3H4-OS-17b-3

Influence of Emoticons and Adverbs on Affective Perception of Japanese Text

Rafal Rzepka*1 Urszula Jagla*2 Pawel Dybala*2 Kenji Araki*1

*1 Hokkaido University *2 Jagiellonian University

In this paper we present preliminary experimental results showing differences in readers affective impressions while reading Japanese sentences with and without adverbs and emoticons, which can be useful for analyzing and generating natural language.

1. Introduction

This paper presents an investigation that aimed at examining whether using adverbs and emoticons while writing in Japanese has any effect on how people estimate various emotions expressed in the output sentence. The main goal of our preliminary experiment was to see how emotive expressions usually used in sentiment analysis are influenced by adverbs and emoticons and how they change the affective evaluation of a sentence in Japanese. In order to investigate the differences, we designed a set of sentences for questionnaires that were later answered by Japanese native speakers. In this paper we introduce the data we utilized and the trends we observed in the results of our survey.

Previously it was confirmed that both emoticons [Hogenboom 13, Solakidis 14] and adverbs [Benamara 07, Takizawa 13] are important in sentiment analysis. In this study we decided to investigate differences in their roles in intensifying affect. This knowledge we need not only for more sophisticated recognition of emotions but also for generating utterances with more precise emotional meaning.

2. Preparing Survey Data

To create credible questionnaires we first retrieved a set of Japanese sentences that included emotive expressions from Nakamura's "Dictionary of Emotive Expressions"*1 [Nakamura 93], adverbs used by [Takizawa 13] and emoticons from [Urabe 13]. The source was YACIS [Ptaszynski 12] corpus. For assuring that emoticons are widely used, we limited them to the ones used in work of Urabe et al; they were emoticons that did not have any additional characters outside the base face borders set by brackets (e.g. symbolizing hands and gestures). To guarantee that emotion expressed by a sentence is compatible with the one expressed by an emoticon we selected only sentences that comply this requirement (which let us limit irregular cases like sarcastic usage). We also normalized politeness and manually or half-manually checked and added usual adverbs if they were missing in the retrieved sentences. The last step in preparing data for our questionnaires was to show the sentences to Japanese native speakers in order to check if they are syntactically and semantically correct. After this process we were left with 22 sentences that included at least one emotive word, one adverb and one emoticon.

Contact: Rafal Rzepka, Hokkaido University, Language Media Laboratory, Sapporo, Kita-ku, Kita 14 Nishi 9, 011-706-6535, rzepka@ist.hokudai.ac.jp

*1 The dictionary contains lexical examples categorized into ten basic emotions characteristic for Japanese: *yorokobi* (joy, delight), *ikari* (anger), *aware* (sorrow, sadness), *kowagari* (fear), *haji* (shame, shyness, bashfulness), *suki* (liking, fondness), *iya* (dislike, detestation), *takaburi* (excitement), *yasuragi* (relief), and *odoroki* (surprise, amazement).

3. Creating Questionnaires

Using those selected 22 sentences we prepared four linked questionnaires. Firstly, we made four versions of every sentence containing all elements (adverb, emotive phrase and emoticon), stripping these elements:

- with adverb and emoticon omitted;
- containing adverb only;
- containing emoticon only;
- containing both adverb and emoticon.

An example set of sentences is given below (last sentence is the unstripped original).

- *Itoshii sugata desu-ne* ("It's such a beloved appearance")
- ***Mechakucha*** *itoshii sugata desu-ne* ("It's such an **amazingly** beloved appearance")
- *Itoshii sugata desu-ne* `(^-^)` ("It's such a beloved appearance `(^-^)`")
- ***Mechakucha*** *itoshii sugata desu-ne* `(^-^)` ("It's such an **amazingly** beloved appearance `(^-^)`")

Then we half-manually divided the sentences into four sets, so that every type appears only once in each set. Finally, we randomized the order of sentences in all sets and generated the questionnaires. The purpose of separating similar sentences and randomization was to make sure that the participants would not be influenced by their previous answers.

Emotion	joy	anger	sadness	excitement	likeness	fear	dislike	surprise	shame	relief
subject 1	4	1	1	1	2	1	1	1	4	2
subject 2	4	1	1	4	4	1	1	1	1	1
subject 3	4	1	1	1	1	1	1	1	1	1
subject 4	5	1	1	1	4	1	1	1	1	1
subject 5	5	1	1	4	5	1	1	1	1	1
average	4.4	1	1	2.2	3.2	1	1	1	1.6	1.2

Table 1: Example of emotional load annotation by different users of a sentence *Kyūkutsu deshita ga, kawaikatta desu (≧ ω ≦)* ("It looked a bit tight but it was cute (≧ ω ≦)")

4. Experiment and Results

Twenty one subjects replied to our survey (one person was asked to fill in 4-5 sets). There were 17 males and 4 females, from 21 to 43 years old (average age 25.2), mostly with background in computer science (16) and few in humanities (5). We asked every subject about their emoticon usage and only two subjects answered they never use them. Six subjects always use emoticons, other six that they use them sometimes, seven admitted to use them rarely.

