Metaphor and Consistency in Text:

A Corpus-based Study

Kaoru Ito
Kyoto University
ito.kaoru.38r@st.kyoto-u.ac.jp

keywords: figure of speech, metaphor, text analysis, metaphor corpus

1. Introduction

In general, figure of speech is thought to be caused by the violation of grammatical or pragmatic conventions. This means figurative language may lead to miscommunication; it sometimes makes people convinced or fascinated, but also it sometimes makes an expression difficult or artificial, and perhaps will be a cause of miscommunication. This study aims to reveal what are the differences between the cases of success and failure on the use of some figures of speech. As a part of this pursuit, this paper particularly concerns the following question:

– Is a lemma (i.e., a lexical item which may appear in different surface forms) used in one meaning in a text? In particular, do we mix literal uses and metaphorical uses of a lemma in a text or not? If we mix them, how often does it occur?

To provide an answer for the question above, this study shows consistency of word meanings in texts from a metaphor corpus, The Vrije Universiteit Amsterdam Metaphor Corpus (VUAMC, see Section 3.1.1.). I provide the proportion between the lemmas which are used in a consistent meaning, and those which are used in different meanings in a text, and the proportional difference according to their word classes. In the study, I calculated the proportion of words used *consistently*, i.e. used in either metaphorical or non-metaphorical meaning through a text, and showed that nouns and adjectives tend to be used consistently, verbs less consistently, and prepositions inconsistently.

This study is broadly classified as a research from the perspective of the second group of the followings:

> … contemporary metaphor scholars can be broadly divided into two groups:
> – those who research metaphorical patterns of thought; and
> – those who research metaphorical language, and seek explanations for it.
>
> (Deignan, Littlemore and Semino 2013: 32)

Papers in Linguistic Science, No. 20 (2014), pp. 83-99

The first group of researchers attempts to understand how the metaphor works in our mind, trying to reveal the mental structure and processing involved in the production and interpretation of metaphor, whereas the second group concerns the description of *linguistic* metaphors which occur in actual contexts, to provide adequate explanatory models. From the perspective of the second group, this paper aims to provide some suggestions about the way we manage the ambiguity of possible interpretations of words distributed in a text.

2. Background

In this section, I introduce some previous studies that give implications for the meanings of a word in a text. After that, I suggest a hypothesis regarding the way in which we manage the meanings of a word when we construct a text.

2.1 Figurative language and ambiguity

In general, figure of speech is thought to be generated by the violation of grammatical conventions. One of the conversational maxims (Grice 1975: 46) "Avoid ambiguity" can be seen as a pragmatic convention that causes a rhetorical effect. Here, I would like to pick up two kinds of figures of speech that involve lexical ambiguity and relate to polysemy and homonymy: antanaclasis and syllepsis. Antanaclasis is "a figure of speech that makes a pun or paronomasia by repeating the same word, or two words sounding alike, (…), but with differing senses (Baldick 2008)." In (1), which is a quote from Benjamin Franklin, the word *hang* is used twice in different meanings. The first *hang* means 'to be united', and the second one describes physical action 'to be suspended by the neck'.

(1) We must, indeed, all *hang* together, or assuredly we shall all *hang* separately.

(Sparks 1840: 408, italics added)

On the other hand, syllepsis is a figure of speech that is more strongly involved in the ambiguity of the word meaning in a text. Among several definitions of the term syllepsis, one most relevant to our discussion is "a figure by which a word or expression is used simultaneously in its literal and figurative senses[1]" (Dupriez 1991: 440). In contrast to antanaclasis, in which a word (or more precisely a lexeme) appears twice or more in a text in different meanings, in syllepsis, two different meanings are attributed to only one occurrence. Consider the following text from BBC news.

(2) **North Sea cod: Is it true there are only 100 left?**

(By Hannah Barnes & Richard Knight BBC News)

If recent reports are to be believed, the North Sea cod's days are numbered. But

> should we believe these reports? What do the experts say about the numbers of fish that are left?
>
> The Daily Telegraph recently ran the headline: "Just 100 cod left in the North Sea". It sounded *fishy*. Trawlermen were furious.
>
> "It just makes my blood boil - 100 cod in the North Sea?" fumes Brian Buchan, who's been fishing in the North Sea for more than 30 years. "More like 100 million cod in the North Sea."
>
> (http://www.bbc.com/news/magazine-19755695, italics added.)

