

The Mapping between the Literal and the Metaphorical: A Latent Semantic Analysis Approach

Qiuhan Zhang¹, Tangfeng Yang² & Ran Tao³

¹ School of Chinese Studies and Exchange, Shanghai International Studies University, Shanghai, China

² Foreign Languages College, Donghua University, Shanghai, China

³ School of Computer Science, Donghua University, Shanghai, China

Correspondence: Tangfeng Yang, College of Foreign Languages, 2999 North Renmin Road, Songjiang District, Shanghai 201620, China. E-mail: yangtangfeng@dhu.edu.cn

Received: November 1, 2025 Accepted: November 15, 2025 Online Published: November 17, 2025

Abstract

With Chinese “飞翔” (*fly/hover*) as the key word, this paper aims to study the mapping of novel metaphor from a semantic computational perspective. The Latent Semantic Analysis was used to measure the relevance between the key word “飞翔” (*fly/hover*) and other lexical items in its literal contexts and metaphorical contexts respectively. Our study demonstrated that 1) the distribution of metaphorical mapping is closely related to the degree of relevance; 2) though the mapping is unidirectional from the literal to the metaphorical, fusion also occurs, suggesting a bi-directional interaction between the source and the target; 3) the metaphorical meanings of a lexical item influence its literal meaning, which can be observed in semantic prosody as well as lexical connotations.

Keywords: metaphor mapping, Latent Semantic Analysis, lexical relevance

1. Introduction

Metaphor has been a hot topic in many research fields such as linguistics, psychology, and philosophy. Apart from analyzing the imagistic and aesthetic property manifested in metaphorical language, modern scholars are more interested in the cognitive function of metaphor, which has been systematically studied by Lakoff and Johnson (1980) in their Conceptual Metaphor Theory (CMT) and many other cognitive linguists. According to CMT, thoughts are metaphorical in nature, because complex and abstract concepts must be understood and formulated by means of metaphor based on simple and concrete ideas of embodied experience.

In CMT, Metaphor is depicted as the mapping from the source domain to the target domain. This was deemed as the fundamental tenet in metaphorical analysis. The term “mapping” was originally derived from set theory in mathematics, referring to a correspondence between elements of two sets. In fact, this term was introduced into metaphor accounts in much earlier studies. In the Interaction Theory, when Richards (1936) and Black (1962) distinguished the primary subject (also called the “tenor”) and the secondary subject (also called the “vehicle”) of the metaphor, they claimed that a set of “associated commonplaces” of the secondary subject is “mapped” onto the primary subject. Lakoff & Johnson’s (1980) Conceptual Metaphor Theory further extended this mapping relationship from the lexical level to the level of conceptual domains, maintaining that more often than not, mapping took place between a number of elements of the two domains rather than merely shared properties.

Metaphorical mapping is characterized by directionality and systematicity (cf. Lakoff & Johnson 1980; Grady 2005). Directionality indicates that metaphorical mapping generally goes from the more familiar and more concrete source domain to a less familiar and more abstract target domain, not vice versa. Systematicity refers to the fact that metaphor mapping is not random, discrete, or transient. Rather, the correspondences of elements of the source domain and the target domain have established stable or even permanent connections in our conceptual system and are stored in our long-term memory, so that they can be activated readily and easily for metaphorical uses.

The mapping account offers a new perspective for metaphor research. However, most researches on metaphorical mapping, with conventional metaphors as the dominant materials, have been based on generalizing specific conceptual metaphors and then identifying sub-correspondences for mapping. These studies are of great significance, for conventional metaphors can best illustrate the stable nature of metaphor and thoughts. However, novel metaphor should not be neglected, either. Therefore this article intends to study the metaphorical mapping

of novel metaphors at the lexical as well as conceptual level by comparing the core elements of the literal use and metaphorical use of a Chinese word “飞翔” (*fly/hover*). Specifically, by means of Latent Semantic Analysis, we intend our research to be a case study without presupposing any conceptual metaphors, but rather to examine relationship between the core elements of literal context (approximate to the source domain) the core elements of metaphorical context (approximate to the target domain). We believe such an analysis is interesting and necessary because it offers a new perspective to reveal the working mechanism of metaphor ----both at the cognitive conceptual level and linguistic lexical level.

