

A MARVS Analysis of Two Chinese Near-Synonymous Verbs of Jumping Based on Chinese Corpora

Yingying Peng

Chinese and bilingual studies

Faculty of Humanities

The Hong Kong Polytechnic University

20113471r@polyu.edu.hk

Abstract

The Module-Attribute Representation of Verbal Semantics (MARVS) has been served as an excellent tool for the analysis and differentiation of Chinese near synonymous verbs. In the current study, two Chinese near-synonymous verbs 蹦 (bèng; to jump) and 跳 (tiào; to jump) which have not been investigated in previous studies are analyzed within the framework of MARVS with their occurrence data and observed distributional differences obtained from two large Chinese corpora, GigaWord2 and the CCL corpus. The results show that compared to bèng, tiào is much more flexible in terms of directionality and locations, which could ultimately be accounted for with the fact that tiào emphasizes more on the displacement of the subject caused by the action.

1 Introduction

Within a semantic domain, words that can be used in a similar fashion to describe similar events are considered as members of the same class of words and further named as near-synonyms (Gao et al., 2016). Near-synonyms are notoriously difficult to be differentiated from each other with explicit explanations even for native speakers. It thus presents great challenges for second languages learners and also for the designation of cross-linguistic dictionaries and domains such as machine translation (Hong, 2014).

In view of the significance of distinguishing near synonyms especially near synonymous verbs and insufficiency or even inadequacy of previous theoretical frameworks (e.g., (Levin, 1993)) designated for

the analysis of English to be extended for the differentiation of Chinese near synonyms, Huang and colleagues (Huang et al., 2000) proposed the representational model called the Module-Attribute Representation of Verbal Semantics (MARVS). The model has been proven to be more adequate in analyzing a broad span of Chinese near-synonymous verbs including pi2juan4 (tired) and lei4 (tired), ge1 (to slice) and qie3 (to cut) (Tsai et al., 1998), and ben1 (to run) and pao3 (to run) (Wang, 2009). Nevertheless, beng4 (to jump) and tiao4 (to jump) as two Chinese near-synonymous verbs denoting the meaning of jumping with frequent usage have never been thoroughly studied.

Thus, the current study aims to adopt the MARVS framework to analyze bèng (to jump) and tiào (to jump) and distinguish them based on their distributional differences obtained from two large Chinese Corpora, namely Gigaword2 (Hong and Huang, 2006) and CCL (Zhan et al., 2019).

2 Literature Review

This session firstly introduces the basics of MARVS. Then insights gained from previous studies on near-synonymous verbs that adopted this model are elaborated in order to lay out a clearer picture of how the current study would be conducted. Then, there will be a brief introduction of these two Chinese verbs followed by the research questions the current study attempts to answer.

2.1 The Module-Attribute Representation of Verbal Semantics (MARVS)

In order to better capture the semantic differences of these two Chinese verbs, Module-Attribute Representation of Verbal Semantics (MARVS) (Huang et al., 2000), a well-tested model built upon Chinese data is applied in the current study. This model postulates that the most essential semantic features of verbs are involved in the configuration of the following four aspects: Event Modules, Event-internal Attributes, Role Module and Role-internal Attribute. In this model, the two modules are the pre-bundled semantic information of the verb while the two attached attributes can give a more fine-grained description of the semantics of verb.

The event module represents the main information on the event structure of the verbs. They are the building blocks of verbal linguistic event structures. Under this proposal, five atomic event modules have been introduced in MARVS as follows : boundary, punctuality, process, state, stage. They are basically concerned with whether the verb encodes a complete or progressive action or homogeneous state or other more complex events.

The event attributes are concerned with the semantics of the whole event and thus are attached to the event module. They can include [control], [effect] (Huang et al., 2000). This can be used to discover more nuanced differences among verbs because verbs with same event modules might vary in their internal attributes.

The role modules represent the participants of the event (Huang et al., 2000) huang2000module, such as agent, theme, causer, manner, location and so on.

