  bian1  dou1  ru2  da3  zai4  ta1  de5,
   
   This time  every  whipped  all  like  whip  she  MOD
   
   shen1  shang4  yi1  ba1  tong4  chu3
   
   body  general  pain
   
   “At this moment, every whipped whip pains her like she is being whipped.”
   
   (3) 爲了防止車輛陷進沙地，不要把輪胎氣 <打> 太足。
   
   Wei4  le5  fang4  zhi3  che1  liang4  jin4  sha1  di4,
   
   For  preven  cars  tuck  into  sand,  don’t  let
   
   lun4  tai1  qi4  da3  de2  tai4  zu2
   
   tires  gas  pump  too  full
“Do not pump the tires too full to avoid the cars being stuck in the sand.”

Moreover, in “hit” sense of da3, we also can divide this sense into two different categories: 1) hand and hand holdings and 2) force and impact. Then, in the second category, we can divide force and impact in additional sub-categories: 1) direct contact and 2) contact by injection.

(4) 他用力<打>門板好幾下，然後說：我是如此的人嗎？
Ta1 yong4 li4 da3 men2 ban3 hao3 ji3 xia4, ran2 hou4 shuo1 wo3 shi4 ru3 ci3 de5 ren2 ma5?
He use force beat door plank several times, then say: I is this MOD man

“He beat forcefully the door plank several times and then said: Am I a person like this?”

(5) 當靜脈注射毒癮者<打>毒品時，通常是不會馬上把毒品立刻注射進去，而是將針頭插在血管上。
Dang1 jing4 mai4 zhu4 she4 du2 yin3 zhe3 da3 du2 pin3 shi2,
When intravenous injection drug addiction person inject drugs time,
tong1 chang2 shi4 bu2 hui4 ma3 shang4 ba3 du2 pin3 li4 ke4
usually will not at once let drugs immediately

zhu4 she4 jin4 qu4, er2 shi4 jiang1 zhen1 tou2 cha1 zai4
inject into, instead use needle head insert in

xie3 guan3 shang4
blood vessel above
“When drug addicts take drugs, they usually rest the needle in the vein instead of injecting the drug directly.”

(6) 一個滿眼夢想的快樂女孩，因為受傷開刀，像釘子進脊椎，像副衣架一般把彎彎的骨頭撐直。
Yi2 ge4  man3 yan3  meng4 xiang3  de5  kuai4 le4  nyu3 hai2,
One full eyes dream MOD happy girl,
yin1 wei4  shou4 shang1  kai1 dao1,  yao4  da3  ding1 zi5  jin4
because hurt operate, will put nail to
ji2 zhu11,  xiang4  fu4  yi1 jia4  yi1 ban1  ba3  wan1 wan1  de5
spine, like set clothes stand in general let bend MOD
gu3 tou2  cheng1  zhi2
bone prop straight

“A happy girl full of dreams due to an injury has to undergo an operation that puts nails into her spine to straighten up the wiggled spine, which is very similar to clothing being hanged on hangers.”

Among these 33 metaphorical event senses of da3, we find two main metaphorical event senses for da3: 1) hand action extension and 2) hand action unextension such as below:

(7)  我從北京坐車去太原，6 個小時的車程，就有 400 餘個電話打到我的手機上。
Wo3  cong2  bei3 jing1  zuo4 ch3  qu4  tai4 yuan2,  6  ge4  xiao3 shi2
I from Beijing by bus to Taiyuan, six hours
de5  che1  cheng2,  jiu4  you3  400  yu2  ge4  dian4  hua4  da3  dao4  wo3
MOD drive, have 400 more telephone call to I
“I took the bus from Beijing to Taiyuan. In these six hours, there were more than 400 calls made to my cell phone.”

“Philips is taking advantages from the new system developed to raise their profit.”

“We follow these categories to explore the common semantic features for these metaphorical event senses of da3 in the SUMO concept system.”
7. Concept analysis

According to Gao’s study (2001), she used a sense division principle to analyze da3, generalize the patterns and features of the polysemy of da3, and thus obtained five major categories.

Gao (2001) mentioned the prototypical meaning of da3. She talked about in the prototypical case the most central part of the meaning of da3 is the physical contact between an agent’s hands and a concrete item. In her paper, she also mentioned that there were three different semantic elements for da3 such as 1) hand, hand holdings or instrument; 2) force direction and 3) impact.

