Verbal Semantics as Eventive Information: The Case of Verb “制服” and “制約” in Chinese Gigaword Corpus

Abstract. Using Chinese Gigaword Corpus, this study aims to explore the similarities and differences between two Chinese verbs “制服” (zhì fú) and “制約” (zhì yuē), so as to understand the lexicosemantics of verbs with and without an endpoint. First, the distributional patterns show that the two verbs occur mainly in the genre of story of journalism (98.4% for “制服” zhì fú and 99.1% for “制約”). Second, based on the MARVS theory and the collocations with the tense and aspect markers (“著” zhe, “了” lǐō, “過” guò), it is found that “制服” is telic (accomplishment), representing an event module with an endpoint, that is BOUNDARY (.), while “制約” is atelic and a PROCESS activity (///). Third, in contrast to the uni-directional verb “制服”, it is found that “制約” can be framed bi-directionally between the agent and the patient, using the adverbials such as “相互” (xiāng hù) and “互相” (hù xiāng). As a continuous PROCESS verb, “制約” can be modified by: a) degree adverbials such as “嚴重” (yán zhòng), b) duration /continuous adverbials such as “一直” (yī zhí), “繼續” (jì xù), and c) starting-relevant adverbials such as “開始” (kāi shǐ). The study sheds light on teaching and learning verb meanings through lexicosemantics in context, particularly the event features suggested by the MARVS theory.

Keywords: corpus linguistics, semantics, tense and aspects, MARVS, coercion

1 Introduction

Understanding how meaning is conveyed in and transmitted through language is an important task for language learners. Researchers try to decode sense and meaning from a variety of perspectives, including for example, syntactically (Chomsky, 1965), innately (Pinker, 2003), and statistically (Ellis, 2002). With the development of computers and internet, data are accumulated at an incredibly fast speed. We now can easily access to large language database (corpora), and many of which are annotated with useful part-of-speech tagging, making corpus-based analysis for understanding patterns of sense and meaning feasible. This study reports a corpus-based analysis, using a tagged corpus and a Sketch Engine called Chinese Gigaword, consisting of about 600 million words from traditional and simplified Chinese news media (Huang, 1996; Kilgarriff et al., 2005).

1.1 Verb sense and meaning

Verbs and nouns are the two basic elements in a language that help us experience and express the world. Nouns stand for entity as they often prime concrete objects and verbs
represent event and state. Verbs are eventual so that sense of verb meaning can be
determined by the event kind and formed through its conceptualization in context
(Huang et al. 2000). Such a conceptualization regarding the lexicosemantic relations of
verbs and their meaning can be understood from Fillmore’s (1982, 2008) frame theory
and Goldberg’s (1995, 2003) Caused-Motion Construction theory, concerning the
thematic relations of agents, patients, themes as well as the placement of verbs in the
“construction grammar”. In addition to the constructional grammar approaches (e.g.,
the application of FrameNet, Fillmore & Baker, 2000), the theory of Module-Attribute
Representation of Verbal Semantics (MARVS) provides a perspective for
understanding lexico-semantics by conceptualizing the event features following the
Event Modules, Event Internal Attributes, Role Module and Role-Internal Attribute
(Ahrens et al., 2003; Huang et al. 2000).

1.2 The MARVS theory

The event module proposed by the MARVS theory (Ahrens et al., 2003; Huang et al.,
2000), which represents the information on the event structure of verbs, includes five
distinguishable atomic event modules:

   BOUNDARY .
   PUNCTUALITY /
   PROCESS ///
   STATE ______
   STAGE ^^^^^^  

These event modules with their conceptual graphs indicate whether an event has an
endpoint (BOUNDARY .) and whether it is a continuous process (PROCESS ///) among others. The benefits of having the MARVS system is to identify patterns in verbs
that make them distinguishable to each other, which has important implications for
teaching, learning, and information retrieval purposes. Therefore, we base our analysis
on the MARVS theory (Ahrens et al., 2003; Huang et al., 2000).