Lastly, the role-internal attributes are the semantic attributes pertaining to a particularly focused role in the event, such as sentience, volition, affectedness, etc (Huang et al., 2000).

2.2 Insights from Previous Studies that Adopted MARVS

Under the framework of MARVS, a considerable number of studies have carried out their investigations on Chinese near-synonymous verbs (Ahrens and Huang, 2001; Liu et al., 2000; Chang et al., 2000a). These studies generally exploited the following methodologies: 1) Extracting occurrences of

near-synonymous verbs in large corpora, summarizing distributional differences between these verbs and making inferences regarding their contrasting grammatical relations; 2) figuring out the semantic implicatures that could explain such distributional differences; 3) verifying the findings through applying them to new syntactic/semantic structures. Their step-by-step analyses and findings provide valuable insights to the current study.

When examining the structures of two Chinese near-synonymous verbs of running *ben1* and *pao2* (Wang, 2009), the author summarized the distributional similarities and differences of these two verbs collocated with the following marking devices: *zai4*, *zhe2*, *kai1shi3* (to begin), *qi3lai2* (denoting the beginning of an action) and *wan2* (finish). Here *zai4* is a progressive marker and *zhe2* is a durational marker. Both of them inform the progressive nature of verbs collocated with them. *Kai1shi3* (to begin) and *qi2lai3* are event starting-point markers and *wan2* (finished) denotes the end of the event. Using these markers, the author successfully identified distributional differences between *ben1* and *pao2* and concluded that *ben1* is an inchoative process because it never collocates with the finish marker while *pao2* is a bounded process because it can collocate with all the above mentioned marking devices. Previous studies (Chang et al., 2000b; Huang et al., 2000) also included other durational phrases such as *yi4zhi2* (continuously) and *le2 + duration*.

For event attributes, previous studies have found that even though both *kuai4 le4* (happy) and *gao1 xing4* (happy) are state verbs (Tsai et al., 1998), they actually differ in terms of their event attributes. To put specific, they differ in the [+/-control] attribute, because *gao1 xing4* can take evaluational sentences as its objects that show the calculated reaction (e.g., *zhen1 de2 gao1 xing4*; really happy) while it does not take wish sentences that embrace impredictive nature such as " *Zhu4 ni2 gao1 xing4** (Wish you happy)" as its sentential objects.

For the role modules, two near-synonyms of doubting *huai2 yi2* and *cai1* (Liu, 2002) were found to play different semantic roles. The former should be represented as \downarrow Agent, Theme \downarrow while the latter is represented as \downarrow Experiencer, Theme \downarrow based on their role module information.

Finally when it comes to the examination of role-

internal attributes, this pair of near-synonyms bai3 (to set) and fang4 (to put) with the sense of putting were carefully studied based on their distributional differences and concluded as possessing distinctive event structure focus. Bai3 can take objects that denotes results while fang4 cannot. Moreover, bai3 is able to be modified with orientational adjuncts while fang4 is not. For example, "ta1 zheng4 zai4 bai2 shu1 (He/she is putting down the books)" is acceptable while "ta1 zheng4zai4 fang4* shu1" would sound problematic for native speakers. Another example could be "ma1ma bai3 chu1 yi4 zhuo1 cai4 (mother cooked and set a table of dishes)" as mentioned in the paper. Fang4 is not acceptable in this sentence largely because it does not denote the result of the action.

Another pair of near-synonymous verbs that show differences in terms of the role-internal attribute of manner is qie1 (to cut) and ge1 (to slice) (Huang et al., 2000). Namely, ge1 involves a cautious and traceable action that has inherent time-duration while qie1 entails a movement with no specification of the manner. Thus, there are a lot more descriptive VR compounds collocated with ge1 instead of qie1.

All these studies provide valuable perspectives for the investigation of beng4 and tiao4. For instance, when probing into the event structure of these two verbs, it will be helpful to look at whether these two verbs could be collocated with the marking devices used in the above studies and the rich findings from the role-internal attributes of verbs prompt the current study to pay more attention to the semantic attributes of roles participating in the action of jumping.

