Semantic Clues for Novel Metaphor Generator

Rafal Rzepka†, Pawel Dybala‡, Koichi Sayama‡and Kenji Araki†
†Hokkaido University, Sapporo, Japan
{kabura,araki}@media.eng.hokudai.ac.jp
‡Otaru University of Commerce, Otaru, Japan
{paweldybala,sayama}@res.otaru-uc.ac.jp

Abstract

In this position paper we introduce our early findings drawn from a new study on an over 30,000 entries collection of metaphorical expressions manually extracted from Japanese literature. Such thoroughly chosen data assure higher standards when figurative speech is analysed from the generation perspective and allow to develop an algorithm for explaining difficult or new notions. Such function is going to be implemented in our artificial tutor project and here we describe a method for generating similes. We also show some statistical discoveries in the semantic level that could be useful knowledge for a dialog system with extended explanation capabilities.

1 Introduction

Figurative speech is one of the most spectacular thought processes used in human communication. If we need to explain some difficult word, to delicately suggest or emphasize something, we often use metaphors. Good teachers and famous poets have been using imaginative examples to help us understand things, abstract phenomena and complex or straightforward emotions. Although human beings usually have no problems with creating such examples, choosing an understandable metaphor that will trigger somebody's imagination is a difficult cognitive process [Carbonell, 1982], [Mcglone, 1996]. The most famous theories on how we understand metaphors are the categorization view [Glucksberg, 2001], the comparison view [Gentner, 1983] and three hybrid views – the conventionality view [Bowdle and Gentner, 2004], the aptness view [Jones and Estes, 2005] and the interpretive diversity view [Utsumi and Kuwabara, 2005]. This paper's goal is not to describe computational models for these methods with their pros and cons but to support researchers working on figurative speech generation as an addition to the existing theories and prepare ourselves to implement them as a part of explanation module for our artificial tutor project [Mazur *et al.*, 2012]. We used more than 12,000 of 30,000 metaphorical expressions (similes) gathered by Onai [Onai, 2005] which allows for a new approach for computing or analysing allegorical utterances in natural language interfaces. First, we show statisitcal data about words usage, then we propose a simple similes generating algorithm and finally describe preliminary experiments for setting understandability treshold. We conclude with discussion about possible usage of the new data set and introduce our plans for the fuller usage of the data.

2 Data Analysis

2.1 Onai's Dictionary

Metaphors used in this study were acquired from Onai's Great Dictionary of Japanese Metaphorical and Synonymical Expressions [Onai, 2005]. The dictionary contains metaphors selected from Japanese modern literature and Japanese translations of foreign works. The dictionary contains approximetely 30,000 metaphorical entries, each of which includes:

- headline, i.e. a word or phrase used to look up metaphors.
- sometimes – sub-headlines, i.e. words or phrases similar to the headline
- index – a phrase that was actually used in that particular metaphor (or its semantic equivalent)
- metaphor – actual metaphor example
- source – reference to the literature work from which the metaphor was selected.

According to the author, the dictionary was compiled to assist in finding interesting and somewhat sophisticated expressions that can be used instead of common phrases. If, as in the example entry in Table 2.1, one needs to find an unusual expression for "a rough woman", first he would have to query the word "woman" (headline), then search for the particular expression in the index and finally check the actual metaphor example.

2.2 Semantic Charactersistcs

To our best knowledge, the data we used is the biggest digitalized collection of Japanese metaphorical expressions and can be analysed from various angles. For the first trials with generation we have chosen the simplest and the most popular metaphorical figure of speech – a simile. A simile differs from a metaphor in that the latter compares two unlike things by saying that the one thing *is* the other thing, while simile directly compares two things through some connective, usually

	Headline	Index	Metaphor	Source
Japanese Transcription	*onna; josei*	*hageshii onna*	*hi no you ni hageshii onna*	*Jakuchou Setouchi*
English Translation	a woman; a lady	rough woman	woman rough like a fire	(a book by Jakuchou Setouchi)

Table 1: An example entry in Onai's dictionary

"like", "as" or by specific verbs like "resembles". In order to select similes from our data set, we used a manually created set of such words (marks) used in Japanese. This allowed us to retrieve 12,214 similes, on which we performed some statistical tests. By using JUMAN morphological parser[1] we have separated and ranked 3,752 unique part of speech patterns that can be helpful while generating figurative expressions. Dependency parser KNP's dictionary[2] and semantic role tagger ASA[3] were then used in order to rank most popular categories and words needed to set weights for deciding on the best simile candidate in the generation process. Most frequent semantic categories characteristic to figurative speech were colors, shapes, and patterns. The words with highest frequency are shown in Table 2 (grouped by part of speech). For comparison, semantic categories characteristic to random web text (3,500 sentences from a blog corpus) were mostly places, currencies, names, dates, organizations, education and the words most characteristic to a random web text were as follows.