2 Methodology

This study focuses on two Chinese verbs, “制服” and “制約” for anlaysis because they
have both meaning overlaps and differences, which make them a good exmaple for
employing corupus-based research to inform language teaching and learning. We first
identify its sense and meaning from Chinese Verbnets. Then we explore the two verbs
in Chinese gigawords corpus for analyzing their distribtional patterns, tense and aspect
collocations, and modifier patterns. Finally we discuss the results and draw a conclusion
on our research findings.
3 Results

3.1 Basic meaning of “制服” and “制約” from Chinese WordNet

From Chinese WordNet dictionary (http://lope.linguistics.ntu.edu.tw/cwn2/query/), we know that the verb “制服” is a transitive verb (VC), which means forcing someone to surrender, and “制約” is a transitive verb (VC) too, which means imposing restrictions on the current state or on the changes.

Figure 1. Basic meaning.

3.2 Analysis from Chinese Gigaword Corpus

Frequency and distribution. We then use the Chinese Word Sketch (CWS) (https://wordsketch.ling.sinica.edu.tw/cws/) (Huang, 1996; Kilgarriff et al., 2005) for analyzing the frequency and distributional patterns of “制服” and “制約” in Chinese Gigaword Corpus, mainly composed of articles from Chinese journalism. First, it is found there are in total 5,496 instances of “制服”, 496 of which are used as verbs, and 13,093 instances of “制約”, all of which are used as verbs. Second, the verbs “制服” and “制約” mainly occur in the genre of story in journalism (98.4% for “制服” and 99.1% for “制約”). The results are presented in Table 1.

Table 1. Frequency and distribution in the Chinese Gigaword Corpus.

<table>
<thead>
<tr>
<th>Genre Verb</th>
<th>Story</th>
<th>Multi</th>
<th>Other</th>
<th>Advis</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>制服</td>
<td>486</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>494</td>
</tr>
<tr>
<td>%</td>
<td>98.4%</td>
<td>0.8%</td>
<td>0.8%</td>
<td>0.0%</td>
<td>100%</td>
</tr>
<tr>
<td>制約</td>
<td>12,972</td>
<td>76</td>
<td>44</td>
<td>1</td>
<td>13,093</td>
</tr>
<tr>
<td>%</td>
<td>99.1%</td>
<td>0.6%</td>
<td>0.3%</td>
<td>0.0%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Story: of a particular topic or event;
Multi: describing a particular topic or event as in “news briefs in finance or sports”;
Advis: what the news service addresses to news editors; not intended readers;
Other: none of the above; items such as sports scores, stock prices, temperatures report.
Collocations with tense and aspect markers. Using the CWS, we conducted searches on the collocations of the keywords “制服” and “制約” respectively with the five tense and aspect markers: “著” (zhe), “了” (liǎo), “過” (guò) (Keyword 制服 /制約 + 著 /了/過), “在” (zài), and “正在” (zhèng zài) (在/正在 + Keyword 制服 /制約). The purpose is to observe whether the two keywords are telic (e.g., accomplishment and achievement verbs like “build” and “retire”) or atelic (e.g., state and activity verbs like “know”, and “walk”) based on the internal temporal constituency of a situation (Comrie, 1976; Smith, 1991, 1997), so that we can use the MARVAS theory to judge the internal event modules of the two verbs— whether they are verbs with boundary, punctuality, process, state, or stage (Ahrens et al., 2003; Huang et al., 2000). The results are presented in Table 2.

Table 2. Collocations with tense and aspect markers in the Chinese Gigaword Corpus (incidences)

<table>
<thead>
<tr>
<th></th>
<th>+ 著 zhe</th>
<th>+ 了 liǎo</th>
<th>+ 過 guò</th>
<th>在 zài</th>
<th>正在 zhèng zài</th>
</tr>
</thead>
<tbody>
<tr>
<td>制服</td>
<td>0</td>
<td>259</td>
<td>2</td>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td>制約</td>
<td>786</td>
<td>1,562</td>
<td>1</td>
<td>24</td>
<td>7</td>
</tr>
</tbody>
</table>

