

# Towards a Comparison of Japanese and English Metaphor

Rowan Hall Maudslay<sup>1</sup> Kanako Komiya<sup>2</sup> Masayuki Asahara<sup>3,4</sup> Simone Teufel<sup>1</sup>

<sup>1</sup>University of Cambridge <sup>2</sup>Tokyo University of Agriculture and Technology

<sup>3</sup>National Institute for Japanese Language and Linguistics

<sup>4</sup>The Graduate University for Advanced Studies, SOKENDAI

{rh635, sht25}@cam.ac.uk, kkomiya@go.tuat.ac.jp,

masayu-a@ninja1.ac.jp

## Abstract

In this paper, we report on preliminary experiments in which we attempt to use large language models to compare Japanese and English metaphors. More specifically, we investigate how well GPT is able to translate Japanese metaphors into English. We find that while GPT is able to produce high quality sentence translations, it is often not successfully able to identify the key metaphorical word in a longer metaphorical phrase. Nevertheless, we find that using GPT we are able to easily identify several cases of Japanese metaphor not present in conventional English.

## 1 Introduction

Language is littered with complex patterns of metaphor. For example, in English many metaphors construe debate using the language of war: just as one can “attack” an opponent in a debate, so too can claims be “defended”, arguments “shot down”, and so on. Lakoff and Johnson [1] call a pattern of this nature a **conceptual metaphor**. A conceptual metaphor is a systematic mapping that frames an abstract domain (such as ARGUMENT) in terms of a concrete domain which relates to real-world experience (such as WAR). The metaphors mentioned above are all instances of the same conceptual metaphor, ARGUMENT IS WAR, which indicates one way that argument is commonly construed by contemporary English-speaking communities.

According to Lakoff and Johnson, variation in conceptual metaphor systems between languages could influence how people in different cultures think. As an example, they suggest that speakers of a hypothetical language that has an ARGUMENT IS DANCING conceptual metaphor rather than ARGUMENT IS WAR would experience argument differently. Lakoff and Johnson’s theory, known as conceptual

metaphor theory (CMT), is the dominant theoretical framework in metaphor research, and has been applied broadly to analyse topics ranging from Nigerian ideology [2] to Shakespeare’s playwriting [3, 4].

There has yet to be any large-scale empirical comparison of the conceptual metaphors in different languages. We do not know how conceptual metaphors vary between languages, or how multilingual metaphor differences correlate with cultural variation. While there have been attempts to study cultural variation of conceptual metaphor between different languages, these are qualitative in nature, and only consider variation in a single conceptual metaphor [5, 6, 7].

Japanese is a language with many properties that differ from Western languages, and Japanese researchers have highlighted limitations in Western research which makes overly broad claims about English metaphors applying universally [8]. Later this year, in a project funded by the Japanese Society for the Promotion of Science (JSPS), we will be attempting to use computational methods to conduct a systematic comparison of English and Japanese metaphor. In this paper, we present preliminary computational explorations of the differences between Japanese and English metaphor.

More specifically, we use a large language model (LLM) to translate metaphorical expressions in a Japanese corpus into English. Our intuition is that if a Japanese metaphorical expression also occurs in English, then the English translation of the Japanese expression will be the same when it is translated in a literal and a metaphorical context. For a small number of Japanese metaphorical sentences (200), we manually evaluate the quality of the synthesized translations, and comment on the viability of this approach for the identification of differences in metaphor between English and Japanese.

## 2 Resources

In this paper, we use the BCCWJ-Metaphor dataset, a resource consisting of metaphor annotations from the Balanced Corpus of Contemporary Written Japanese (BCCWJ) [9]. The BCCWJ is a comprehensive corpus comprising 123 million words across diverse genres, including books, magazines, newspapers, white papers, blogs, online bulletin boards, textbooks, and legal documents. As of January 2025, it remains the only balanced corpus available for the Japanese language. Metaphor annotations were added by Kato et al. [10, 11] on a subset of the core data from the BCCWJ, which includes samples from magazines, newspapers, and books, covering a total of 347,094 tokens. Annotations were performed using the Metaphor Identification Procedure (MIP) [12], and metaphor classes were assigned accordingly. According to MIP, a metaphor is recognized when there is a shift from the basic meaning of a word to a more abstract or figurative meaning in a particular context. In addition to metaphor annotations for the BCCWJ, annotations have also been collected for metonymy and synecdoche, but we do not use that data in this paper.

