

Construction of Adverb Dictionary that Relates to Speaker Attitudes and Evaluation of Its Effectiveness

Toshiyuki Kanamaru¹, Masaki Murata¹, and Hitoshi Isahara¹

National Institute of Information and Communications Technology (NICT)
3-5 Hikaridai, Seikacho, Sorakugun
Kyoto, 619-0289, Japan
{kanamaru, murata, isahara}@nict.go.jp

Abstract. Adverbs can express a speaker’s attitude in a given situation on a specific matter. We constructed an adverb dictionary in which the attitude of the speaker is described. We also looked into whether or not adverbs could effectively be used as basic data to analyze reputations. We conducted three kinds of experiments to verify how effective and precise our dictionary was. First, we calculated the coverage ratio of the dictionary by comparing the ratios of appearances of all adverbs to the ratios of appearances of our dictionary items. Next, we attached a tag to 988 adverbs, and found that the coverage ratio of the tagged adverbs was 97.76% in the open data. Finally, we classified whether sentences were positively or negatively represented using the adverb dictionary and calculated that the accuracy of the classification was 86.5%.

1 Introduction

Japanese adverbs are mainly classified into two kinds. The first is adverbs that indicate the state and mode of motion. The second is adverbs that indicate the psychological attitude or feeling of the speaker. [1, 2].

Reputation analysis¹, which excludes adverbs from the objects being processed, has been used in previous research to handle sensitive information such as the extraction of evaluative expressions. Therefore, little research has been done to investigate the importance of adverbs in language processing. Adverbs are mostly excluded during information retrieval as stop words, because they are not part of the propositional content of sentences.

Adverbs are not used in language processing such as reputation analysis or opinion extraction, because it involves the same process as information retrieval. In addition, when doing such processing, either adverb forms are not defined, or their relative frequency is not sufficiently high. Few adverbs have been extracted as important words to determine reputations in reputation analysis. We need to define adverbs manually in advance in a dictionary for reputation analysis

¹ “Reputation” for our purposes refers to whether speakers appear to evaluate a given subject as positive, negative, or neutral.

since it is different from other evaluations of expressions in that it does not use automatic acquisition or automatic classification.

We therefore constructed an adverb dictionary that describes speakers' attitudes. We found that the dictionary provides effective data for analyzing reputations.

We verified how effective the dictionary we constructed was. First, we examined its coverage ratio using a Web text. Next, we verified the accuracy of the classifications using randomly selected dictionary data from the Web text. Finally, we assessed how effective the adverb dictionary was.

2 Construction of the adverb dictionary

As was mentioned above, we need to define adverbs manually in advance in a dictionary as reputation analysis is different from other evaluations of expressions in that it does not use automatic acquisition or automatic classification. Therefore, we constructed an adverb dictionary that describes the attitudes of speakers. This section explains the method we used to construct it.

2.1 Target adverbs and attaching the value tag to them

We referred to two existing dictionaries when we constructed the adverb dictionary. The first was the Japanese EDR [3] word dictionary. The second was ipadic, which is used with ChaSen [4]. We extracted all 2592 adverbs defined in the two dictionaries. We called these extracted adverbs the "all adverb" group.

We next added to the adverbs information that had been manually extracted on the attitudes of speakers. There are various evaluations of speakers' attitudes such as how advantageous speakers think a certain matter is and how likely it is that they are thinking certain things.

The information tags we attached to the adverbs regarded attitudes that related to affirmative or negative characteristics. Adverbs that are used when speakers think that a matter is affirmative were called affirmation characteristic evaluations, and a "p" tag was attached to them. Conversely, adverbs that are used when speakers think that a matter is not affirmative were called negative characteristic evaluations, and an "n" tag was attached to them. Adverbs that were not used affirmatively or negatively or that were used affirmatively or negatively according to context were called neutral adverbs, and a "0" tag was attached to them. Adverbs with any of the above evaluative tags attached were called evaluation adverbs.

We referred to "the adverb use dictionary [5]" when attaching tags. The connotations of the adverbs in this dictionary, which are independent of the context, are placed in seven groups. The connotations of the adverbs in this dictionary depending on the context are placed in neutral group. Although the adverbs were assigned to seven groups in "the adverb use dictionary," we only used p, 0, and n to simplify our assessment of accuracy. We assumed that the seven groups could be evaluated as neutral, positive, or negative. Our evaluation

Table 1. Tagged and other adverb appearance in web text

All adverbs	74.10% (2087938 / 2817672)
Adverbs with value tags	47.76% (1345796 / 2817672)
Coverage ratio for all adverbs	64.46% (1345796 / 2087938)

of classification would have become too fine grained had we adopted the same seven levels for this dictionary.