The task was to mark emotional load of every sentence by choosing an emotion degree on 1-5 scale where 1 is no load and 5 is a strong load. An example how users evaluated a sentence is given in Table 1.

As expected, sentences with only emotive words scored lower than those containing also adverbs and emoticons.

What was more surprising, the highest scores were not so much lower for sentences with only emoticons (36.4%) and both emoticons and adverbs (40.9%) – see Figure 1.

There were five cases (22.7%) where emotive phrase and adverb sentences conveyed the most distinct emotions, which could mean the emoticons have a stronger influence on reader's affective perception on a sentence than adverbs. However, a closer look at these examples does not allow to jump to any certain conclusions. In this section we discuss various trends we have observed.

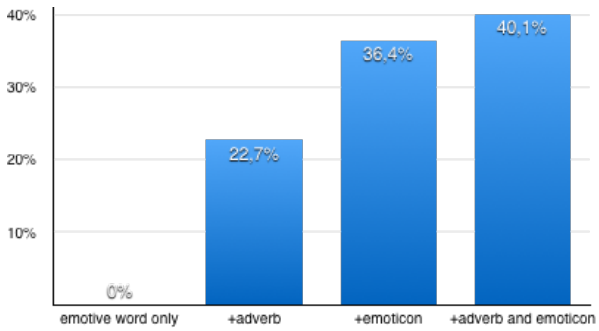

Figure 1: Rate differences between the highest scores depending on a sentence variation.

Sentences created from *Unto tanoshinde kudasai* `(^o^)` ("Please have lots of fun `(^o^)`") were evaluated as expected, meaning the highest score (joy: 4.8, like: 4.6) for the original with adverb and emoticon, lower evaluation for emotive phrase only (joy: 3.4, like: 2.8) and for emoticon only (joy: 3.8, like: 3.8). However, its most simplified version with emotive phrase only scored second high (joy: 3.7, like: 4.0). *Tanoshinde kudasai* ("Please have fun") as a dry request-like expression may be concealing emotional associations which cannot be properly conveyed by pure text, therefore a reader might be assuming involvement of other emotions. Average sums of scores for all types of sentences do not differ significantly, but in 27.3% cases the simplest sentences had the highest average totals, showing that subjects had quite different opinions and were assigning scores to emotions which were evaluated as non-existent in sentences with emoticons and adverbs. For instance, sentence *Kyūkutsu deshita-ga, kawaikatta desu (≧ ω ≦))* ("It looked a bit tight but it was cute (≧ ω ≦)") had average total of scores 20.6 (sums of all points given by subjects to a sentence), while its most simplified version's average was 21.7, because subjects have chosen all types of emotions except sadness (we do not see this phenomenon in Table 1 containing results for an emoticon version of the sentence). In case of the semantically richer original, only half categories were chosen (joy, likeness, relief, excitement and surprise, no negative interpretations). Subjects seem to be guessing what the utterer could feel in the semantically the poorest versions and probably for that reason there were also scores for anger, sadness and shame.

The same trend was seen in sentences as *Kyō-wa shinpai sugiru hi desu-ne* ("Today is the day for worrying too much"), or *Hikōki-ni noru toki-wa tanoshii desu-ne* ("I'm happy when I get on a plane"). Scores for sadness and dislike may suggest that some subjects suspect traces of irony or they project their own emotional associations while evaluating the sentences.

In other example, shame / shyness in *Watashi-mo jitsu-wa hazukashigarimono desu-kara!* ("Because actually I'm a shy guy, too!") scored higher than the version with adverb *hijō-ni* ("very") and the version with emoticon `(*^_^*)`. We think we can witness two phenomena in such cases. When it comes to using an adverb to underline one's own weakness, it might be perceived as an intensifier of author's modesty, not of the shyness itself. Smiling emoticon, on the other hand, seems to weaken the statement, making it more funny and lighter than just a straight confession.