2.3 An introduction of two near-synonymous verbs: beng4 and tiao4

In Chinese, both beng4 and tiao4 have the meaning of 'stepping on the ground and making use of the reaction force to move the body off the ground.', which is demonstrated clearly in the following figure extracted from Chinese WordNet (<http://lope.linguistics.ntu.edu.tw/cwn/>), a dictionary-like resource presenting word meanings and word/sense relations that have all been attested based on corpus data. (Huang et al., 2010). Here we can see that both of these two verbs share the exact same sense of jumping and they are also mutually

considered as synonyms. In this study, we only consider the sense of jumping of these two verbs as explained above because the two verbs also have other meanings which are beyond the scope of the current study.

These two verbs are not only highly similar in terms of this particular sense but also frequently used in Chinese as both verbs are in the list of 2,500 most frequently used Chinese characters (Fu, 1988). Thus, as far as we consider, it is important to make explicit of any possible nuanced differences between these two verbs to advance our understanding of near-synonymous verbs and extend the findings to second language learning of Chinese for benefiting more second language learners.

2.4 Research Questions

Since no previous studies have ever enquired the differences of these two near-synonymous verbs, the current study plans to carry out the first investigation of beng4 and tiao4 and aims to answer the following questions:

- 1) Can the two near-synonymous verbs be used interchangeably in all contexts?
- 2) If not, can any of their distributional differences be identified based on the four major components of the MARVS model?
- 3) If there are any differences in terms of any major components of MARVS, what are the underlined semantic implicatures?
- 4) Can the semantic implicatures be verified through new grammatical/semantic structures observed in the corpora?

3 Method

3.1 Two Corpora

The current study makes use of two corpora to obtain the instances of beng4 and tiao4 and their contexts. The first one is called the CCL corpus (Zhan et al., 2019), which was developed by the center for Chinese Linguistics of Beijing University. This corpus contains 783,463,175 Chinese characters and its data was extracted from both modern and classical Chinese. The other large corpus adopted in the study is Chinese Gigaword2 (Huang, 2009) with more than 800 million PoS tagged words from three varieties of Mandarin Chinese: Taiwan, Mainland

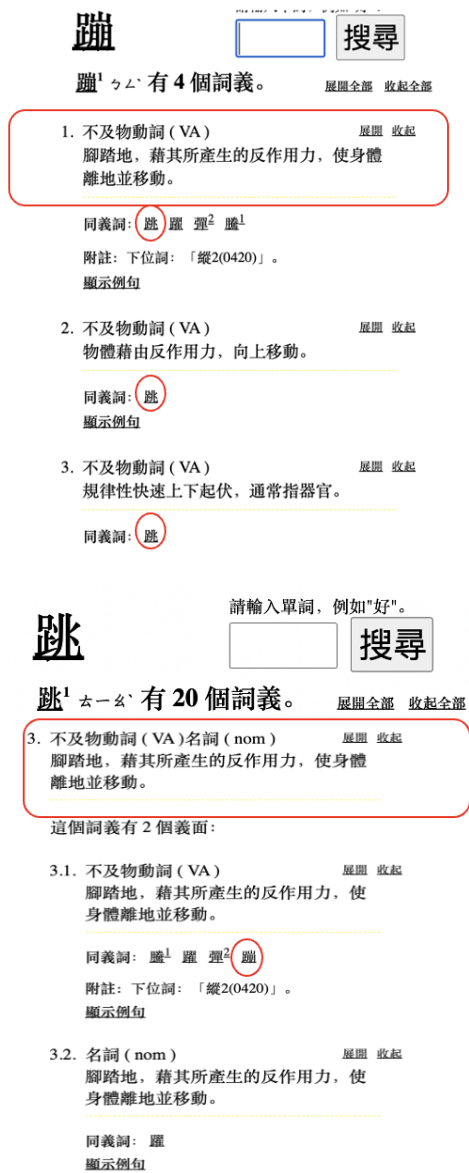


Figure 1: Captions of the searching results of beng4 and tiao4 from Chinese WordNet

China, and Singapore. It is implemented on Chinese Word Sketch (Huang et al., 2005). These two corpora are selected for the current study because they contain large amount of data and are well-accepted.