From all our senses of da3, we can divide three main categories: 1) physical activity senses such as da3 zhuo1 zi5 (to tap the table), da3 shou3 bei4 (to hit the back of a hand), ba3 wan3 da3 po4 (to break a bowl) ... and so on; 2) metaphorical event senses such as da3 jiao1 dao4 (to develop the interpersonal relationship/ to come into contact with), da3 dian4 hua4 (to call), da3 ru2 yi4 suan4 pan2 (to expect things to turn out as one wishes) ... and so forth and 3) additional senses include idiomatic senses and extended senses such as da3 zhe2 kou4 (to give a discount), da3 ke1 shui4 (to nod), and da3 jian1 (to stop for refreshments when traveling). However, in this study, we just focus on physical activity senses and metaphorical event senses. Based on Gao’s analysis (2001), we know that physical activity senses of da3 include these features such as hand, hand holdings, instrument; force and impact, so we thoroughly examine our physical activity senses by SUMO concept system. In here, we can observe that there are several concept of SUMO for physical activity senses of da3 such as below Table 5:

<table>
<thead>
<tr>
<th>The phrase of da3 in Chinese</th>
<th>Translation in English</th>
<th>SUMO concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>da3 shou3 bei4</td>
<td>to hit the back of a hand</td>
<td>impacting</td>
</tr>
<tr>
<td>da3 zhuo1 zi5</td>
<td>to tap the table</td>
<td>touching</td>
</tr>
<tr>
<td>ba3 wan3 da3 po4</td>
<td>to break a bowl</td>
<td>impacting</td>
</tr>
<tr>
<td>lun4 tai1 da3 qi4</td>
<td>to pump gas into tire</td>
<td>putting</td>
</tr>
<tr>
<td>da3 shi1 li4 kang1</td>
<td>to inject silicon</td>
<td>putting</td>
</tr>
</tbody>
</table>

Table 5: SUMO concept for physical activity senses of da3
Also based on Gao’s study (2001), metaphorical event senses of da3 include these features such as bodypart and instrument, so we could explore our metaphorical event senses by SUMO concept system again. Therefore, we would obtain that there are several kinds of concepts in SUMO ontological concept system for metaphorical event senses of da3 such as below Table 6:

 mass_table(Chinese,|Translation in English|SUMO concept|
d|da3 dian4 hua4|to call|communication|
d|da3 zhi1 ming2 du4|to build the well-known|intentional process|
d|da3 ru2 yi4 suan4 pan2|to scheme the wishful thinking|intentional process|
d|da3 zhao4 mian4|to greet|communication|
d|da3 jiao1 dao4|to develop the interpersonal relationship|intentional psychological process|

Table 6: SUMO concept for metaphorical event senses of da3

We follow SUMO concept system to obtain these concepts of physical activity senses and metaphorical senses of da3. These physical activity concepts are such as impacting, touching, putting. We also know that the SUMO concept identifications correspond with the divisions and definitions in WordNet. For this reason, we need make sure the WordNet definitions of these concepts for physical event senses of da3.

|The phrase of da3 in Chinese|The English correspondence of da3|WordNet definition|SUMO concept|
d|da3 shou3 bei4|hit|deal a blow to, either with the hand or with an instrument|impacting|
d|da3 zhuo1 zi5|tap|a light touch or stroke|touching|
|ba3 wan3 da3 po4|break|destroy the integrity of; usually by force; cause to separate into|impacting|
The Polysemy of Da3: An ontology-based study

<table>
<thead>
<tr>
<th>lun4 tai1 da3 qi4</th>
<th>pump</th>
<th>deliver forth</th>
<th>putting</th>
</tr>
</thead>
<tbody>
<tr>
<td>da3 shi1 li4 kang1</td>
<td>inject</td>
<td>force or drive (a fluid or gas)</td>
<td>putting into by piercing</td>
</tr>
</tbody>
</table>

Table 7: WordNet definition and SUMO concept for physical activity senses of da3

In this same way, we also follow SUMO concept system to obtain these metaphorical event concepts are such as communication, intentional process and intentional psychological process. We also make sure the WordNet definitions of these concepts for metaphorica event senses of da3.