We first chose 883 adverbs that were in the dictionary to add tags to. If an adverb was polysemous, we took its number of meanings or contexts where it was used into consideration when attaching the evaluation tag. We added a neutral tag to most of these adverbs with plural meanings. We also added a neutral tag to adverbs with both positive and negative meanings.

2.2 Survey of coverage ratio of adverbs with value tags

We first investigated adverbs with value tags that had about the some coverage ratio. The coverage ratio is not the ratio of sentences including an adverb to all sentences, but the ratio of sentences with tagged adverbs to sentences with adverbs in the all-adverb group. If the coverage ratio in the dictionary was high enough, attaching more tags would have been unnecessarily time consuming. However, if the coverage ratio was too low, adding tags to additional adverbs that were not in the dictionary would have been necessary.

We used 5,635,345 sentences that were extracted from a Web text to assess the coverage ratio. We divided these in half to designate closed and open data for evaluating the coverage ratio. First, we searched using character string agreement for any of the adverbs in the all adverbs group in the closed data. Next, we searched for adverbs with value tags in the same way. Many sentences correspond when searching for character string agreement. We searched for character string agreement in the sentences so as to be able to find as many sentences with adverbs as possible. Finding the coverage ratio of parts of speech is usually done using morpheme analysis; however, some adverbs are not recognizable with morpheme analysis.

Table 1 lists how many sentences from the “closed” half matched, using character-string agreement, adverbs in the all-adverbs group, how many sentences matched adverbs with tags, and the ratio of the tagged-adverb matches to the all-adverbs group matches.

The ratio of sentences with appraisal-tagged adverbs to sentences with an adverb that belongs to the all-adverb group was 64.46%. The coverage ratio was low, and it was necessary to tag more adverbs. We extracted a list of adverbs with no value tag from all adverbs that appeared in over 0.01% of sentences to select adverbs that could be newly tagged.

The following kinds of adverbs in Table 3, including those in Table 2, were selected for possible tagging. The probability that these adverbs will match with

Table 2. Adverbs not yet tagged (20 most frequently used)

Adverbs	Frequency	Appearance ratio	Adverbs	Frequency	Appearance ratio
<i>you</i>	153717	5.455461	<i>nochini</i>	10132	0.359588
<i>ito</i>	59405	2.108301	<i>iya</i>	9848	0.349508
<i>ete</i>	55248	1.960768	<i>minnna</i>	9797	0.347698
<i>ika</i>	37263	1.322475	<i>kaku</i>	9543	0.338684
<i>chira</i>	33609	1.192793	<i>poi</i>	9216	0.327079
<i>zen</i>	26095	0.926119	<i>ichido</i>	9019	0.320087
<i>ten</i>	21767	0.772517	<i>ryaku</i>	6421	0.227883
<i>sou</i>	17294	0.613769	<i>masa</i>	6369	0.226038
<i>ban</i>	17231	0.611533	<i>koremade</i>	6287	0.223127
<i>mochi</i>	16504	0.585732	<i>issai</i>	6124	0.217343

Table 3. Adverbs or strings removed from the all-adverb group

Continuous form of adjective
Word stem of adjective
Word formed below two Hiragana characters
Word formed below two Katakana characters
Word formed from one Chinese character
Word formed with noun and particle
Word common to another part of speech
Word included in other word

portions of words is high when using a character string search. As a result, there is the possibility of erroneous results. Accordingly, adverbs corresponding to the conditions in Table 3 were manually removed from the adverb group.

Table 4 shows how many sentences matched adverbs in the all-adverbs group before and after removing the adverbs fitting categories in Table 3, how many sentences matched adverbs with tags, and the ratio of the tagged-adverb matches to the corrected all-adverbs group matches. Furthermore, because this correction was applied only to the all-adverb group, there was no change in the number of tagged adverbs.

When adverbs and strings conforming to these conditions were removed from the all-adverbs group, sentences in the web text containing matching adverbs in the search decreased 20.46% to become 53.64%. When the ratio of tagged adverbs to all adverbs was calculated again, the coverage ratio was 89.04%.

After extracting adverbs whose appearance ratio was over 0.01% from the all-adverb group and removing adverbs and strings that conformed to the conditions in Table 3 from the extracted adverbs as described above, 105 adverbs remained.

The coverage ratios in the open data after adding adverbs is shown in Table 5. The coverage ratio for all tagged adverbs became 97.76% in the open data after the adverbs were added.