In the second paragraph (which begins with "The Daily…"), the word *fishy* has a double meaning – "seeming dishonest or false," and "tasting or smelling of fish." The first meaning of *fishy* is compatible with the local construction *it sounds X*. On the other hand, the second meaning is incompatible with *sounded*, and no words in the sentence evoke this meaning. Rather, this meaning seems to be activated by some elements in other sentences in the text, such as *cod*, *fish*, and *fishing*. It is declined in the phase of interpretation of the sentence, but still provides the basis that the sentence to be recognized as a pun.

2.2 Reference Chain and Domain Mapping in Text

One important notion to capture the relationship between the same words in a text is *cohesion*, defined as follows:

> Cohesion occurs where the INTERPRETATION of some elements in the discourse is dependent on that of another. The one PRESUPPOSES the other, in the sense that cannot be effectively decoded except by resource to it. When this happens, a relation of cohesion is set up, and the two elements, presupposing and presupposed, are thereby at least potentially integrated into a text.
>
> (Halliday and Hasan 1976: 4)

Halliday and Matthissen (2014: 9.2) lists four types of cohesion: (i) conjunction, (ii) reference, (iii) ellipsis, and (iv) lexical cohesion[2]. What matters here is (iv) lexical cohesion.

A speaker or writer constructs cohesion in discourse through the choice of lexical items. Halliday and Matthissen observes this type of cohesion as following: "lexical cohesion comes about through the selection of items that are related in some way to those that they have gone" (*ibid.*: 462). One of the relations that form lexical cohesion is repetition. For example, consider the example (3), in which the second occurrence harks back to the first.

(3) Algy met a bear. The **bear** was bulgy

(Halliday and Matthissen 2014: 644)

Halliday and Matthissen points out that, in order for a lexical item to be recognized as repeated, it need not be in the same morphological shape. Inflectional variants, for example *dine, dining, diner,* and *dinner,* are the same item, so they are recognized as repeated. In the case of derivational variants, when they are based on a "living" derivational process, they may be recognized as repeated. For example, the semantic relation between *rational* and *rationalize* are rather transparent, thus easily recognized as repeated, whereas the relation between *ration* and *rational* are opaque, difficult to recognized as repetition.

Because the notion of *cohesion* is established on the perspective of *text-as-a-product*, it should not be straightforwardly applied to online processing of the reader. In the perspective of text-as-a-product, the meanings of each word are already given. This means that when we judge whether two occurrence of a polysemous lexical item form lexical cohesion, we need only consider the meaning of the item in the context. For example, the other polysemous meanings do not matter that the relationships between *hang - hang* in (1) and *fish - fishy,* and *fishy - fishing* in (2) do not form lexical cohesion.

However, considering a reader's mental process, lexical cohesion can be seen as a reader's expectation for the words which have similar or the same forms to share a similar meaning. This expectation is crucial for the figures of speech mentioned above. If we did not have such expectations, we would not feel any rhetorical effect, like what we feel when we read "It sounded fishy." without any context like (2). Therefore, in this study, I assume that a reader expects words with similar or the same form to share a similar meaning, and if not, the reader feels rhetorical effect or incoherency of the text under some conditions.

Viewed from a broader perspective, this issue relates not only to repetition, but also to other types of lexical cohesion, such as synonym and hyponym. For example, in (2), not only the relations between *fish - fishy,* and *fishy – fishing,* but also *cod – fishy* are problematic in semantic aspect.

2.3 Sense Disambiguation in Natural Language Processing

Metaphorical meaning has a lot to do with sense disambiguation task in natural language processing. Metaphorical extension is one of the primary sources of polysemy and dictionary definitions of meanings of a word often include dead metaphors. Gale *et al*. (1992) conducted an experiment and suggests "one sense per discourse" hypothesis. They examined 54 pairs of polysemous or homonymous nouns extracted from the same article, and found that 51 of them shared the same sense. They conclude that with probability about 94 percent, two nouns drawn from the same article will have the same sense, and excluding doubtful nouns

as to whether they are polysemous or not, the probability moves up to 98 percent. Together with the other hypothesis "one sense per collocation" (Yarowsky 1993), "one sense per discourse" hypothesis plays an important role in the research of sense disambiguation in natural language processing.