<table>
<thead>
<tr>
<th>The phrase of da3 in Chinese</th>
<th>The English correspondence of da3</th>
<th>WordNet definition</th>
<th>SUMO concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>da3 dian4 hua4</td>
<td>call</td>
<td>get or try to get into communication by telephone</td>
<td>communication</td>
</tr>
<tr>
<td>da3 zhi1 ming2 du4</td>
<td>build</td>
<td>build or establish something abstract</td>
<td>intentional process</td>
</tr>
<tr>
<td>da3 ru2 yi4 suan4 pan2</td>
<td>scheme</td>
<td>form intrigues (for) in an underhand manner</td>
<td>intentional process</td>
</tr>
<tr>
<td>da3 zha04 mian4</td>
<td>greet</td>
<td>express greetings upon meeting someone</td>
<td>communication</td>
</tr>
<tr>
<td>da3 jiao1 dao4</td>
<td>develop</td>
<td>generate gradually</td>
<td>intentional psychological process</td>
</tr>
</tbody>
</table>

Table 8: WordNet definition and SUMO concept for metaphorical event senses of da3

In this way, we obtain the common semantic elements from SUMO concepts for physical activity senses of da3. The semantic features are hand, instrument, and force. In addition, we can detect when we do these actions, the manners are impact, direct contact or contact by injection. We may visual this as:
(1) Agent + hand, hand hold object or instrument

--> Patient (person or physical object)

(2) Agent --> impact (direct contact)

--> Physical object

(3) Agent with an instrument --> force (contact by injection)

--> Patient (person or physical object)
From the SUMO concept system for the physical activity senses of *da3*, we see these concepts imply the following semantic features: hand, hand holdings, instrument; force and impact (direct contact or contact by injection). Also, we would like use the same way—taking SUMO concept system in order to explore the semantic elements from their concepts of the metaphorical event senses of *da3*. From these concept of the metaphorical event senses of *da3* --- communication, intentional process and intentional psychological process, their concepts connote several semantic elements such as bodypart, instrument and mentation. For example, in “*da3 dian4 hua4* (to call)”, agent need use his hand to finish this event, therefore, this event would imply the bodypart element. Moreover, in “*da3 shui3* (to take some water)”, we know that we need use pails or some containers to obtain some water. From “*da3 shui3*”, we would expect the instrument element in this event. Finally, in “*da3 ru2 yi4 suan4 pan2* (to expect things to turn out as one wishes)”, the concept would connote the mentation element in this event. Last but not least, when we precede these events, the manners are two different contacts as hand action and visual action.

We use SUMO ontological concept system to find out all possible concepts for the physical activity senses and metaphorical event senses of *da3*, while Gao (2001) used semantic features to analyze and explain physical actions of *da3*. Following our analyses, explanations, comparison and demonstrations, we discover that our analyses and main categories correspond with Gao’s study result for the physical activity senses and metaphorical event senses of *da3*.

8. Conclusion
In this study, we explore all possible concepts for physical activity senses and metaphorical event senses of da3 through the SUMO ontological concept system. These physical activity concepts imply the semantic features: hand, hand holdings, instrument; force and impact, while these metaphorical event concepts connote the features: bodypart, instrument, presentation and contact.

We use the concept-based approach in this study, while Gao (2001) took a semantic-feature-based approach to examine the physical activity senses and metaphorical event senses of da3. We use SUMO to find out their concepts in order to group and cluster these the same concept words. For that reason, we can get different concept groups for the physical activity senses or metaphorical event senses of da3.

In this study, we attempt to use concept-based approach to analyze Polysemy in Chinese. Therefore, this leads us to propose that a concept-based approach is a viable one when exploring the sense of polysemous verbs in Chinese.

References
Hong, Jia-Fei, Xiang-Bing Li and Chu-Ren Huang. 2004. Ontology-based Prediction of Compound Relations: A study based on SUMO. Presented at PACLIC18. 8-10,


蔡美智. 2002。講「清楚」、說「明白」——漢語動詞近義、多義、詞義劃分研究。第三屆中文詞彙語意學研討會。台北，南港：中央研究院。

Website Resources

Chinese Word Sketch Engine: http://wordsketch.ling.sinica.edu.tw/

English Word Sketch Engine: http://www.sketchengine.co.uk/