Nouns:
learning, media, bad, work, information, method, enterprise, understanding, company, strength, area, necessity, relationship, usage, utilization, direction, United States, system, administration, thought, two, city, money, district, caution
Verbs:
to be visible, to know, to be divided
Adjectives:
individual, many, old

Further semantic analysis data can broaden the system's knowledge and become also helpful for recognizing figurative speech because when understanding users' utterances as metaphorical or idiomatic expression they need to be processed by using different comprehension strategies.

3 Example Usage: Novel Similes Generation

The topic of machine generating metaphors is not as quite popular as the understanding task. Most first trials were limited to narrow categories of target (topic/tenor) and source (vehicle) as in [Hisano, 1996]. Ten years later [Abe *et al.*, 2006] have tackled problem of metaphorical data insufficiency by using statistical analysis of language data to represent large scale human language knowledge stochastically.

[1]Juman System, a User-Extensible Morphological Analyzer for Japanese. Version 7.0: `http://nlp.ist.i.kyoto-u.ac.jp/index.php?Juman`

[2]Japanese Dependency and Case Structure Analyzer KNP 4.0: `http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP`

[3]ASA 1.0 - Semantic Role Tagger for Japanese Language: `http://cl.it.okayama-u.ac.jp/study/project/asa/`

Nouns	Verbs	Adjectives
eye (524)	to look like (236)	white (265)
voice (437)	to flow (204)	black (156)
water (338)	to shine (197)	beautiful (138)
sound (330)	to stand (182)	cold (135)
face (325)	to look (169)	heavy (107)
heart (309)	to fall down (161)	dark (101)
breast (282)	to put out (146)	sharp (101)
light (265)	to go (145)	big (96)
sky (235)	to move (139)	small (93)
head (218)	to feel (137)	detailed (86)

Table 2: Top 10 nouns, verbs and adjectives out of 10,658 morphological tokens found in corpus (similes only).

In order to examine the applicability of a generation task, the experimenter must conduct a metaphor generation task with a huge number of concepts, therefore Abe et al. used Japanese newspaper corpus as a base for their language model. Researchers also use the Web as a source for their models [Veale and Hao, 2007][Masui *et al.*, 2010] and utilize the latest thesauri and ontologies as WordNet [Miller, 1995] to build sophisticated generation algorithms [Huang *et al.*, 2013]. Numerous examples extracted from Onai's dictionary could be helpful for all existing approaches. Therefore we are planning to test most popular approaches in nearest future. For the basic preparations we have used the 12,214 similes mentioned in the previous section. Currently we are working on Ortony's salience imbalance theory [Ortony, 1979] which predicts possible source-target shared attributes and their positions in each ranking. Together, these concepts imply that low-high topic-source pairings should cause increases in salience of topic attributes. On Figure 1 we show the idea of two lists of attributes that describe a word in an order of occurrences. So "sweet honey" is more natural than "sweet voice" but the same adjective can describe both nouns. However, as Ortony's theory suggests, if two adjectives are from the top or bottom of the lists (distance between them increases), it is less likely that they can form an apt simile. We propose a method for calculating this distance in the following section.

3.1 Toward Calculating The Salience Imbalance

To observe which attributes could be used for a ground combining a source and a target (as "strong" in *man strong as a bear*) we experimented with two attribute lists. Preliminary tests with different data sets suggested that it is fairly probable to find common attributes between distant source and target nouns using the Japanese case frames database [Kawahara and Kurohashi, 2001], but as it is mostly verb-centered, we also created attributes lists for nouns used in our