2. Research Method and Materials

Latent Semantic Analysis is a statistical method to uncover the higher-order structure implicit in the relationships between words and documents. It captures the relationships between words and texts (word-word, word-text, and text-text) in a computable form, which is believed to approximate the underlying “meaning” of words or texts. (cf. Dumais et al. 1988; Deerwester, et al. 1990; Landauer et al. 1998)

Latent Semantic Analysis is predicated on the hypothesis that there exist certain underlying semantic structures in language use, but the choice of different words often obscures these structures. However, through specific statistical techniques that filter out the “obscuring” noise, the underlying semantic structure can be revealed. That technique is Singular Value Decomposition (SVD). SVD decomposes the large word-by-document matrix into a set of approximately 50 to 150 orthogonal factors, enabling the original matrix to be represented as an approximate linear combination. Texts, and words within texts are described and represented based on vectors within this 50-150 dimensional “semantic” space. The result of the SVD ensures that this semantic space (1) retains only the major associative patterns within the data, disregarding smaller, less significant variations; and (2) places semantically-close words, texts, and word-text pairs in proximate locations, while distancing unrelated ones within the semantic space.

“飞翔”(fēixiáng, *fly/hover*) was used as the key word for retrieval in CCL Peking University, one of the largest modern Chinese corpus with the size of 15 billion Characters. The context window includes 40 characters before and after the keyword respectively. The initial search yielded 1540 instances. After manually removing repeated entries due to unspecified reasons and those where “飞翔” appeared in proper nouns (e.g., person or company names), 1331 valid instances were obtained. Among them, 709 instances were literal usage (e.g., “鸟儿在飞翔”, *birds are flying*), while 622 instances were metaphorical usage (e.g., “文学与绘画是纪伯伦艺术生命飞翔的双翼”, *Literature and painting are the twin wings that help Gibran's artistic life fly*).

Subsequent procedures include word segmentation, part-of-speech tagging, and stop-word filtering. The *Jieba* segmentation and POS tagging toolkit (one of the most popular toolkit in processing Chinese materials in NLP research) were employed for initial processing, which was followed by a human checking and manual corrections when necessary. For stop-word filtering, the Harbin Institute of Technology stop-word list was applied, primarily to remove function words such as particles, prepositions, auxiliaries, and conjunctions (e.g., “着”, “了”, “过”, “吧”, “呢”, “并且”), with an addition of 88 modal or frequency adverbs (e.g., “可能”, “也许”, “经常”, *possibly, perhaps, often*). To enhance salience, words occurring only once across the corpus were also filtered out.

As a result, the lexical bag for the literal usage comprises 1,254 word tokens, while the lexical bag for the metaphorical usage contains 1,052 word tokens. They consist exclusively of content words: nouns, verbs, adjectives, and adverbs, all expressing substantive concepts. We posit that these two lexical bags broadly represent the source domain and target domain respectively in the use of Chinese word “飞翔”(fly/hover).

3. Results and Analysis

3.1 Semantic Relevance in Computation

We employed Latent Semantic Analysis to compute the semantic relevance between the key word “飞翔” (*fly/hover*) and other words in the lexical bags ($k=50$). The computations were performed separately for the literal sentences and the metaphorical sentences. Table 1 shows the top 5 words in semantic relevance for the two categories.

Table 1. The top 5 words in semantic relevance to “飞翔”(fly/hover)

Rank	literal		metaphorical	
	word	relevance	word	relevance
1	天空(<i>sky</i>)	0.114328	翅膀 (<i>wings</i>)	0.124375
2	翅膀 (<i>wings</i>)	0.093737	自由 (<i>free</i>)	0.094835
3	自由 (<i>free</i>)	0.089423	天空 (<i>sky</i>)	0.076897
4	上空 (<i>the sky above</i>)	0.066094	展翅 (<i>spread wings</i>)	0.050161
5	展翅 (<i>spread wings</i>)	0.05975	世界 (<i>world</i>)	0.045155

Of the 5 most relevant words in the literal bag and the metaphorical bag, 4 words overlap (though the order is slightly different). The only different word is “上空(*the sky above*)” in the literal bag and “世界 (*world*)” in the metaphorical bag.