2.4 Selectional Restriction and Metaphor Identification

Metaphor scholars have proposed some rules on how we identify metaphor. For example, Kittay (1984) posits several conditions under which an expression is interpreted as metaphorical. The most relevant here are the following two conditions. The first one is *selection restriction* violation, which refers to incongruity between intra-sentential elements. For example, in the case of *sweet voice*, the adjective *sweet* modifies noun that is not categorized as FOOD, which otherwise would modify a noun to be tasted. She suggests another criterion relating to cohesion. In the discussion of the second criterion, she illustrates the following metaphorical expression not explainable by selectional restriction violation.

(4) *The rock is getting brittle with age*. He responds to his students' questions with none of his former subtlety. His lectures also lack the verve which was characteristic of them.
(*ibid.*: italics added)

The first sentence in (4) is ambiguous, because this can be literal in the context of a geological exposition, and also it can be interpreted as metaphorical when spoken of a professor emeritus. This is disambiguated with the cohesive relationship with *he* in the second and third sentence. Kittay's study suggests both of collocation and textual environments relate to understanding metaphor, like the studies mentioned in 2.3.

2.5 Implications of Previous Studies

Considering the previous studies shown above, it would be better to assume at least two heuristics to understand an expression in a text properly: *collocational heuristic,* which focuses on a single sentence, and *textual heuristic*[3], which concerns the part of the text that the reader has already read. In other words, collocational heuristic concerns the local context, while textual heuristic concerns the global context. As suggested by previous studies referred earlier, collocational heuristic appears to be stronger than textual one. In fact, the interpretation of *fishy* in (2) is expected to be "tasting or smelling of fish" with textual heuristic, but actually it is understood as collocationally appropriate one, "seeming dishonest or false" after all. Even if we try to interpret it as textual meaning, we cannot interpret the meaning of the whole sentence. A textually invoked meaning would only survive as a punning effect. Since the question about how we interpret metaphorical expressions has

already been discussed from collocational perspective in many publications, here I shed the light on an explanation from textual heuristic. As shown here, there seem to be some motivations to use a word in a consistent meaning in a text. To investigate the detail of these motivations, the author conducted a study on consistency of word meanings, as a first step.

3. Details of the Study

3.1 Purpose

This study aims to reveal how meanings of words are used consistently throughout a text. Here, the word "consistent" is used to mean that all the tokens of a lemma are used exclusively either in metaphorical or non-metaphorical meaning throughout a text. Conversely, when a word is used in different meanings in a text, the word is used *inconsistently*. Remember that "one sense per discourse" hypothesis limits attention to the cases in which the meaning is established as one SENSE. Since this study includes the cases in which the metaphorical meaning is not conventionalized and rather novel, this study enables us to attest whether "one sense per discourse" hypothesis is applicable to metaphorical/non-metaphorical distinction in general.

Of course, this distinction of being either metaphorical or non-metaphorical is not sufficient to determine whether or not a meaning of a word is consistent. A word may have more than one literal or metaphorical meaning, and there are many subtypes in non-metaphorical meanings, such as literal and metonymic meaning. For example, *sake* has two literal meanings: *benefit* and *a kind of drink*. In this study, however, we will see the consistency in terms of metaphorical and non-metaphorical meanings primarily because the corpus is binary annotated between the two.

3.2 Data

This study uses The Vrije Universiteit Amsterdam Metaphor Corpus (VUAMC), a metaphor corpus hand-annotated to a subset of BNC baby. It consists of approximately 200,000 words with the texts classified into four genres: academic, news, conversation, and fiction. Each genre consists of approximately 50,000 words. It should be noted that the metaphor-related words in the corpus are not necessarily creative or novel; rather, metaphorical expressions are identified to be primarily based on conceptual metaphor theory. For example, in (5), the word *donkey* is marked as metaphor-related and seems to be novel metaphor, whereas in (6) the preposition *in* is also marked as metaphor-related because of its non-spatial meaning, but the metaphorical expression is already well-conventionalized, with no rhetorical novelty.