## 3 Preliminary Experimentation

In this section, we report on initial experimentation in which we use GPT-4o [13] (accessed December 2024) to translate Japanese metaphors into English. We experiment with two varieties of the translation task, described below.

### 3.1 Translating Metaphorical Sentences

As a first translation task, we elicit sentence-level translations from GPT. Each input corresponds to a Japanese sentence containing a metaphorical expression. We experiment with 200 Japanese sentences of this nature, which are sourced from the BCCWJ-Metaphor dataset.

To preserve information about the location of the metaphorical expression in each translation, we use square brackets to demarcate the position of the metaphor in the Japanese sentence. We use a system prompt to instruct GPT to preserve the square brackets when translating. A **system prompt** is a text prompt that is given to GPT before it receives any input, which describes the task that it needs to perform on each input. The system prompt that we used for this translation task is as follows:

#### System Prompt 1

You are a machine translator. Translate the given Japanese sentences into English. In each input sentence, one expression is enclosed in square brackets. When you translate the sentence, put the corresponding expression in the English sentence in square brackets also. Take care to make sure the correct part of the translation is in square brackets.

In general, we found that GPT produced high quality Japanese to English translations, which were much better than those produced by other software such as DeepL. However, a common issue with GPT was incorrect bracketing in the translation. This occurred in 20.0% of cases (40/200). In 62.5% of these erroneous cases (25/40), GPT put the wrong phrase in the output sentence in square brackets. An example of one of these cases is shown below, with metaphorical phrases underlined:

- (1) a. その上につくられる躯体は、独自のプレウォールパネル工法により、従来の木造軸組工法の2・5倍の強度を発揮。
- b. The structure built on top demonstrates strength that is 2.5 times greater than the conventional wooden frame construction method, thanks to the unique pre-wall panel construction method.
- c. The structure built on top demonstrates strength that is 2.5 times greater than the conventional wooden frame construction method, thanks to the unique pre-wall panel construction method.

Example (1a) contains the original Japanese sentence. Example (1b) contains GPT's predicted translation. GPT has not correctly aligned the metaphorical phrase from the original with the translation. Example (1c) contains GPT's translation, but with the correct metaphorical phrase manually identified.

In the remaining 37.5% of erroneous cases (15/40), the English translation split a metaphorical phrase into several discontinuous segments, but GPT only marked one of these segments as metaphorical. For example:

- (2) a. 日本の中国侵略へのあからさまな策謀の引き金が引かれようとしていた。
- b. The blatant trigger of the scheme for Japan's invasion of China was about to be pulled.
- c. The blatant trigger of the scheme for Japan's invasion of China was about to be pulled.

In this case, the original Japanese sentence in example (1a) contains a metaphorical phrase which is split in two in the translation. GPT’s prediction in example (1b) identifies part of the phrase, but misses another part shown in (1c).

These translations reveal that what is considered to be a metaphor in the BCCWJ-Metaphor corpus is different to what is considered to be a metaphor in English resources such as the VUAMC [14]. The VUAMC is a popular corpus in which individual tokens were labelled for metaphoricity following MIP. In the BCCWJ-Metaphor corpus, unlike in the VUAMC, long multi-words spans are labelled as metaphorical. These spans include not only metaphorical words, but also other words that are relevant to the metaphorical transformation. In order to identify the metaphorical lexical item in a metaphorically-labelled phrase in the BCCWJ-Metaphor corpus, a more sophisticated approach is needed.

### 3.2 Explaining Metaphors via Translation

As a second translation task, we use GPT to automatically select the metaphorical word in a metaphorical phrase, and to additionally explain the metaphor via translation. For an input metaphorical phrase from the BCCWJ-Metaphor corpus, we request four pieces of information from GPT:

- The Japanese word in the input that is metaphorical.
- The part of speech of the selected word.
- The literal English translation of the selected word.
- The metaphorical English translation of the selected word in the context of the utterance.

These four pieces of information are obtained by requesting an output in the form of a JSON object. This is achieved using the following system prompt:

**System Prompt 2**

You are a translator. You will be given a phrase in Japanese. This word contains a target word that is metaphorical. You will return a JSON object with four fields:

The first field is called "word", and should contain the target Japanese word in the phrase that exhibits metaphor.

The second field is called "POS", and should identify the part of speech of the target word (e.g. verb, noun).

The third field is called "metaphorical\_translation", and should contain a translation of the contextual meaning of the target word.

The fourth field is called "literal\_translation", and should contain a direct translation of the target word, on its own. It should be as literal as possible, and should capture the imagery of the basic meaning of the word.