3.2 Overlaps between the Literal Bag and the Metaphorical Bag

Table 1 suggests that the metaphorical usage of “飞翔”(fly/hover) tends to “borrow” the most relevant words from the literal usage. Or technically speaking, there are overlaps between the literal bag and the metaphorical bag. However, does the overlap have anything to do with the semantic relevance? Specifically, can we conclude that the higher the semantic relevance (of a word) is, the more likely it appears in the metaphorical usage?

To answer this question, we compared all the words that appear in both bags. The results showed that the common words numbered 440 (not including “飞翔” itself), constituting 35.1% (440/1254) of the lexicon in the literal bag, and 41.8% (440/1052) in the metaphorical bag. To observe the distribution more clearly, we divided each lexicon bag in a descending order with 100 words for each scale, and examined their distribution. The results are presented in Diagram 1.

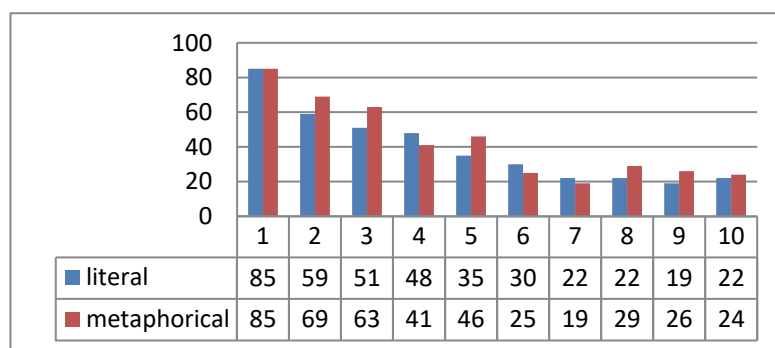


Diagram 1. The number of shared words for every 100 words (in a descending semantic-relevance order)

Diagram 1 shows that 85 words (85%) from the top 100 words in the literal bag also appear in the metaphorical bag. As for the top 300 words, 195 words (65%) from the literal bag and 217 words (72%) from the metaphorical bag are present in the other bag. As the semantic relevance decreases, the number of shared words decreases as well.

3.3 A Comparison of the Most-Related Words

The results above show that the number of shared words in the two bags is directly proportional to the semantic relevance. As there exist similarities as well as differences between two subjects in a metaphor---though only part of the similarities are highlighted in most metaphorical interpretations---it is necessary to find out the most semantically-related words in both the source domain (the literal bag) and the target domain (the metaphorical bag). The shared words represent the ground, while the different words capture the distinct features of each domain themselves. Therefore, we examined the shared and different words in the two lexical bags separately and selected the top 10 words for each type, as is shown in Table 2 and Table 3:

Table 2. The top 10 shared words in the semantic-relevance order (Note 1)

Literal Usage			Metaphorical Usage	
Rank	Word	Relevance	Word	Relevance
1	天空 (<i>sky</i>)	0.114328	翅膀 (<i>wings</i>)	0.124375
2	翅膀 (<i>wings</i>)	0.093737	自由 (<i>free</i>)	0.094835
3	自由 (<i>free</i>)	0.089423	天空 (<i>sky</i>)	0.076897
4	上空 (<i>the sky above</i>)	0.066094	展翅 (<i>spread wings</i>)	0.050161
5	展翅 (<i>spread wings</i>)	0.05975	世界 (<i>world</i>)	0.045155
6	鸟儿 (<i>birdie</i>)	0.055573	人们 (<i>people</i>)	0.040864
7	海鸥 (<i>seagull</i>)	0.052797	想象 (<i>imagine</i>)	0.037764
8	看到 (<i>see</i>)	0.048111	感觉 (<i>feel</i>)	0.035004
9	飞机 (<i>airplane</i>)	0.042023	孩子 (<i>child</i>)	0.03254
10	飞行 (<i>flight</i>)	0.035608	艺术 (<i>art</i>)	0.032045

Table 2 shows that, except for the first 4 shared words in the metaphorical bag (corresponding to the first three words and the fifth word in the literal bag), the list begins to diverge sharply even in the shared lexical bag.