Shōshō tanoshimi-ni shite ita-no da-ga, sore-hodo-no kangeki-wa nakatta ("I was looking for it a bit but there was no deeper emotions") is an example where adverb *shōshō* ("a bit") was influential enough to made this version score significantly higher in sadness category than the richest ver-

sion (4.3 vs 3.0). Our first intuition was that also in this case the emoticon has soften the interpretation but when we confirmed average scores, it appeared that emoticon version scored lower (3.2), and the lowest when enriched also by adverb (3.0) as mentioned above. Emotive word only sentence scored slightly higher (3.4) leaving the adverb only version to be felt as distinctly sadder. The same trend is visible in evaluation of sentence *Hannya-no Kawashima-san-ni ore-wa totemo tekii-o mochimashita* ("I feel such a hostility toward Mr. ≫Hannya≪ Kawashima") and *Nodo-ga kakudan-ni raku-ni narimashita* ("My throat is distinctly better now"). In the first case, emoticon that caused a big drop in anger evaluation (from 4.4 to 2.2) was (` ε'). As a less common one it is possible that users felt puzzled by it, especially with the epsilon symbol between the eyes. Another potential explanation is that because more than half of cases where adverbs were the most influential were conveying negative emotions, the lack of emoticons could be intensifying these emotions. Another example of interesting adverb – emoticon relation is contained in a sentence *Zutto akogareteita-kara kiai jūbun* (*^^*) ("Because I admired it for ages, I have enough of spirit (*^^*)"). It seems that omitting *zutto* ("for ages") and adding the emoticon increases the emotional strength of the sentence.

5. Conclusions and Future Work

Although the total number of sentences evaluated in this experiment was 88, they were semantically similar variants of the basic 22 originals and this number is not sufficient to draw indubitable conclusions. However, to acquire more accurate measurements of emotions in the written language, one will need to add various semantic features to diversify them and calculate their correlations. The results may be useful for language analysis and generation (e.g. opinion mining, sentiment analysis) and also for language generation (e.g. dialog systems with affective processing). The main messages for researchers in these fields are as follows.

- Ignoring emoticons may carry important consequences of incomplete emotion measurements and losing this semiotic information might be more costly than utilizing only adverbs weighting.
- Adverbs intensifying function differ from context to context so clustering methods for setting strengths as used by [Takizawa 13] must be learned on as big data sets as possible.

We concentrated on data from a blog corpus but it is not perfect in many aspects. Firstly, it reflects lack of balance between emotion categories (59.1% of all sentences were positive combinations of joy and likeness or relief). To exclude influence of different levels of politeness, we have normalized all sentences to the basic, neutral one (*-desu, -masu*). However, to acquire more accurate measurements of emotions in the written language, one will need to add various semantic features to make distinctions and calculate correlations between them. Human beings have problems with interpreting emotions in texts, therefore a perfect measurement is very difficult by default. Nevertheless, more accurate manipulation of adverbs and emoticons may help creating useful agents and tools using emotions for solving problems as depression by cognitive therapy bots described by [Rzepka 15].

References

[Benamara 07] Benamara, F., Cesarano, C., Picariello, A., Recupero, D. R., and Subrahmanian, V. S.: Sentiment Analysis: Adjectives and Adverbs are better than Adjectives Alone., in *Proceedings of The International AAAI Conference on Web and Social Media (ICWSM)* (2007)

[Hogenboom 13] Hogenboom, A., Bal, D., Frasincar, F., Bal, M., Jong, de F., and Kaymak, U.: Exploiting Emoticons in Sentiment Analysis, in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, SAC '13, pp. 703–710, New York, NY, USA (2013), ACM

[Nakamura 93] Nakamura, A.: *Kanjo hyogen jiten [Dictionary of Emotive Expressions]*, Tokyodo Publishing (1993)

[Ptaszynski 12] Ptaszynski, M., Dybala, P., Rzepka, R., Araki, K., and Momouchi, Y.: YACIS: A five-billion-word corpus of Japanese blogs fully annotated with syntactic and affective information, in *Proceedings of The AISB/IACAP World Congress*, pp. 40–49 (2012)

[Rzepka 15] Rzepka, R. and Araki, K.: ELIZA Fifty Years Later: An Automatic Therapist Using Bottom-Up and Top-Down Approaches, in Rysewyk, van S. P. and Pontier, M. eds., *Machine Medical Ethics*, Vol. 74 of *Intelligent Systems, Control and Automation: Science and Engineering*, pp. 257–272, Springer International Publishing (2015)

[Solakidis 14] Solakidis, G. S., Vavliakis, K. N., and Mitkas, P. A.: Multilingual sentiment analysis using emoticons and keywords, in *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 02*, pp. 102–109IEEE Computer Society (2014)

[Takizawa 13] Takizawa, M., Rzepka, R., and Araki, K.: Evaluating Polarity Considering Influence of Adverbs (in Japanese), in *Technical Reports of Kotoba Kogaku*, Vol. SIG-LSE-B302-3, pp. 13–18 (2013)

[Urabe 13] Urabe, Y., Rzepka, R., and Araki, K.: Emoticon recommendation for Japanese computer-mediated communication, in *Proceedings of the 2013 IEEE Seventh International Conference on Semantic Computing (ICSC)*, pp. 25–31 (2013)