We use this prompt to elicit translations for the same 200 examples that were used in section 3.1. Example outputs are shown in Table 1.

We conduct two preliminary evaluations of the quality of GPT’s predicted outputs. These evaluations were performed by two native speakers of Japanese, who are not specially trained in metaphor annotation. First, an author of this paper went through each of the 200 examples, and identified cases which they thought to be reasonable explanations of the metaphor from the input. In this evaluation setting, we found that 82.0% of GPT’s predictions were reasonable (164/200). Second, another author of this paper went through each of the 200 examples, and produced their own explanation of the metaphorical phrase. There was no fixed form for possible explanations; we note that a wide variety of explanations for metaphors are theoretically possible, for instance based directly on the MIP, or based on violations of subcategorisation preferences. The author then determined whether GPT’s predictions aligned with their explanations. In this evaluation setting, we found that that 50.5% of GPT’s predictions aligned with the author’s own explanations (101/200). We emphasise that these numbers are indicative rather than definitive, as there were a large number of borderline cases.

The difference between the results in our two evaluation settings suggests that while GPT often produces a reasonable explanation of a metaphorical phrase, the explanation that it produces is not necessarily the explanation that would be most apparent to a human annotator. As an example, consider the metaphorical phrase 名をつける, which means “to name”. This phrase consists of three words: the noun 名 (meaning “name”), the direct-object particle を, and the verb つける (literally meaning “to attach”). For this input phrase, we expected GPT to state that the verb was used metaphorically, because an abstract

	Input Metaphorical Phrase	Selected Word	POS	Literal Translation	Metaphorical Translation
1	特徴をつかん	つかん	verb	grasp	understand
2	職を退い	退い	verb	step back	resign
3	ボール紙の山	山	noun	mountain	pile
4	バランスを壊す	壊す	verb	break	disrupt
5	人気の風向き	風向き	noun	wind direction	trend
6	歴史に幕を下ろす	幕を下ろす	verb	lower the curtain	come to an end
7	徹底的な英語指導	徹底的	adj	penetrate to the bottom	thorough
8	市場に参戦	参戦	verb	participate in battle	enter
9	金が払底	払底	adj	bottom scraped out	depleted
10	人生の荒波にもまれた	荒波	noun	rough waves	challenges
11	修業を積み	積み	verb	pile up	gain experience
12	格差縮小を視野	視野	noun	field of vision	consideration
13	依存浮き彫り	浮き彫り	verb	relief sculpture	highlighted
14	麻葉密輸ルートが芋蔓式に手繰られた	芋蔓式	adj	like a potato vine	in a chain reaction

**Table 1** Examples of GPT’s generated output for metaphorical phrases in the BCCWJ-Metaphor corpus

object such as a name cannot be physically attached to something. Instead, GPT’s explanation stated that a different sense of 名 is evoked in the metaphor, namely a sense that refers to “reputation” rather than “name”. This is not the metaphor that we expected, but it is nonetheless a reasonable explanation of a metaphor that is present.

The examples in Table 1 are all cases that both annotators thought were reasonable. There were a total of 94 cases of this nature. For each of these cases, we manually determined whether or not English has an equivalent metaphor. We found that 38.3% (36/94) of the metaphors successfully explained by GPT are common to English. Examples of these cases are shown in rows 1–6 in Table 1. For instance, the metaphors for 退い (step back → resign; row 2), 山 (mountain → pile; row 3), and 幕を下ろす (lower the curtain → come to an end; row 6) are all used in English. Some of the Japanese metaphors are not lexicalised in exactly the same way in English, but nevertheless occur in both languages. For example, while the phrase “wind direction” is not used verbatim to refer to a “trend” in English (see row 5), the common English expression “the way the wind is blowing” captures the same idea.

The remaining 61.7% (58/94) of the successfully-translated Japanese metaphors do not occur in English; examples of these are shown in rows 7–14. One case is the adjective 徹底的, which means “thorough”, but literally translates to “penetrate to the bottom” (row 7). In some cases, the Japanese metaphors appear to be based on conceptual metaphors that are not present in English. For example, in the phrase 市場に参戦 (“enter the market”; row 8), the word 参戦 means “enter” but literally

means “participate in battle”. This suggests the presence of a conceptual metaphor like ECONOMIC MARKETS ARE BATTLEFIELDS, which does not exist in English. In other cases, English has similar conceptual metaphors, but different facets of the conceptual metaphors are realised. For example, the use of 荒波 (rough waves → challenges; row 10) evokes a LIFE IS SAILING conceptual metaphor. In English many nautical phrases such as “charting a course”, “battening down the hatches”, and being “in the doldrums” can be applied metaphorically to lived experience, based on the same conceptual metaphor. However, in English rough waves are not usually construed as challenges.