(5) The backs were mainly pedestrian but the fundamental problem lay elsewhere. Apart from Kevin Moseley's steady supply from the line-out, there was nothing much to

commend in a Welsh forward effort which reverted to the very worst Eighties stereotype of static, cumbersome *donkeys* only too willing to slow the game to walking pace.

(VUAMC: a1n-fragment09, italics added)

(6) Tiphook, which yesterday unveiled a 124 per cent increase *in* half year profits to £10million, hopes the court will block SeaCo's purchases of its own shares.

(VUAMC: a8u-fragment14, italics added)

Following the annotational criteria adopted in the corpus, in the following study we do not distinguish the types of metaphor (i.e. novel or dead).

In this study, 92 texts are chosen among various texts in the corpus. For a comparable study with the result provided in the previous study (Gale *et al*. 1992) which exclusively uses written texts, I also exclude spoken texts, using three types of written texts: Academic, News, and Fiction[4].

3.3 Method

3.3.1 Summarizing Data

The study was conducted in the following procedure, which is illustrated in Figure 1. The author extracted lemmas used more than once in a text in VUAMC, and counted the frequency of actual forms of the lemmas according to the semantic distinction between metaphor-related and non-metaphor-related meanings[5]. If the meanings of the lemma are all non-metaphorical or all metaphorical, the lemma is marked as *consistent* (CONS. in Figure 1), and, if the meanings of a lemma switch between the two uses, it is judged as *inconsistent* lemma (shown as INCONS. in Figure 1). The ratios between consistent and inconsistent lemmas were calculated for each text. The word classes examined here are nouns, verbs, adjectives, and prepositions, and the results were summed up according to their word classes.