Table 3. The top 10 different words in the semantic-relevance order

Literal			Metaphorical		
Rank	Word	Relevance	Rank	Word	Relevance
8	鸟类 (<i>bird category</i>)	0.049003	15	思想 (<i>idea</i>)	0.026057
12	鸽子 (<i>pigeon</i>)	0.035156	16	灵魂 (<i>soul</i>)	0.024593
20	盘旋 (<i>hover/circle</i>)	0.028549	21	心 (<i>heart</i>)	0.022942
39	使用 (<i>use</i>)	0.019065	22	理想 (<i>ideal</i>)	0.022922
40	觅食 (<i>forage</i>)	0.018838	32	天使 (<i>angel</i>)	0.0194
41	水鸟 (<i>waterbird</i>)	0.018609	36	人生 (<i>life</i>)	0.018243
46	昆虫 (<i>insect</i>)	0.017152	46	爱情 (<i>love</i>)	0.016251
49	蜜蜂 (<i>bee</i>)	0.016702	54	作品 (<i>work</i>)	0.014853
61	成群 (<i>in flocks</i>)	0.014245	64	伟大 (<i>great</i>)	0.014437
64	湖面 (<i>lake surface</i>)	0.014223	66	文学 (<i>literature</i>)	0.013209

The contrast is most obvious in Table 3 (the words ranged between the 8th to 66^h in the order of semantic relevance), with the different words belonging to the physical domain for the literal bag and to the nonphysical domain for the metaphorical bag.

3.4 The Semantic Landscape of “飞翔” (*fly/hover*)

Based on the core lexical items identified above, we can preliminarily construct a holistic semantic landscape of “飞翔” (*fly/hover*), as is shown in Diagram 2.

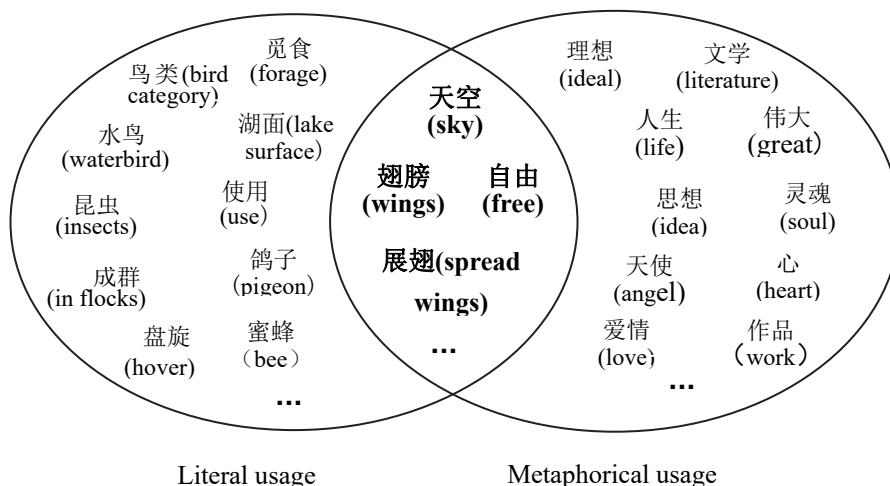


Diagram 2. The semantic landscape of “飞翔”(fly/hover)

The semantic landscape illustrates that the source domain is primarily constituted by core physical elements and actions, such as the spatial context (“天空” *sky*), the instrumental means (“翅膀” *wings*), or the dynamic process (“展翅” *spread wings*). These elements are “borrowed” and used directly in the metaphorical expressions. One might say these elements, namely the exact words from the source domain, constitute the conceptual framework for the metaphorical expressions. These expressions, together with the keyword “飞翔”(fly/hover), form a set of extended metaphors. On the other hand, the different words from each bag are the ones that actually characterize the source domain and the target domain. For example, “鸟类”(bird category), “水鸟”(waterbird), “昆虫”(insects), “鸽子”(pigeon), which are the agents of the action “飞翔”(fly/hover), correspond to “灵魂”(soul), “心”(heart), “思想”(idea) in general; The literal physical space of “天空”(sky), “上空”(the sky above), 湖面(lake surface) corresponds to the realm of “世界”(world), “文学”(literature), “艺术”(art). This well illustrates that the metaphor of “飞翔”(fly/hover) is a mapping from the concrete physical space to the abstract intangible realm, which is instantiated by both shared words as well as different vocabulary.