3.2 Procedures

The current study first extracts all occurrences of two jumping verbs and see if these two synonyms share any common texts and then examines whether there is any distributional difference in terms of the four major components under the framework of MARVS (i.e., event modules, event-internal attributes, role modules and role-internal attributes) taking reference of previous explorations. Next, the distribution differences are used to infer their semantic implicatures. Lastly, new scenarios where these two verbs occur are introduced in order to verify the semantic implicatures.

4 Results and Discussion

In the two corpora, there are respectively 268 (Gigaword2) and 2914 (CCL) instances of beng4 and 8243 (Gigaword2) and 53204 (CCL) of instances of tiao4. It could be seen that tiao4 is more frequently used than beng4. However, here as noted above, we only take into consideration of these two verbs with the meaning of jumping, which eliminates all occurrences of tiao4 with its other frequent collocates such as tiao4wu2 (to dance).

4.1 Common Contexts of Beng4 and Tiao4

To start with, these two verbs are frequently used in the same contexts due to their highly similar nature when referring to jumping. Table.1 indicates the number of co-occurrences of these two verbs in the same contexts extracted from the two corpora. Examples of their frequently co-occurring patterns can be found in Figure 2. These examples show that when used together, these two verbs tend to indicate that the actions of Beng4 and Tiao4 happen simultaneously, which imply the highly overlapping or even same semantic meanings of these two verbs. Otherwise these two actions cannot take place at the same time since they require the same body movements as denoted in Chinese WordNet. As we can see the total number of beng4 in these two corpora are only 268 and 2914 and the large number of co-occurrences of these two verbs in the same context

shows the prevalence of these two verbs used with the same meaning.

Corpus	co-occurrences
Gigaword2	106
CCL	1247

Table 1: Co-occurrences of Beng4 and Tiao4 from Two Corpora.

Frequently- observed Co-occurrence
又蹦又跳
一蹦一跳
蹦著跳著
愛蹦愛跳
蹦啊跳啊
活蹦亂跳
蹦蹦跳跳

Figure 2: Example of Co-occurrences of Beng4 and Tiao4 from CWS.

4.2 Findings Concerning the Event Module of Beng4 and Tiao4

Since the previous studies have already provide insights regarding possible aspect markers that can give hints to differentiating the event modules of near-synonymous verbs, the current study would firstly incorporate these marking devices including: qi2lai2 to mark the beginning of the event, ASP-zhe2, yi4zhi2 to show that the event can be progressive and finally le2ji2beng4/le2ji2tiao4 to demonstrate whether these jumping events could be clearly defined based on their starting and ending time points. The distributional information of these two verbs with the above mentioned marking devices are summarized and presented in Table 2 and Table 3.

Table 2 summarized the distributional information of beng4 and tiao4 with qi2lai2 and le2ji2beng4/tiao4. The former denotes the beginning of the verbs. The latter shows that this action could be counted as to how many times the jumping events have been acted. Thus, it implies the fact that

the action has a clear starting and ending time point. Since both of these two verbs could be collocated with qi2lai2 and le2ji2beng4/tiao4, it can concluded that both beng4 and tiao4 are bounded at the beginning and end of the action.

In addition, these two verbs can also occur with marking devices zhe2 and yizhi2, which indicate that the action is a process. For example, in the following sentences:

1) Xiao3Ping2 tiao4 zhe2 chu1 lai2

Xiaopeng came out by jumping

2) Ta1 Beng4 zhe2 wang2 shan1shang4 pao2

He was jumping while running to the mountain.

With these examples, it could be recognized that beng4 and tiao4 can be referred to a process in the sense that they can keep repeating when the main action is to go somewhere by either walking or running. However, even in this case, these two verbs should also be identified as a bounded process because the jumping events can be started and ended along with the main actions of walking and running. Hence, beng4 and tiao4 can also be seen as bounded processes.

