Argument/Adjunct Distinction Criteria in Japanese
日本語における必須・副次補語区別クライテリア

Pierre Marchal[1,2] Thierry Poibeau[3] Yves Lepage[2]
[1]ER-TIM, INaLCO, Paris, France
[2]IPS, Waseda University, Kitakyūshū, Japan
[3]LaTTiCe, CNRS – ENS – Université Paris III, Montrouge, France
pierre.marchal@inalco.fr, thierry.poibeau@ens.fr, yves.lepage@waseda.jp

1 Introduction

This paper tackles the problem of distinguishing, among verbal complements, between arguments and adjuncts (hereafter AAD for "argument/adjunct distinction"). This topic has been studied extensively in natural language processing, notably in dependency analysis, predicate frame acquisition, and semantic role labeling. However, only a few works tried to account for the existence of a continuum between arguments and adjuncts.

In this paper, we introduce two linguistically motivated AAD criteria developed for the case of Japanese, with respect to the continuum between arguments and adjuncts, and compared them to a baseline, following different evaluation methods, on more than 30,000 predicative structures.

2 Related work

First attempts to automatically distinguish arguments from adjuncts appear in the early 90's with works on subcategorization frame acquisition and prepositional phrase attachment ambiguity. Since then, only a few of studies [8] took into account the hypothesis of a continuum between arguments and adjuncts, often expressed in traditional linguistics.

As for Japanese, linguists have shown interest in developing specific distinction criteria [5] as well as validating existing criteria formerly developed for other languages [4]. To our knowledge, however, no work tackles the AAD problem in Japanese in an automatic way. Studies on related topics (*e.g.* predicate frame acquisition, corpus annotation) barely evoke the subject.

The contribution of this paper to the AAD is double. To our knowledge, this is the first work to treat in an automatic way the AAD problem in Japanese with respect to the continuum between arguments and adjuncts.

3 Methodology

3.1 Modeling the continuum between arguments and adjuncts

We approximate a representation of the continuum between arguments and adjuncts by modeling the two extremes of the continuum with reliable examples of arguments and adjuncts. We call these complements *prototypical arguments* and *prototypical adjuncts* respectively.

We define a prototypical argument as a type of complement which appears in *every* predicate frame (*i.e.* usage) of a given verb. We define a prototypical adjunct as a type of complement which appears in *none* of the predicate frames of a given verb.

Given a predicate frame lexicon we propose to build two sets of verbs: verbs which lead to the identification of prototypical arguments, and verbs which lead to the identification of prototypical adjuncts.

We model each extreme of the continuum by extracting predicative structures of verbs from these two sets from a parsed corpus.

3.2 The distinction criteria

Distinction criteria are implemented so as to compute a degree of autonomy for a (head-noun,case-marker,verb) triple noted (h, c, v). A degree of autonomy ranges from 0 to 1, 0 corresponding to an argument and 1 corresponding to an adjunct. When a situation is undecidable the degree of autonomy should be equal to 0.5.

Ordering of complements According to [9], arguments tend to be closer to the verb than adjuncts do. From a theoretical point of view, this criterion may seem irrelevant to Japanese because the order of the complements before the verb is not strictly defined. However, according to [3], there exists a natural order of the complements.

We express the following constraints for implementation of this criterion:

- Complements close to the verb should have a lower degree of autonomy than more distant complements.
- When only one complement is attached to the verb, or when a complement is at a median position – the degree of autonomy should be neutral (*i.e.* 0.5).
- The measure should not penalize prototypical arguments which eventually get distant from the verb, or prototypical adjuncts which get close to it.

Let us consider a verb v and a set of n complements (*i.e.* a verb and all its syntactic dependents). Complements are numbered from 1 to n. 1 is the most distant complement to the verb, and n is the closest complement to the verb. We define the distance as the ratio between the position i of a complement and the total number of complements. We add $\frac{1}{2}$ to the numerator so as to capture undecidable cases.

$$dist(h_i, c_i, v) = \frac{(n-i)+\frac{1}{2}}{n} \tag{1}$$

The degree of autonomy of a (h, c, v) triple corresponds to the arithmetic mean of all of its distance scores computed from all the examples it appears in.

$$auto_{ord}(h, c, v) = \frac{\sum dist(h, c, v)}{C(h, c, v)} \tag{2}$$

Joint productivity According to [2], adjuncts tend to appear with a broader range of verbs than arguments. Also, verbs tend to have a stronger selection for case-markers introducing arguments than for those introducing adjuncts. That is, (h, c) pairs yielding a high productivity[1] tend to indicate adjuncts, while (c, v) pairs yielding a high productivity tend to indicate arguments.

We express the following constraints for implementation of this criterion:

- Complements exhibiting a high productivity should have a higher degree of autonomy than complements exhibiting a low productivity.
- (c, v) pairs with a higher productivity introduce complements with a lower degree of autonomy. (c, v) pairs with a lower productivity introduce complements with a higher degree of autonomy.
- When there is only one occurrence of a pair, its degree of autonomy should be neutral (*i.e.* 0.5).
- The two measures of productivity should be combined so as to compute the degree of autonomy of a (h, c, v) triple.