Table 4. Tagged and other adverb appearance in web text (after correction)

All adverbs (before correction)	74.10% (2087938 / 2817672)
All adverbs (after correction)	53.64% (1511389 / 2817672)
Adverbs with value tags	47.76% (1345796 / 2817672)
Coverage ratio for all adverbs	89.04% (1345796 / 1511389)

Table 5. Appearance ratio and frequency of appearance of adverbs in Web text (after adding adverbs)

All adverbs (before correction)	69.90% (1969440 / 2817673)
All adverbs (after correction)	50.20% (1414559 / 2817673)
Adverbs with value tags (before adding new adverbs)	44.52% (1254317 / 2817673)
Adverbs with value tags (after adding new adverbs)	49.08% (1382827 / 2817673)
Coverage ratio for all adverbs	97.76% (1382827 / 1414559)

2.3 Attaching value tags to adverbs

In the foregoing, new adverbs chosen to be tagged were taken from those occurring in at least 0.01% of sentences, with certain types of strings disqualified from counting as adverbs. Tags were newly attached to the remaining 105 adverbs. We used the “Bunrui Goi Hyo database [6]” when affirmation, denial, and neutral tags were attached to newly added adverbs. The adverbs in the Bunrui Goi Hyo database that are in its taxonomic vocabulary chart are lined up in numerical sequence, where they are changed into other closely related adverbs. Because of this, some adverbs have already had tags from that database (as opposed to our tags) attached, and others that are closely related to them but have no tags yet are lined up with them in the listing. Murata et al. [7] gave an example of a dictionary construction where meanings were sorted. Our method can be considered the same as theirs. There is an actual example listed in Table 6.

We found that appraisal tags could be attached beforehand using the Bunrui Goi Hyo database. We could attach tags efficiently and quickly. The adverbs marked with *, are new ones to be added to Table 6. As we can see, there are already adverbs with tags from the database (not our tags) at the top and the bottom of the table. We could attach tags to adverbs by referring to these. In the case described above, we attached a “0 ” tag to the marked adverbs. Furthermore, we manually attached tags to adverbs that were not in the Bunrui Goi Hyo database.

3 Appraisal of the adverb dictionary

3.1 Appraisal experiment on the adverb dictionary

We did an experiment to check the precision of tagged adverbs. We extracted 1000 sentences at random from web texts, attached tags to adverbs that corre-

Table 6. Example of attaching tags using Bunrui Goi Hyo

Category number	Adverb	Value
3100030201	<i>kesshite</i>	0
3100050101	<i>douyara</i>	p
3100050102	<i>tabun</i>	0
3100050103	<i>osoraku</i>	n
3100050104	<i>ookata</i>	0
3100050105	<i>tashika*</i>	“0”
3100050106	<i>taigai</i>	0

Table 7. Experimental results for adverb dictionary (before adding new adverbs)

	Recall	Precision	F-measure
All	85% (175 / 205)	85% (175 / 205)	85%
Neutral	89% (151 / 170)	94% (151 / 160)	91%
Positive	78% (18 / 23)	55% (18 / 33)	64%
Negative	50% (6 / 12)	50% (6 / 12)	50%
Positive and negative	69% (24 / 35)	53% (24 / 45)	60%

sponded to those in the adverb dictionary, and classified them into affirmative, negative, and neutral sentences.

Affirmative and negative sentences are defined in the following.

Affirmative sentences were defined as sentences whose contents are regarded as desirable by the writer. In contrast, negative sentences were defined as sentences whose contents are regarded as undesirable by the writer. Neutral sentences were neither affirmative nor negative. On the basis of these criteria, answer tags were attached manually to extracted sentences.

When there were adverbs from the dictionary in sentences, we classified them as affirmative “p”, negative “n”, or neutral “0”.

The appraisal experiment used 883 adverbs before data was added to the dictionary and then 988 adverbs after data was added.

3.2 Results and considerations

The results of the experiment are listed in Tables 7 and 8. Table 7 has the results before the new adverbs were added to the adverb dictionary. Table 8 has the taxonomic results after adverbs were added to the dictionary.

All the correct answer ratios in Tables 7 and 8 correspond to affirmative, negative, and neutral classifications in which the adverbs were in the dictionary as intended. Affirmation and denial, from judgments of affirmation, denial, and neutrality that use the dictionary, are only precise where the output is affirmation or denial calculated by recall factors. Neutrality, affirmation, and denial in the sentences were tabulated both manually and using the tagged adverbs to provide precision and recall factors for every adverb classified by the dictionary. For

Table 8. Experimental results for adverb dictionary (after adding new adverbs)

	Recall	Precision	F-measure
All	86% (200 / 233)	86% (200 / 233)	86%
Neutral	89% (166 / 187)	95% (166 / 175)	92%
Positive	84% (27 / 32)	60% (27 / 45)	70%
Negative	50% (7 / 14)	54% (7 / 13)	52%
Positive and negative	74% (34 / 46)	59% (34 / 58)	65%

example, only affirmations corresponding to sentences targeting those that are judged both by manual affirmative appraisal and the dictionary were used to calculate precision and recall factors.