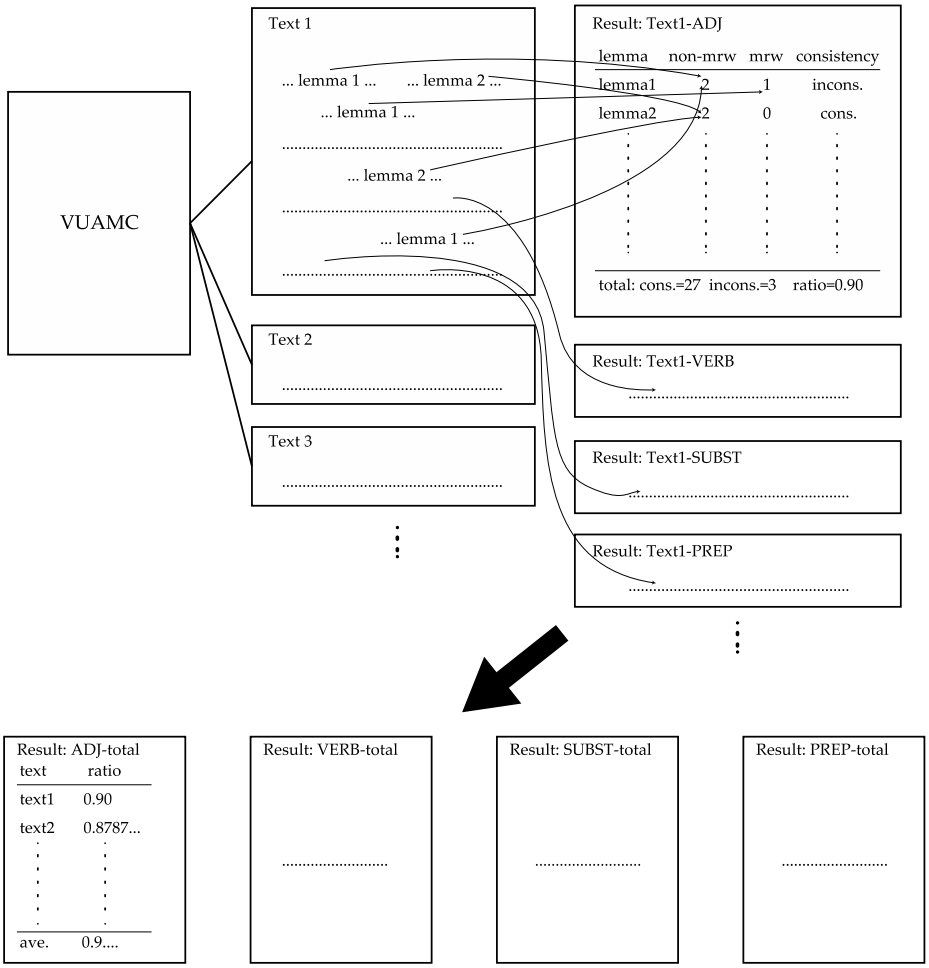

Fig. 1: Outline of the study

3.3.2 Comparison with Previous Study

To compare with Gale *et al.* (*ibid.*) which shows the probabilities of two tokens of the same lemma sharing the same meaning, I counted the frequencies of lemmas used more than once, according to their word classes. The numbers of consistent lemmas c and inconsistent lemmas i were counted for each text, and then the numbers of consistent and inconsistent lemmas in the corpus, $L_{consistent}$ and $L_{inconsistent}$, were calculated by summing up $c_1 \ldots c_n$ and $i_1 \ldots i_n$ of $\text{text}_1 \ldots \text{text}_n$, as shown in the following formulae:

$$L_{consistent} = \sum_{text\ in\ corpus} c_i$$

$$L_{inconsistent} = \sum_{text\ in\ corpus} i_i$$

As mentioned in 2.3, Gale *et al*. reported probabilities for two cases, i.e. the case of polysemous nouns (98%), and the one without excluding doubtful nouns (94%). For testing differences between the results of this study and of the previous study, one-tailed binomial test was used.

3.3.3. Comparison between Word Classes

Gale *et al*. (1992) examined tokens of the same noun if they share the same meaning or not. However, they did not consider the possibility that the consistency of their meanings may vary depending on their word classes. As mentioned above, the numbers of consistent and inconsistent lemmas are counted separately for the four word classes. I further test whether or not the consistent–inconsistent ratios between word classes significantly differ, using Steel-Dwass test.

3.4 Result

Figures 2–5 show the distribution of the word-meaning consistency in a text[6].