Often, GPT translated metaphors which did not exist in English using alternative English metaphors. For example, the phrase 浮き彫り literally means “relief sculpture”, but has the metaphorical meaning of bringing something to attention. GPT captured this metaphorical meaning using a word with a similar metaphorical meaning in English: “highlighted” (row 13). Similarly, the phrase 芋蔓式 literally means “like a potato vine”, but is used metaphorically to describe a series of events that happen one after another. GPT captures this meaning using the English metaphorical phrase “in a chain reaction” (row 14).

## 4 Conclusion and Future Work

Our preliminary investigations indicate that LLMs could be used to automatically compare English and Japanese metaphor systems. In the future, to support quantitative analyses, we need to refine our annotation criteria, thereby making it possible to conduct systematic annotation. We also need to develop improved prompting procedures.

## Acknowledgments

This research was supported by the JSPS KAKENHI Grant Number JP22K12145, the collaborative research project “Empirical Computational Psycholinguistics Using Annotated Data” of the National Institute for Japanese Language and Linguistics, and the Kayamori Foundation of Informational Science Advancement Research Grant for “Extracting Conceptual Metaphors Using Natural Language Processing.”

## References

- [1] George Lakoff and Mark Johnson. Metaphors We Live By. University of Chicago Press, 1980.
- [2] Akin Odebunmi. Ideology and body part metaphors in nigerian english. Review of Cognitive Linguistics. Published under the auspices of the Spanish Cognitive Linguistics Association, Vol. 8, No. 2, pp. 272–299, 2010.
- [3] Sandra Peña Cervel. Macbeth revisited: A cognitive analysis. Metaphor and Symbol, Vol. 26, No. 1, pp. 1–22, 2010.
- [4] Bruce Lackie. Catch a falling kingdom: A meditation on King Lear. Metaphor and Symbolic Activity, Vol. 6, No. 4, pp. 259–270, 1991.
- [5] Shahrzad Pirzad Mashak, Abdolreza Pazhakh, and Abdolmajid Hayati. A comparative study on basic emotion conceptual metaphors in English and Persian literary texts. International Education Studies, Vol. 5, , January 2012.
- [6] Erich A. Berendt and Keiko Tanita. The ‘heart’ of things: A conceptual metaphoric analysis of heart and related body parts in Thai, Japanese and English. Intercultural Communication Studies, Vol. 20, No. 1, 2011.
- [7] Xiao Liu and Guodong Zhao. A comparative study of emotion metaphors between English and Chinese. Theory and Practice in Language Studies, Vol. 3, , January 2013.
- [8] Kazuko Shinohara and Yoshihiro Matsunaka. An analysis of japanese emotion metaphors. Journal of Yokohama Linguistic Circle, Vol. 4, pp. 1–18, 2003.
- [9] Kikuo Maekawa, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka, and Yasuharu Den. Balanced Corpus of Contemporary Written Japanese. Language Resources and Evaluation, Vol. 48, pp. 345–371, 2014.
- [10] Sachi Kato, Rei Kikuchi, and Masayuki Asahara. 『現代日本語書き言葉均衡コーパス』に対する MIP に基づく比喩表現情報の付与 (Assigning MIP-based figurative expression information to the Balanced Corpus of Contemporary Written Japanese). Proceedings of NLP2022, pp. 1427–1431, 2022.
- [11] Sachi Kato, Rei Kikuchi, and Masayuki Asahara. BCCWJ-Metaphor における比喩表現認定と情報付与作業手順 (Procedures for recognizing figurative expressions and assigning information in BCCWJ-Metaphor). Proceedings of NLP2025, 2025.
- [12] Peter Crisp, Raymond W. Gibbs, Alice Deignan, Low, Graham, Gerard Steen, Cameron, Lynne, Elena Semino, Joe Grady, Alan Cienki, and Zoltan Kövecses. MIP: A method for identifying metaphorically used words in discourse. Metaphor and Symbol, Vol. 22, No. 1, pp. 1–39, 2007.
- [13] OpenAI. GPT-4 technical report, 2024.
- [14] Gerard J. Steen, Aletta G. Dorst, J. Berenike Herrmann, Anna Kaal, Tina Krennmayr, and Trijntje Pasma. A Method for Linguistic Metaphor Identification: From MIP to MIPVU. John Benjamins, 2010.