Taken together, no distributional differences have been identified for beng4 and tiao4 in terms of event modules.

Word	qi2lai2	le2ji2tiao4/beng4
Beng4	55	2
Tiao4	920	5

Table 2: Distributional Information of Beng4 and Tiao4 with Qi2lai2 and Le2jibeng4/tiao4.

Word	Zhe2	Yi4zhi2
Beng4	82	2
Tiao4	787	7

Table 3: Distributional Information of Beng4 and Tiao4 with zhe2 and yi4zhi2

4.3 Findings concerning the Role Module and Role-Internal Attributes of Beng4 and Tiao4

Next, we would continue our investigation of beng4 and tiao4 by examining whether they differ with regard to role modules and role-internal attributes.

Since the role module is concerned with the focused roles of an event which contain all required arguments and optional arguments and adjuncts (Huang et al., 2000), we start with summarizing all possible grammatical constructions formed by these two verbs.

V: 1) 從頭到尾在座位上又蹦又跳
Preposition + Location + V: 1) 不停地從教練席蹦上蹦下 2) 从车上跳下来
V + Preposition +/- location: 1) 單腿蹦上起跳臺，然後衝著教練 2) 剛要蹦下來 3) 賈小谷就唰的一下子跳上去 4) 它跳上窗臺
Preposition + direction + V: 1) 從炕裏向上蹦出去 2) 在火車上向上跳
V + asp(着) + V: 1) 小和小平跳著出來 2) 他跳著，躍著，蹦著，拼命往上跑
V + de(得) + adv: 1) 就像皮球，越摔打，越是蹦得高 2) 復出的喬丹甚至沒有 45 號跳得高，主要靠經驗打球
Preposition + location + V + preposition + location: 1) 從窗臺上跳到四樓 2) 從床上蹦到地上

Figure 3: Possible grammatical constructions of Beng4 and Tiao4.

From Figure 3, we could see all types of sentence structures that beng4 and tiao4 form with data obtained from the two large corpora. It can be summarized that both beng5 and tiao4 allow at least these seven distinctive grammatical constructions as listed in the figure and there is no observable differences between these two verbs in this regard.

4.4 Findings concerning the Role Internal Attributes of Beng4 and Tiao4: directionality

Since no apparent differences have been identified between these two near synonymous-verbs regarding their event and role modules, the current study takes a step further to examine whether they are distinctive in terms of role-internal attributes. The first finding is that these two near-synonymous verbs differ in terms of the flexibility in direction. Based on the two corpora, we extracted all grammatical constructions of 'Preposition + direction + V' out of all the occurrences of these two verbs and divided them into six directions including shang4 (up), xia4 (down), qian2 (forward), hou4 (backward), zuo3 (left) and you4 (right) with two prepositions xiang4 and wang2 meaning 'towards'. The Tables below show the number of occurrences of these two verbs with each directional word. From the table, the distributional differences of these two words with different direction words can be observed. Namely, tiao4 can accommodate all directions when referring to jumping while beng4 can only take forward, upward and downward.

Word	Forward	Backward
Beng4	5	0
Tiao4	24	24

Table 4: Distributional Information of Beng4 and Tiao4 with Qian2(forward) and Hou4(backward)

Word	Up	Down
Beng4	8	3
Tiao4	47	129

Table 5: Distributional Information of Beng4 and Tiao4 with Shang4(up) and Xia4(down)

Word	Left	Right
Beng4	0	0
Tiao4	3	1

Table 6: Distributional Information of Beng4 and Tiao4 with Zuo3(left) and You4(right)