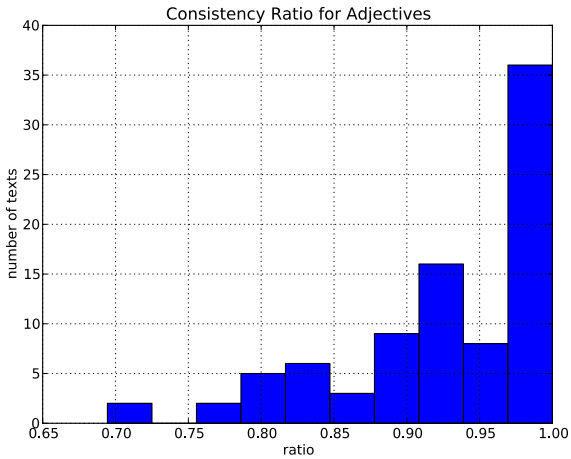

Fig. 2: Distribution of the consistency of adjectives in a text

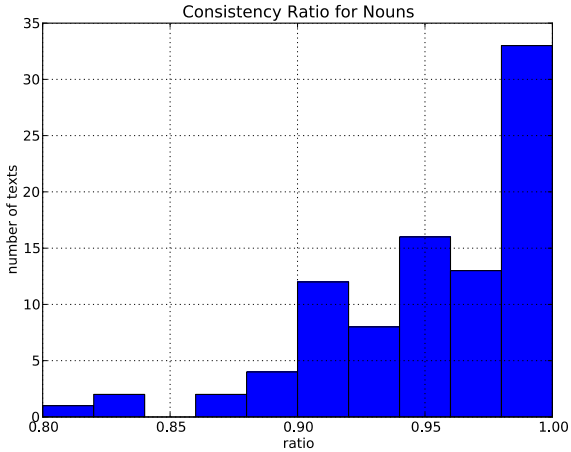

Fig. 3: Distribution of the consistency of nouns in a text

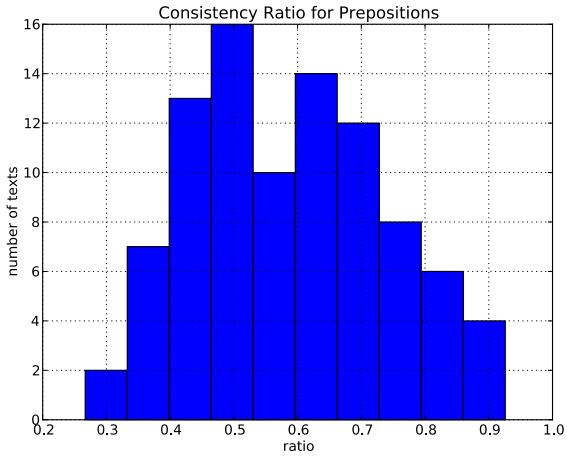

Fig. 4: Distribution of the consistency of prepositions in a text

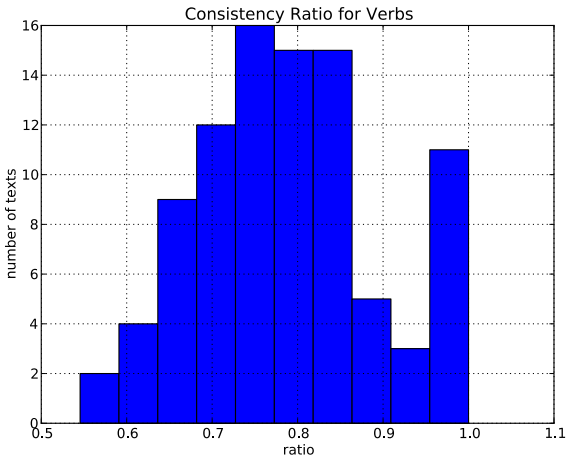

Fig. 5: Distribution of the consistency of verbs in a text

Table 1: Means and standard deviations on the consistency ratio

Word class	Number of texts	Consistency ratio	
		M	*SD*
Adjectives	87*	0.93	0.076
Prepositions	92	0.58	0.15
Nouns	91*	0.95	0.044
Verbs	92	0.79	0.11

*Texts that have no lemmas used more than once were removed.

Table 1 shows the means and standard deviations for consistency ratios according to each word class. The means of consistency ratios in the cases of nouns and adjectives were 95 percent and 93 percent respectively, notably higher than the case for prepositions (58 percent).

Table 2: Binomial test

Word class	Number of consistent lemmas ($L_{consistent}$)	Number of lemmas ($L_{consistent} + L_{inconsistent}$)	Ratio	*p*-value	
				$P = 0.98$	$P = 0.94$
Adjectives	2517	2757	0.91	< 0.001	< 0.001
Nouns	5168	5476	0.94	< 0.001	0.89
Prepositions	730	1239	0.59	< 0.001	< 0.001
Verbs	2257	2973	0.76	< 0.001	< 0.001

Note. P: Supposed probability of a lemma to be consistent in a text.

Table 2 shows the result of one-tailed binomial tests for each word class. The test was applied to two cases, depending on probability P of a lemma that appears more than once to be used in one meaning. As mentioned in 3.3.2, the values of P considered here are $P = 0.98$ and $P = 0.94$. The result of the binomial tests strongly suggests that consistency ratio is lower than the result reported in Gale *et al.* (1992), except for the case of nouns under the condition that $P = 0.94$.

Table 3: Steel-Dwass multiple comparison test

Word class comparison	*t*-value
Adjectives-Nouns	1.311
Adjectives-Prepositions	11.06***
Adjectives-Verbs	8.017***
Nouns-Prepositions	11.56***
Nouns-Verbs	9.123***
Prepositions-Verbs	8.222***

Note. ***$p < 0.001$.

Table 3 shows the result of Steel-Dwass test. The result suggests that the means of consistency ratio are significantly different in all pairs, except for the pair of adjectives and nouns.

3.5 Summary of the Study

The study shows the distributions of word-meaning consistency in a text and basic statistics. Compared with the result shown in Gale *et al*. (1992) using binomial test, as for adjectives, prepositions and verbs, the consistency ratios were significantly lower, whereas there was no significant difference between the consistency ratio for nouns and the result of previous study. Also, the result of multiple comparison suggests that nouns and adjectives tend to be used consistently, verbs less consistently, and prepositions inconsistently.