“正在 zhèng zài +” is also included in “在 zài +”
**Collocations with modifiers.** Using the CWS, we conducted search on the collocations of the keywords “制服” and “制約” with modifiers (modifier + Keyword 制服 /制約) “相互” (xiāng hù)、“互相” (hù xiāng)、“彼此” (bǐ cǐ)、“共同” (gōng tóng)、“严重” (yán zhòng)、“一直” (yī zhí)、“继续” (jì xù)、and “開始” (kāi shǐ). The results are presented in Table 3. The results indicate that among the many modifiers, “制服” can only collocate with “共同” (“共同制服”). None of the other modifiers that work well with “制約” collocate with “制服”. These modifiers work with “制約” to describe the bidirectional agent and object (“相互制約”、“互相制約”), the degree of severeness (“嚴重制約”), the duration relevance (“一直制約”、“继续制約”、“開始制約”).

It is because first of all, “制服” is uni-directional, so that the bidirectional agent and object (“相互”、“互相”) do not collocate with it. Second, “制服” denotes an endpoint according to the MARVS theory (Ahrens et al., 2003; Huang et al., 2000), there is no need to put a degree and duration to “制服”, which is the same as the accomplishment verbs such as “cross the road”, “build a house”. It is noted that due to the boundary that instantly make an endpoint (like semelfactive verbs such as “jump”, “sneeze”, “hit”), neither is it necessary for “制服” to collocate with starting-relevant adverbials such as “開始” (kāi shǐ). In contrast, “制約” with a continuous active course, can be modified by the duration or continuous adverbials, as well as degree adverbials and starting-relevant adverbials.

<table>
<thead>
<tr>
<th></th>
<th>相互+</th>
<th>互相+</th>
<th>彼此+</th>
<th>共同+</th>
<th>严重+</th>
<th>一直+</th>
<th>開始+</th>
<th>制服</th>
<th>制約</th>
</tr>
</thead>
<tbody>
<tr>
<td>xiāng hù</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>320</td>
</tr>
<tr>
<td>hù xiāng</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>121</td>
</tr>
<tr>
<td>bǐ cǐ</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>gōng tóng</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>yán zhòng</td>
<td>1,174</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>yī zhí</td>
<td>40</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>1,174</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>jì xù</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1,174</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>kāi shǐ</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1,174</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

We observed from the modifier collocations in the Word Sketch Differences analysis that that the two transitive verbs “制服” and “制約” can have similar agent and patient regarding thematic relations (Fillmore, 1982), but they can also be modified by different adverbials, because the semantic event modules of “制服” has an endpoint while “制約” is a continuous process oriented verb. The results of Word Sketch Differences analysis are presented in Figure 2.
Figure 2. Word Sketch Differences analysis

Additional analysis of “制服” and “制约” in the CWS, we also observed the phenomenon of “coercion”, extending the basic meaning of “制服” from “making someone to surrender” to a metaphorical meaning of “controlling disease” and “controlling inflation by the government”. In that sense, the object of “制服” is not necessarily a human (“someone”) but can be problems such as “diseases” and “inflations”. We summarize the results of in Figure 3.
4 Discussion and Conclusion

In this section, the results are summarized regarding the verb similarities and differences. There are probably three main similarities between the verbs “制服” and “制約”:

1. both are transitive activity verbs that may have an object;
2. both have concrete animated agent and patient;
3. both can be modified by adverbials such as “finally”.

There are also significant differences between the two verbs. First, according to the MARVS theory, “制服” represents an event module with an endpoint, that is BOUNDARY (.), while “制約”, a continuous PROCESS (\). Second, “制服” is uni-directional, and has telic aspect (like an accomplishment), so it cannot be modified by:

a) bi-directional adverbial, b) degree adverbial, c) duration adverbial, and d) starting-relevant adverbial.

5 Implications

From the analysis of lexicosemantic of the verbs “制服” and “制約” in this article, it is argued that the MARVS theory (Huang et al., 2000; Chung & Ahrens, 2008) can be used to decode and understand verb meanings. The event features described in the Module-Attribute representations can be employed to conceptualize verb senses and meanings. Therefore, it is suggested that teachers and students can use the MARVS theory and the large language corpora such as the 600 million Chinese Gigaword corpus for gaining contextualized meaning understanding of individual words, distributional patterns, collocations, as well as meaning differences among words.
References