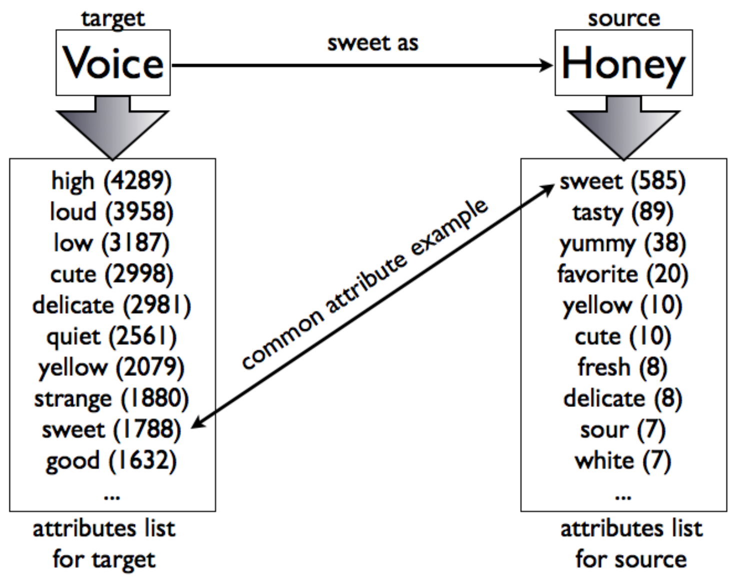

Figure 1: An example of attributes lists for a source - target pair created by ranking of adjective-noun bigram occurrences in a blog corpus. Salience imbalance should, according to the Ortony's theory, occur between attributes placed higher and lower on such lists.

trials (an example is shown in Figure 1). We concentrate on web-based textual resources because we aim at agent's dialog capabilities for using figurative speech mainly among Japanese teenagers studying English. Newspapers have not many metaphorical expressions and freely available corpus for Japanese literature consists mostly of old books with old written language. We chose one simple pattern: $Source$ - $Mark$ - $Ground$ - $Target$, which in English would be $Target$ - $Ground$ - $Mark$ - $Source$ as in *Train small like a matchbox*. The algorithm we created uses original source and target pair and compares lists of semantically related phrases retrieved from the web. For instance associations list for *fire* contains phrases as "to set", "to put out", "weak", "fast" or "extinguisher". When we input 100 random phrases according to a chosen pattern, only 16 source-target pairs had common attributes and this is because there are less adjectives than verbs in the data. We have calculated the positions on the attribute lists that are sorted according to the frequency in the web corpus so *to set fire* is on the top *to carry fire* is closer to the bottom. We have calculated the distance value between common attributes using the following formula.

$$distance = \frac{SourcePosition}{Total_{SourceAttr}}100 - \frac{TargetPosition}{Total_{TargetAttr}}100$$

For example, from the metaphor "train small as a matchbox", the system first extracts "matchbox" (source) "train" (target) and "small" (ground). Next, the rankings of attributes of "train" and "matchbox" are extracted from the case frames database, and the average position of "small" in each ranking is checked. Finally, the system (after multiplying each position by 100 to avoid dealing with very small fractions) calculates the difference of grounds position in these two rankings.

- Metaphor: Train small as a matchbox
- Source: matchbox
- Target: train
- Ground: small
- Total source attributes: 64
- Ground position in source attributes ranking (SourcePosition): 21
- Total target attributes: 7444
- Ground position in target attributes ranking (TargetPosition): 5088

$$distance = \frac{21}{64} \cdot 100 - \frac{5088}{7444} \cdot 100 = 35$$

The results are shown in Tables 3 and 4.

3.2 Preliminary Evaluation

As we plan to use the generator for a English tutoring dialog system, we need to be sure that the example chosen by computer is not too metaphorical and difficult to understand. To set the tresholds we performed a preliminary evaluation experiment and asked 5 Japanese language speakers (males only) to evaluate phrases used in the previous section. The subjects were asked to rate understandibility and aptness on the 5 points scale. The results (see Table 3) show that human evaluation averages for both aspects have set a borderline for distinguishing good similes ("voice sweet as honey" or "hand cold as ice") from lesser ones as "hand delicate as woman" which is a repetition abbreviation of *hand delicate as woman's hand*.

By using human made examples from Onai's dictionary we were able to set an average distance (D) to 20. Because setting only distance was generating too many candidates, we

source	position in A-list	attribute usualness	common ground	attribute usualness	position in A-list	target	attributive distance
daughter	2892/29614	10	YOUNG	86	8531/9872	wife	76
fire	2340/3330	70	ROUGH	9	4371/47106	woman	61
horse	6193/12416	50	LONG	0.5	494/92105	face	49.5
matchbox	21/64	33	SMALL	68	5088/7444	train	35
honey	390/1162	34	SWEET	0.7	1124/151360	voice	33.3
blade	729/1934	38	THIN	7	527/7388	lips	31
fish	7798/17082	46	COLD	21	1673/7882	hand	25
zori thongs	201/310	65	BIG	42	376/893	oysters	23
be on the watch	367/1122	33	SHARP	10	1925/19620	eye	23
ice	2604/6532	40	COLD	20	10702/53506	hand	20
bleached	6/258	2	WHITE	7	758/10686	skin	5
elephant	688/3438	20	SMALL	16	15209/93516	eyes	4
death	425/4821	9	DEEP	10	246/2436	sleep	1
blood	225/7318	3	RED	2	164/7388	lips	1
paper	66/6002	1	THIN	0.1	11/10686	skin	0.9
woman	90/47106	0.2	DELICATE	0.3	164/53506	hand	0.1