In order to better confirm the findings, instances

of tiao4 and beng4 collocating with de2 (marker prior to the resultant complements) gao1 (high) and yuan2 (far) are also extracted because these two adjuncts can also imply the directions of the jumping events, namely upward in the case of gao1 (high) and forward in the case of yuan2 (far). The results are displayed in Table 7. As can be seen from the table, high and far can well be the results of tiao while for beng4, it seems that gao1(high) is the better collocation due to its much higher occurrences than yuan2 (far). Also, two out of four sentences where beng4 is collocated with yuan2 have grasshopper as the subject of beng4 whose main movement is achieved via hopping. Thus, it leads to the speculation that yuan2(far) is just a by-product of beng4, because when one jumps, it is very possible or sometimes inevitable that the body also moves forward. Therefore, it could be summarized that beng4 focuses mostly on the upward movement or vertical level movement as indicated by the considerable amount of occurrences of beng4 de2 gao1 (jumping high),

Word	high	far
Beng4	24	4
Tiao4	108	23

Table 7: Distributional Information of Beng4 and Tiao4 with High (de2gao2) and Far(de2yuan2)

Therefore, as a interim summary, tiao4 enjoys much more flexibility in the direction of the movement while beng4 is more confined and only collocates with vertical (up and down)and forward (much less common) directions.

4.5 Findings Concerning the Role-internal Attributes of Beng4 and Tiao4: Degree of Diversity in Locations

Except for the distributional differences of beng4 and tiao4 regarding directionality, there is another difference observed between these two near-synonymous verbs when they are used in the grammatical construction of 'preposition + location + V + xia4 (from + location + V + down; jumping down from somewhere)', The main differences are recognized in the degrees of diversities embedded in these location words. The following figures respec-

tively show the three location words appearing prior to beng4 xia4 (down) and the 28 different locations occurring before tiao4 xia4 (down) to indicate from where the subject jumps down.

It can be seen from the two figures that the locations from which subjects tiao4xia4 and beng4xia4 are different in the sense that the locations that collocate with tiao4 involve all locations that collocate with beng4 (i.e., bed, car and desk) and these locations are fairly confined in terms of quantity and distance, because they all denote shorter distance. Nevertheless, tiao4 can also be used with locations that are well beyond the height of bed, car and desk, and much more importantly, are considerably more diverse than beng4, such as building, rooftop, third floor to eighth floor, bridge, tree, tower, etc. This seems to suggest that compared to beng4, the jumping event tiao4 can take place in more different locations with various starting points in terms of space.

4.6 The Real Cause of These Differences in Directionality and Location: +/- Emphasis on the Resultant Displacement

With the findings of the different distributional patterns of beng4 and tiao4 with regard to different direction words and diversity of locations, now we can contemplate what is the real cause of these distributional differences. Since there are some noticeable similarities of beng4 and tiao4 with bai3 (to set) and fang4 (to put) (Huang et al., 2000), we can refer to the findings of bai3 and fang4 to infer how beng4 and tiao4 can be differentiated. Bai3 and fang4 are two near-synonymous verbs that differ in terms of their role internal attribute of [design], which points to the fact that bai3 requires the resulted state of its objects to be attained. This semantic requirement thus initiates for instance the grammatical construction of "orientation + bai3 (e.g., chao2 dong1 bai3; set something towards the east)".

Based on the previous findings and the current data, it is proposed that the differences of beng4 and tiao4 are essentially arisen from the fact that tiao4 emphasizes the displacement of the subject caused by the jumping event. That should be the reason why tiao4 occurs with more types of directions and much higher frequencies with directional words, since these directions indicate the displacement of the jumping event in terms of directional-

Beng4 xia4
床上(chuang2shang2; bed)
車上(che2shang4; car)
炕上(kang4shang4; bed)
桌上(zhuo1shang4; desk)