4. Discussion

There would be several factors responsible for whether a word is used consistently or inconsistently, so this section discusses possible factors to explain some of the result shown in Section 3.

4.1 Motivations to Use a Word in Consistent Meaning in a Text

As mentioned in Section 2, one of the motivations to use a word in consistent meaning can be that we should obey the rule "Avoid ambiguity". However, the result of this study has shown that prepositions are used in an inconsistent way, with different meanings of the same lemmas observed in a text. This fact may reside in the difference in their word class. Thus, this section provides possible explanation for the meaning consistency of nouns, verbs, and prepositions.

First, the meaning consistency of nouns can be explained by *conceptual autonomy* (Langacker 2008). In cognitive grammar, archetype for nouns is accounted as follows:

1. A physical object is composed of material substance.
2. We think of an object as residing primarily in space, where it is bounded and has its own location.
3. In time, on the other hand, an object may persist indefinitely, and it is not thought of as having any particular location in this domain.
4. An object is **conceptually autonomous**, in the sense that we can conceptualize it independently of its participation in any event.

(Langacker 2008: 104)

What matter here are 3 and 4. Considering these properties, the referent of a noun can exist thorough the text unfolds, and it can be referred to again and again within different events or relationships (for the relationship between cognitive archetype and behavior of each word class in discourse, see Croft (1991: 104-121)). However, this theory cannot explain why adjectives are also used consistently, or why prepositions are used less consistently than verbs.

Second, the consistency of verbs and preposition can be explained by the difference in discourse function between function words and content words. In discussing the nature of complement clauses, Thompson (2002) points out that the words or phrases which carry grammatical information, such as epistemic and evidential meanings (e.g. *I think*, or *I guess*), are treated as less prominent in a discourse, seldom traced in a subsequent utterances. This is the case in point for prepositions. Since prepositions are function words and also carry grammatical information like the verbs mentioned above, it is unsurprising that the information carried by prepositions are also seldom traced in discourse.

4.2 Entropy of Information

This study shows that the proportions of the lemmas notably differ between different word classes. It does not mean, however, that this result is unlikely to occur. In other words, this study did not show whether the result is not statistically significant. That is, the result can be caused by the mere improbability of metaphorical usage, or there can be another reason for the high consistency of the meanings. For example, let us consider the difference between the results of nouns and prepositions. Empirically, it is likely that one preposition occurs more frequently than a noun. In neutral condition, i.e. supposing there is no motivation to use a word in a consistent meaning, non-metaphorical use and metaphorical use would follow a Bernoulli distribution. This means that the more frequently a word is used in a text, the more likely the meanings of the words to vary.

4.3 What Causes the Rhetorical Effect?

As we have seen in Section 2, ambiguity is one of the causes which create a rhetorical effect. In particular, antanaclasis is, by definition, strongly relate to word-meaning inconsistency in a text. It should be noted, however, that lemmas used inconsistently in a text do not always produce a rhetorical effect. This suggests that we manage ambiguities with some devices, or we need to consider further conditions to show the ambiguity that cause rhetorical effect. One possible condition under which such an effect occur is proximity of two tokens of the same lemma. In the example of antanaclasis (1), two usages of *hang* occur proximally, which may create a conflict of two meanings. Another condition is contrast. At least, when we become aware of the effect of an antanaclasis, we must notice the difference between the two distinct meanings for each instance. If the two meanings are contrasted explicitly, we easily find the difference – regardless whether it is funny, interesting or cheesy. For example, proximity might be a device that realizes contrasting.

5. Conclusion

This study has discussed **what expectation we have when we understand a text**, and examined how a word is used in consistent meaning in a text. The result of the study has shown **nouns and adjectives tend to be used consistently, whereas prepositions tend to be used inconsistently, and verbs are between them.** This study also suggests its motivations from a conceptual and discursive perspectives, with its indeterminacy mentioned.