4. General Discussions

The term “metaphor” derived from the Ancient Greek word, meaning “to transfer” or “to carry across”. What is “carried across” with the metaphorical vehicle are not merely words, but also the imagery, conceptual structures, and underlying inferential patterns they evoke. The result of this transfer is the establishment of semantic connections (the mapping) between two conceptual domains. However, once the connections are built, the semantic influence of the two usages tends to be bidirectional rather than unidirectional, for the frequent usage of the metaphor will lead to a fusion, which allows for bidirectional flow of the meanings of the two usages.

(I) Metaphor of 飞翔(fly/hover) is More of Fusion than of Mapping

Though Conceptual Metaphor Theory (Lakoff & Johnson 1980) and Conceptual Blending Theory (Fauconnier 1997) are both important theories in cognitive linguistics, there still exist apparent differences in their accounts of metaphor. The former emphasize mapping, a unidirectional process from the source domain to the target domain; while the latter stresses blending, the results of which are a third blending space that draws upon partly from the source and partly from the target (therefore is different from the original two spaces).

What best describes the mental activities of metaphor, mapping or blending? This might be a difficult question. Fauconnier & Turner (2006) assume that in metaphor, some elements are mapped, while others undergo blending. Wolff & Gentner (2011) also note that metaphor has a dual identity: on the one hand, it is a unidirectional mapping from a source to a target domain; on the other hand, it is a process of discovering commonalities between the two domains, during which our understanding of both the target and source can be altered.

At the conceptual level, it might be true that mapping and blending is a matter of viewing perspective, for the former views the two domains separately and explores their relations, while the latter takes a holistic vantage point. However, when metaphor is examined at the lexical level as we have done, we might tentatively conclude that there are more blending than mapping---- even in the case of a novel metaphorical use as “飞翔”(fly/hover). In other words, we propose that blending dominates in the metaphorical use while mapping is more salient in simile. We illustrate our viewpoint with the following examples.

[1] 他40多年如一日，展开想象的翅膀，艰难地飞翔在湍流理论研究的特殊天空之中。(For Each day in the past 40 years, he tirelessly spread the wings of imagination and flew hard in the unusual sky of turbulence theory research.)

We can easily figure out the elements of mapping the example above: “翅膀”(wings) onto “想象”(imagination), “天空”(sky) onto “研究”(research area), which appear as lexical items in the sentence. Two of the most important elements in the source domain are used to set up the framework and make the mapping salient. However, what does “飞翔”(fly/hover) mean exactly here? The activities, the progress or the exploring (sometimes unexplainable in nature) of the research? The metaphorical meaning, which is subject to a few different interpretations, is quite vague. The mapping account obviously cannot adequately explain this. In this sense, it is the blending that leads to the emergence of the uncertain but understandable metaphorical meanings.

(II) Metaphorical Mapping is Unidirectional, But the Influence of Lexical Connotations Might be Bidirectional.

Of all the semantically-related words of “飞翔”(fly/hover), “自由”(free/freely) ranks the third in the literal bag and the second in the metaphorical bag. “自由”(free/freely) is a subjective feeling from an observer's perspective, not an inherent property of the act of flying itself. The association arises from two levels: 1) At the conceptual level, flying typically takes place in the sky, which is vast and endows people with the feeling of being free. 2) At the linguistic level, the highly-frequent cooccurrence of “自由”(free/freely) and “飞翔”(fly/hover) makes “自由飞翔”(fly freely) almost a fixed collocation. This is the case both in literal usage and metaphorical usage. Even in sentences where “自由”(free/freely) does not appear lexically, the trace of this connotation can still be detected, as in the literal use of example 2 and metaphorical use of example 3:

[2] 我望着窗外蓝天上漂浮的白云、飞驰的景物和楼群间飞翔的白色鸽群，心里有一种格外的轻松。(As I gazed at the white clouds drifting in the blue sky outside the window, the speeding scenery, and the flock of white pigeons flying between the buildings, I felt extremely at ease.)