The recall of items that were classified by the dictionary as correct is the ratio of those that were correctly interpreted. The precision is the ratio of correct answers for classified items. In addition, the F-measure was calculated to evaluate performance. The F-measure was calculated as the harmonic average of precision and recall factors here.

There were 205 sentences that had an adverb in the dictionary before new adverbs were added, and 233 sentences after adverbs were added.

The classification results for the attitudes of speakers obtained by the adverb dictionary were as follows; when we used the dictionary before newly adding adverbs for the judgment, the total precision was 85% and the F-measure of affirmative or negative characteristics was 60%. When we used the dictionary after newly adding adverbs for the judgment, the total precision was 86% and the F-measure of affirmative or negative characteristics was 65%. When a base line of three classifications, 33%, is considered, there is relatively high precision.

The F-measure in classifying affirmative characteristics using the dictionary before adverbs were added for the judgment is 64% and 70% after. The F-measure for negative characteristics is 50% before and 52% after, which is low compared to the classification of affirmative attitudes.

4 Related Studies

Little research that designates adverbs as objects has been done focusing on speaker attitudes that are not influenced by context. No research has tried to classify general documents in terms of affirmation and denial.

Research on the subjective characteristics of adjectives such as affirmative and negative connotations has been done [8, 9]. Adjectives were chosen that had affirmative and negative characteristics to become a seed, and they were used to extract other adjectives that co-occurred to judge the subjective characteristics of each in view of the co-occurrence. However, no experiments to calculate the precision of document classification have yet been done that have made use of the affirmative and negative characteristics of adjectives.

5 Conclusion

We constructed an adverb dictionary in which the attitude of the speaker was described. We also looked into whether or not an adverb could be effectively used as basic data for reputation analysis. We conducted three kinds of experiments to verify how effective and precise our dictionary was. We first calculated the coverage ratio of the dictionary by comparing the ratios of appearances of all adverbs to the ratios of appearances of our dictionary items. We finally attached a tag to 988 adverbs, and found that the coverage ratio of the tagged adverbs was 97.77% in the open data.

We then classified various sentences using the adverb dictionary, and calculated how accurately they were classified. We classified whether sentences had a positive or negative connotation with the dictionary and found that the accuracy was 86.5%. The F-measure in classifying affirmative connotations using the dictionary after adverbs were added to the database for the judgment was 70%, and the F-measure in classifying negative connotations using the dictionary after adverbs were added for the judgment is 52%.

If we use this dictionary, it is possible to discover any misuse of adverbs since positive or negative connotations of a sentence can be separately evaluated.

References

1. Nakamichi, M. and Ariga, C.: Function of the adverb in feeling expression. *new-block Nihongo Gaku (Japanese language study)* **12**(1) (1993) 85–93
2. Morimoto, J.: *Hanashite no Shukan wo arawasu fukushi ni tsuite* About the adverb which displays the subjectivity of the speaker. Kuroshio Publishing (1994)
3. EDR: EDR Electronic Dictionary Technical Guide. EDR (Japan Electronic Dictionary Research Institute, Ltd.) (1993)
4. Matsumoto, Y., Kitauchi, A., Yamashita, T., Hirano, Y., Matsuda, H., and Asahara, M.: Japanese morphological analysis system ChaSen version 2.0 manual 2nd edition. (1999)
5. Hida, Y. and Asada, H.: *Gendai Fukushi YouHou Jiten* The adverb use dictionary. Tokyodou Publishing (1994)
6. NLRI: *Bunrui Goi Hyou*. Shuuei Publishing (1964)
7. Murata, M., Kanzaki, K., Uchimoto, K., Ma, Q., and Isahara, H. In: *Meaning Sort — Three examples: dictionary construction, tagged corpus construction, and information presentation system* —. Springer Publishing (2001) 305–318
8. Hatzivassiloglou, V. and McKeown, K.R.: Predicting the semantic orientation of adjectives. In Cohen, P.R., Wahlster, W., eds.: *Proceedings of the Thirty-Fifth Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, Somerset, New Jersey, Association for Computational Linguistics (1997) 174–181
9. Hatzivassiloglou, V. and Wiebe, J.: Effects of adjective orientation and gradability on sentence subjectivity (2000)