Notes

1. Dupriez supposes the notion of syllepsis as such. However, in my opinion, the meanings are not necessarily limited to these two. For example, they can be two literal meanings of a homonym.
2. This type of cohesion is introduced as "lexical organization" at first (Halliday and Matthissen 2014: 603), while it is referred to as "lexical cohesion" in other parts. This study refers to this type as "lexical cohesion".
3. Psychologically speaking, this heuristic may relate to priming effect.
4. A text ("b1g-fragment02") was removed from the data because of the partial lack of annotations.
5. Technically, the distinction between non-metaphorical and metaphorical uses corresponds to the XML element <w> (i.e. word element) that does not have <seg> tag and <w> that has <seg> tag, respectively. Also, the word classes mentioned above correspond to POS tags that belong to SUBST, VERB, ADJ, PREP. These classes are the groups of the tags listed in *Reference guide to BNC baby*. See the appendix for details of this classification.

6. If some texts included no adjectives or nouns that are used more than once, they were removed from the histograms.

Appendix

Table 1: Correspondence between word classes and tags listed in Burnard (2008)

Word classes	Tag
ADJ (adjective)	AJ0, AJC, AJS, CRD, DT0, ORD
PREP (preposition)	PRF, PRP, TO0
SUBST (substantive)	NN0, NN1, NN2, NP0, ONE, ZZ0, NN1-NP0, NP0-NN1
VERB (verb)	VBB, VBD, VBG, VBI, VBN, VBZ, VDB, VDD, VDG, VDI, VDN, VDZ, VHB, VHD, VHG, VHI, VHN, VHZ, VM0, VVB, VVD, VVG, VVI, VVN, VVZ, VVD-VVN, VVN-VVD

References

Baldick, Chris. 2008. *The Oxford Dictionary of Literary Terms*. 3rd Edition. Oxford: Oxford University Press.

Burnard, Lou 2008. *Reference Guide to BNC Baby.* 2nd edition. From http://www.natcorp.ox.ac.uk/baby/manual.pdf

Croft, William. 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information.* Chicago: University of Chicago Press.

Deignan, Alice, Jeannette Littlemore, and Elena Semino. 2013. Figurative Language, Genre and Register. Cambridge: Cambridge University Press.

Dupriez, Bernard Marie. 1991. *A Dictionary of Literary Devices: Gradus, A-Z*. Toronto: University of Toronto Press.

Gale, William A, Kenneth W Church, and David Yarowsky. 1992. One Sense per Discourse. *Proceedings of the Workshop on Speech and Natural Language*. 233–237. Association for Computational Linguistics.

Grice, H Paul. 1975. Logic and Conversation. In Cole, Peter, and Jerry L. Morgan (eds.), *Syntax and Semantics Vol.3: Speech Acts,* 41–58. New York: Academic Press.

Halliday, Michael A. K., and Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.

Halliday, Michael A. L., and Christian M. I. M. Matthiessen. 2014. *Halliday's Introduction to Functional Grammar*. 4th edition. London and New York: Routledge.

Kittay, Eva Feder. 1984. The Identification of Metaphor. *Synthese.* vol.58(2): 153–202.

Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.

Steen, Gerard J, Aletta G Dorst, J Berenike Herrmann, Anna Kaal, Tina Krennmayr, and Trijntje Pasma. 2010. A Method for Linguistic Metaphor Identification: From MIP to MIPVU. Amsterdam: John Benjamins.

Sparks, Jared. 1840. *The Works of Benjamin Franklin*. vol.1. Boston: Hilliard Gray.

Yarowsky, David. 1993. One Sense per Collocation. *Proceedings of the workshop on Human Language Technology*: 266–271. (Association for Computational Linguistics.)

Corpus

VU Amsterdam Metaphor Corpus (VUAMC)

テクストにおけるメタファーと一貫性
―コーパスに基づく研究―

伊藤 薫

あるテクストの中で語が異なる意味で用いられるということは、異義反復などの修辞的効果をもたらす要因の一つであると考えられる。また、ある語がメタファー的かどうかを判断する上で文脈が関わっていることからも、我々がテクスト中の文を理解するメカニズムを探る上で、ある語の意味がテクスト中でどの程度一貫しているかを探ることは重要であると考えられる。本論では、名詞、形容詞、動詞、前置詞の一貫性について VU Amsterdam Metaphor Corpus を用いて調べ、名詞・形容詞＞動詞＞前置詞の順で意味の一貫性が高いことを示した。また、意味的な側面から、上述した一貫性の差異の原因や、修辞的効果が起きる条件について考察を行った。