Table 3: Sixteen metaphors which had common Ground in a Kyoto Case Frames-based attributes lists. Sorted in order in distance which is the difference between attribute usualness values calculated from Ground position in both lists.

simile	average understandability	average aptness	difference between understandability and aptness	difference between attribute positions
train small as a matchbox	4,8	4,2	0,6	35
skin thin as paper	3,966	2,8	1,166	0,9
hand cold as ice	4,4	4,4	0	20
woman rough as fire	4	4	0	61
hand delicate as woman	3	3	0	0,1
skin white as bleached	5	4,6	0,4	5
hand cold as a fish	2,6	2,6	0	25
eyes small as elephant	2,2	2,8	-0,6	4
voice sweet as honey	4,2	3,8	0,4	33,3
eye sharp like being on the watch	3,6	3	0,6	23
lips thin as a blade	3,4	3,48	-0,08	31
wife young as a daughter	4,4	3,4	1	76
lips red as blood	4,8	4,4	0,4	1
oysters big as zori thongs	4	3,4	0,6	23
sleep deep as death	5	4,8	0,2	1
face long as a horse	4	3,4	0,6	49,5
averages for tresholds:	**3,96**	**3,63**		

Table 4: Results for understandability and aptness evaluation experiment. Lines in gray show similes which had both understandability and aptness higher than the averages.

have performed multiple aditional experiments with different parameters to see which conditions are helping and which are not. Results of one strict set of conditions is shown in Table 5. In most cases, if semi-correct results are counted as positive, the newly generated similes were significantly better that a random generation, but further filtering semantically strange outupts is needed.

4 Conclusions and Future Work

In this short paper we have introduced new possibilities for figurative speech generation by using a new vast collection of Japanese metaphorical expressions. Because this is an early stage of our study, we have performed only preliminary tests and experiments in order to get a grasp of which tools and other repositories can be combined before we start implementing the data to known theories about human's ability to use examples while explaining physical and abstract objects. We have already started to work with more simile patterns, also including verbs to fully utilize Kyoto Frames database. We are experimenting with N-gram frequencies of target, ground and source triplets to create vectors which should help us discover more statistical dependencies. We are also testing WordNet and ConceptNet [Liu and Singh, 2004] as a source for further calculation of semantic dependencies and show the latest progress during the workshop.

References

[Abe *et al.*, 2006] Keiga Abe, Kayo Sakamoto, and Masanori Nakagawa. A computational model of the metaphor generation process. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, pages 937–942, 2006.

[Bowdle and Gentner, 2004] Brian Bowdle and Dedre Gentner. The career of metaphor. *Psychological Review*, 111(1):193—216, 2004.

[Carbonell, 1982] Jaime Carbonell. Metaphor: An inescapable phenomenon in natural language comprehension. In Wendy Lehnert and Martin Ringle, editors, *Strategies for Natural Language Processing*, pages 415–434. Lawrence Erlbaum, 1982.

[Gentner, 1983] Dedre Gentner. Structure mapping: A theoretical frame- work for analogy. *Cognitive Science*, 7:155–170, 1983.

[Glucksberg, 2001] Sam Glucksberg. *Understanding figurative language: From metaphors to idioms*. Oxford University Press, 2001.

[Hisano, 1996] Masaki Hisano. Influence of topic characteristics on metaphor generation. In *Proceedings of 38th Annual Conference of Japanese Association of Educational Psychology (in Japanese).*, page 449, 1996.

[Huang *et al.*, 2013] Xiaoxi Huang, Huaxin Huang, Beishui Liao, and Chihua Xu. An ontology-based approach to metaphor cognitive computation. *Mind and Machines*, 23:105–121, 2013.

[Jones and Estes, 2005] Lara Jones and Zachary Estes. Metaphor comprehension as attributive categorization. *Journal of Memory and Language*, 53:110—124, 2005.

[Kawahara and Kurohashi, 2001] Daisuke Kawahara and Sadao Kurohashi. Japanese case frame construction by coupling the verb and its closest case component. In *Proceedings of the first international conference on Human language technology research*, HLT '01, pages 1–7, Stroudsburg, PA, USA, 2001. Association for Computational Linguistics.