Tiao4 Xia4:
床(chuang4shang4; bed)
車上(che1shang4; car)
炕上(kang4shang4; bed)
桌上(zhuo1shang4; desk)
臺上(tai2shang4; stage)
八樓(ballou2; 8th floor)
窗口(chuang1kou2; window)
二樓(er4lou2; 2nd floor)
樓頂(lou2ding2; roof)
樓上(lou2shang4; building)
鋪上(pu4shang4; bed)
岸邊(an4bian1; bank)
空中(kong1zhong1; in the air)
牆頭(qiang2tou2; wall)
三樓(san1lou2; 3rd floor)
樹上(shu4xia4; tree)
四樓(si4lou2; 4th floor)
五樓(wu2lou2; 5th floor)
高空(gao1kong1; sky)
高處(gao1chu4; high place)
碼頭(ma2tou2; pier)
屋頂(wu1ding3; roof)
橋上(qiao2shang4; bridge)
塔頂(ta3ding3; tower)
巖頂(yan2ding3; rock)
房上(fang2shang4; building)
上鋪(shang4pu4; upper bunk)
高樓(gao1lou2; tall building)

ity. The displacement proposal can also account for the diverse and abundant starting locations of tiao4 vs. beng4, since displacement should naturally take into consideration of the places prior to and after the jumping events. The final proof of this proposal comes from the relatively abundant instances of the construction of "cong2 (from) + (one place) + tiao4 + dao4 (to) + (another place)", which has 36 occurrences while the same construction for beng4 only occurs twice. The final set of corpus data makes it clear that tiao4 concerns with the starting and ending point of the jumping event because its semantic implicatures demand the emphasis on the displacement as the result of this event.

On the other hand, beng4 seems to be more concerned with the jumping action itself since it mostly occurs with upward and downward directions and much lessly forward, which is a natural consequence of the event beng4 without specific need to focus on other directions to indicate the results of the action. Apart from this, beng4 often appears as the predicate of the sentence without any other grammatical roles such as prepositions and locations, which can be seen from the fact that the co-occurrences of beng4 and tiao4 as the only element in the VP can account for a large proportion of the instances of beng4 (i.e., gigaword2: 39.5 percent; the CCL corpus: 42.7 percent). This can also serve as a piece of evidence that beng4 is more concerned with the jumping action itself instead of the displacement as a result of the action.

4.7 Verification

The postulation that tiao4 emphasizes the displacement while beng4 only emphasizes the action itself can also be borne out by a new construction that shows a distributional difference for beng4 and tiao4. The new construction is beng4/tiao4 guo4qu4 (to jump over there). Since the phrase clearly demonstrates the fact of being displaced after the jumping event, it is predicted that tiao4 will be used with guo4qu4 (over there) much more frequently than beng4. This hypothesis is supported by the corpus data. For tiao4, there are in total 122 instances of such construction while for beng4, there are only 6 occurrences.

Also, in Chinese, "jumping off a building" is expressed as 'tiao4lou2' rather than 'beng4lou2'.

Figure 4: Locations from which subjects jump down via Beng4xia and Tiao4Xia4.

This also serves as another piece of evidence that tiao4 refers to longer distances as well as displacement of the subject as a consequence of the action jump, namely from certain floor to the ground in this phrase.

Another proof can be the fact that only tiao4gao1 (jump high) and tiao4yuan2 (jump far) are used as the term for high jump and long jump in Chinese rather than beng4gao1 and beng4yuan2, because in these sports, the distance of the displacement is considered as the most crucial part.

5 Conclusion

Beng4 and tiao4 are two near-synonymous verbs that share a very high degree of similarity and occur frequently in Chinese. This might construct difficulties for second language learning or other domains where cross-linguistic equivalents are needed.

Under the MARVS framework and with previous analysis and findings, the current study extracts occurrences of these two verbs from two large corpora, identifies and analyzes their distributional differences in terms of directionality and locations and finally comes to the conclusion that these differences emerge from the fact that tiao4 entails a semantic focus on the displacement of the action while beng4 largely focuses on the action itself. This proposal can also be used to explain new grammatical constructions and terms where these two verbs are concerned.

This study can serve as another exploration of Chinese near synonyms and it reconfirms the robustness of MARVS model in differentiating Chinese near synonyms. Future studies can continue the exploration with more corpus data and carry out diachronic or cross-linguistic research.