[3] 只要你有能耐，能够登上竞争的舞台，就有一片天地让你尽情飞翔。(As long as you have the ability to step onto the stage of competition, you will have a realm to fly at will.)

Though “自由”(free/freely) does not appear, “轻松”(at ease) in example 2 and “尽情”(at will) in example 3 still creates similar feelings. In the metaphorical use, one might even argue that “自由”(free/freely) serves as the primary and permanent “ground” of “飞翔”(fly/hover), considering the uncertainty of other possible components in interpretation. Some other important collocations from both the literal bag and metaphorical bag provide more evidence. The following are modifying collocations of “飞翔”(fly/hover) in our materials collected:

Literal Usage: 自由(free), 快乐(happy), 欢快(joyful), 成功(successful), 巨大(tremendous), 不同(different), 欢乐(joyous), 优美(graceful), 喜悦(delighted), 优雅(elegant), 轻巧(defit).

Metaphorical Usage: 自由(free), 广阔(vast), 快乐(happy), 伟大(great), 重要(important), 幸福(fortunate/blissful), 欢乐(joyous), 年轻(young), 美妙(wonderful), 优美(graceful), 完美(perfect).

All the words above are positive in semantic prosody, indicating that there is also a transfer of emotions between the literal usage and metaphorical usage. Is this a unidirectional flow from literal usage to metaphorical usage? Psychologically, this is unlikely: the frequent collocation of 自由(free/freely) with 飞翔(fly/hover) in both the literal usage and metaphorical use will strengthen the association and make the connotation entrenched. The result is a reinforcement of the “ground” by both the target and source domains, causing the “ground” to become the most salient aspect in both literal and metaphorical usage. In this manner metaphor is capable of constituting the “central capacity for meaning construction” unique to humans (Fauconnier 1997).

5. Conclusion

This study employed Latent Semantic Analysis to investigate the lexical distribution patterns in both literal and metaphorical usages of “飞翔”(fly/hover). The case study contributed to metaphorical study by demonstrating a quantitative approach to characterizing and analyzing the mapping between the source and target domains of a metaphor. It allows for a clearer delineation of the semantic relationships among mapped elements. By utilizing computational natural language processing techniques, this study provides empirical support for theoretical research on metaphor and represents a novel attempt in the computational processing of metaphorical language.

References

- Black, M. (1962). *Models and metaphor: Studies in language and philosophy*. Cornell University Press.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391–407. [https://doi.org/10.1002/\(SICI\)1097-4571\(199009\)41:6<391::AID-ASII>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASII>3.0.CO;2-9)

- Dumais, S. T., Furnas, G. W., Landauer, T. K., Deerwester, S., & Harshman, R. (1988). Using latent semantic analysis to improve access to textual information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 281–287). Association for Computing Machinery. <https://doi.org/10.1145/57167.57214>
- Fauconnier, G. (1997). *Mappings in thought and language*. Cambridge University Press.
- Fauconnier, G., & Turner, M. (2006). Conceptual integration networks. In D. Geeraerts, R. Dirven, & J. R. Taylor (Eds.), *Cognitive linguistics: Basic readings* (pp. 303–372). Mouton de Gruyter.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25(2–3), 259–284. <https://doi.org/10.1080/01638539809545028>
- Richards, I. A. (1936). *The philosophy of rhetoric*. Oxford University Press.
- Tao, R., Wei, Y. P., & Yang, T. F. (2021). Metaphor analysis based on latent semantic analysis. *Journal of Donghua University (English Edition)*, 38(1), 83–90. <https://doi.org/10.19884/j.1672-5220.202010087>
- Wolff, P., & Gentner, D. (2011). Structure-mapping in metaphor comprehension. *Cognitive Science*, 35(8), 1456–1488. <https://doi.org/10.1111/j.1551-6709.2011.01194.x>

Notes

Note 1. This word list is a refined one of a similar list in a technical report in Tao, Wei, and Yang (2021). In our list, only the content words were retained.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).