[Liu and Singh, 2004] Hugo Liu and Push Singh. Conceptnet: A practical commonsense reasoning toolkit. *BT Technology Journal*, 22:211–226, 2004.

[Masui *et al.*, 2010] Fumito Masui, Rafal Rzepka, Yasutomo Kimura, Jun-Ichi Fukumoto, and Kenji Araki. Www-based figurative descriptions for japanese word. *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, 22(6):707–719, 2010.

[Mazur *et al.*, 2012] Michal Mazur, Rafal Rzepka, and Kenji Araki. Proposal for a conversational english tutoring system that encourages user engagement. In *Proceedings of the 19th International Conference on Computers in Education*, pages 10–12. Asia-Pacific Society for Computers in Education (ICCE2011), 2012.

[Mcglone, 1996] Matthew S. Mcglone. Conceptual Metaphors and Figurative Language Interpretation: Food for Thought? *Journal of Memory and Language*, 35(4):544–565, August 1996.

[Miller, 1995] George A. Miller. Wordnet: A lexical database for english. *Communications of the ACM*, 38:39–41, 1995.

[Onai, 2005] Hajime Onai. *Great Dictionary of 33800 Japanese Metaphors and Synonyms (in Japanese)*. Kodansha, 2005.

[Ortony, 1979] Andrew Ortony. Beyond literal similarity. *Psychological Review*, 86:161–180, 1979.

[Utsumi and Kuwabara, 2005] Akira Utsumi and Yuu Kuwabara. Interpretive diversity as a source of metaphor-simile distinction. In *In Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, pages 2230–2235, 2005.

[Veale and Hao, 2007] Tony Veale and Yanfen Hao. Comprehending and generating apt metaphors: a web-driven, case-based approach to figurative language. In *Proceedings of the 22nd national conference on Artificial intelligence - Volume 2*, AAAI'07, pages 1471–1476. AAAI Press, 2007.

English translation	Japanese in Roman letters
Lips as red as they were licking blood	Chi-o nameru you-na akai kuchibiru
Lip as thin as somebody press a blade against it	Ha-o oshitsukeru you-na usui kuchibiru
Voice so sweet as somebody was sucking up honey	Mitsu-wo suiageru you-na amai koe
Wife as young as somebody was talking to daughter	Musume-ni hanasu you-na wakai tsuma
Wife so young that resembles daughter	Musume-ni niru you-na wakai tsuma
Woman so rough like she was dropping fire	Hi-o otosu you-na hageshii onna
Sleep as deep as somebody was avoiding death	Shi-o kaihi-suru you-na fukai nemuri
Hand as cold as somebody was biting ice	Koori-o kamu youna tsumetai te
Hand as cold as somebody was putting it into ice	Koori-ni ireru you-na tsumetai te
Skin as thin as somebody was cutting off paper	Kami-o kiritoru you-na usui hifu
Skin as thin as somebody was peeling off paper	Kami-o hagu you-na usui hifu
Skin as thin as somebody was scratching off paper	Kami-o hikkaku you-na usui hifu
Skin as thin as somebody could stick paper (to it)	Kami-ni haritsukeru you-na usui hifu
Face as long as somebody was being put on a horse (back)	Uma-ni noseru you-na nagai kao
Face as long as somebody was aiming at a horse	Uma-ni ateru you-na nagai kao
Face as long as somebody was separated from a horse	Uma-kara hanasu you-na nagai kao
Wife as young as somebody was passing to daughter	Musume-ni watasu you-na wakai tsuma
Wife so young that you could get used to daughter	Musume-ni nareru you-na wakai tsuma
Wife so young that she could take away daughter	Musume-wo ubau you-na wakai tsuma
Wife so young that you could find in daughter	Musume-ni mitsukeru you-na wakai tsuma
Wife so young that you could be worried about daughter	Musume-ni kizukau you-na wakai tsuma
Wife so young that you could let her go to daughter	Musume-ni hanasu you-na wakai tsuma

Table 5: Examples for an Kyoto frames-based generation experiment for candidates with the same particles and grounds, where there were more than 30 verb-source and 30 grounds in Ameba corpus. Conditions for salience imbalance distance values were $D < 100$ and $D > 5$. First group of examples were classified by authors as correct, second as semi-correct and third as incorrect, although there was no agreement in case of few examples from the last group suggesting that imagination plays a big role in evaluating novel metaphors.