References

- Ahrens, K. and Huang, C.-R. (2001). A comparative study of english and chinese synonym pairs: An approach based on the module-attribute representation of verbal semantics. In *Proceedings of the 15th Pacific Asia Conference on Language, Information and Computation*, pages 27–32.
- Chang, L.-l., Chen, K.-j., and Huang, C.-R. (2000a). Alternation across semantic fields: A study on mandarin verbs of emotion. In *International Journal of Computational Linguistics & Chinese Language Processing, Volume 5, Number 1, February 2000: Special Issue on Chinese Verbal Semantics*, pages 61–80.
- Chang, L.-l., Chen, K.-J., and Huang, C.-R. (2000b). A lexical-semantic analysis of mandarin chinese verbs: Representation and methodology. *Computational Linguistics and Chinese Language Processing*, 5(1):1–18.
- Fu, Y. (1988). Frequentwordsofmodernchinese. *LanguageConstruction*, 2.
- Gao, H. H., Wang, H., and Nicoladis, E. (2016). The delineation of ‘throw’verbs in mandarin chinese: Behavioural and perceptual approaches. *Journal of Cognitive Science*, 17(1):95–131.
- Hong, J.-F. (2014). Chinese near-synonym study based on the chinese gigaword corpus and the chinese learner corpus. In *Workshop on Chinese Lexical Semantics*, pages 329–340. Springer.
- Hong, J.-F. and Huang, C.-R. (2006). Using chinese gigaword corpus and chinese word sketch in linguistic research. In *Proceedings of the 20th Pacific Asia conference on language, information and computation*, pages 183–190.
- Huang, C.-R. (2009). Tagged chinese gigaword version 2.0, ldc2009t14. *Linguistic Data Consortium*.
- Huang, C.-R., Ahrens, K., Chang, L.-L., Chen, K.-J., Liu, M.-C., and Tsai, M.-C. (2000). The module-attribute representation of verbal semantics: From semantic to argument structure. In *International Journal of Computational Linguistics & Chinese Language Processing, Volume 5, Number 1, February 2000: Special Issue on Chinese Verbal Semantics*, pages 19–46.
- Huang, C.-R., Hsieh, S.-K., Hong, J.-F., Chen, Y.-Z., Su, I.-L., Chen, Y.-X., and Huang, S.-W. (2010). Chinese wordnet: Design, implementation, and application of an infrastructure for cross-lingual knowledge processing. *Journal of Chinese Information Processing*, 24(2):14–23.
- Huang, C.-R., Kilgarriff, A., Wu, Y., Chiu, C.-M., Smith, S., Rychly, P., Bai, M.-H., and Chen, K.-J. (2005). Chinese sketch engine and the extraction of grammatical collocations. In *Proceedings of the fourth SIGHAN workshop on Chinese language processing*.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago press.
- Liu, M.-c. (2002). Corpus-based lexical semantic study of verbs of doubt: Huaiyi and cai in mandarin. *Concentric: Studies in Linguistics*, 28(2):43–55.
- Liu, M.-c., Huang, C.-R., Lee, C. C., and Lee, C.-Y. (2000). When endpoint meets endpoint: A corpus-based lexical semantic study of mandarin verbs of throwing. In *International Journal of Computational Linguistics & Chinese Language Processing, Volume 5, Number 1, February 2000: Special Issue on Chinese Verbal Semantics*, pages 81–96.

- Tsai, M.-C., Huang, C.-R., Chen, K.-J., and Ahrens, K. (1998). Towards a representation of verbal semantics—an approach based on near-synonyms. In *International Journal of Computational Linguistics & Chinese Language Processing, Volume 3, Number 1, February 1998: Special Issue on the 10th Research on Computational Linguistics International Conference*, pages 61–74.
- Wang, J. (2009). A corpus-based study on the chinese near-synonymous verbs of running. *Volume 2/edited by Yun Xiao.*, 2:399.
- Zhan, W., Guo, R., Chang, B., Chen, Y., and Chen, L. (2019). The building of the ccl corpus: its design and implementation. *Corp. Ling*, 6:77–86.