[1] Productivity is somehow comparable to the notion of conditional entropy, except for using raw counts of unique cooccurrences instead of probabilities.

Let us consider h and c, a head-noun and a case-marker respectively, and V the set of all verbs. In order to produce a value between 0 and 1 we normalize the productivity of the (h, c) pair using its frequency of occurrences. We add 1 to the frequency (*i.e.* the denominator) to capture undecidable cases.

$$prod_{h,c}(h, c) = \frac{|\{v' \in V : \exists (h, c, v')\}|}{\left(\sum_{v' \in V} C(h, c, v')\right) + 1} \tag{3}$$

Let us consider c and v, a case-marker and a verb respectively, and H the set of all head-nouns. The same normalization and smoothing techniques as in (3) are applied here too. Also, contrary to the (h, c) productivity measure, here a high productivity indicates an argument. To comply with our convention (*i.e.* 0 for arguments, 1 for adjuncts) we take the difference to 1.

$$prod_{c,v}(c, v) = 1 - \frac{|\{h' \in H : \exists (h', c, v)\}|}{\left(\sum_{h' \in H} C(h', c, v)\right) + 1} \tag{4}$$

Finally, these two measures of productivity are combined by computing their geometric mean.

3.3 Evaluation

Evaluation is performed on predicative structures of the model, containing at least one prototypical argument and one prototypical adjunct.

We compute the degree of autonomy of each prototypical argument and each prototypical adjunct. If the complements are properly ordered along the continuum – that is, if all prototypical arguments were assigned a lower degree of autonomy than any prototypical adjunct – then we consider the predicative structure as correctly analyzed.

4 Experiments

4.1 Setup

Model We built two sets of verbs using 動詞項構造シソーラス (*dousikoukouzou sisoorasu*, verb-argument structure thesaurus, hereafter VAST) [10] as a predicate frame lexicon. We considered verbs containing a *wo*-marked argument in all their predicate frames as clues for identifying prototypical arguments, and verbs containing a *de*-marked argument in none of their predicate frames as clues for identifying prototypical arguments. A description of the two sets of verbs is given in Table 1.

We used the 2009 edition of the Balanced Corpus of Contemporary Written Japanese [7] as a corpus of raw text. Predicative structures of verbs were extracted using the two sets of verbs and CaboCha[2] [6]. A description of the two sets of examples is given in Table 2.

[2] CaboCha version is 0.64. We use MeCab (version 0.993) for part-of-speech tagging and morphological analysis, and ipadic (version 2.7.0) as a Japanese part-of-speech dictionary. We use the default configuration.

Table 1: The two sets of verbs. Ratios in parentheses indicate the coverage of VAST.

	Verbs	Predicate-frames
wo-set	2,560 (49%)	4,671 (46%)
de-set	4,954 (95%)	9,416 (91%)
VAST	5,190 (100%)	10,364 (100%)

Table 2: The two sets of examples. Table shows the number of examples retrieved as well as the number of prototypical arguments and prototypical adjuncts.

	Count	Arguments/adjuncts
wo-set	1,041,818	504,391 (*wo*)
de-set	1,890,151	144,481 (*de*)

Baseline As baseline, we use two measures used in [1]. *Selectional preference* is defined as the probability for a head-noun to cooccur with a given (c, v) pair. *PMI* measures the association between a verb and a case-marker.

Evaluation As test data, we built a set of examples containing at least one prototypical argument and one prototypical adjunct, *i.e.*, examples appearing in both sets of examples and containing at least one *wo*-marked complement and one *de*-marked complement. The test data consists in 31,531 unique examples.

In addition to our own evaluation method, we considered evaluations from previous works on the AAD. At a lexical level, we evaluate the number of correctly classified complements. At a clause level, we evaluate the number of correctly analyzed clauses, that is, clauses where all complements have been correctly classified.

Both methods above require a threshold to classify complements between arguments and adjuncts. We used an optimized threshold inspired by [1]. It is defined as the value that splits the test data into two groups of complements which sizes correspond to the exact number of prototypical arguments (for the lowest values) and the exact number of prototypical adjuncts (for the highest values).

4.2 Results

The results of the evaluation are given in Table 3.

The results of the two baselines, as evaluated with M1, are similar to those obtained by [1] with the same evaluation method. This gives credit for the validity of our model.

With our continuum-based evaluation method M3, the ordering criterion performed the best with almost 82 percent of correctly analyzed examples. The selectional preference baseline performed the worst with only just over 35 percent of correctly analyzed examples. The accuracy obtained with the joint probability criterion is surprisingly low, and should be thoroughly investigated. Combining the ordering criterion with the joint productivity criterion did not yield any improvement.

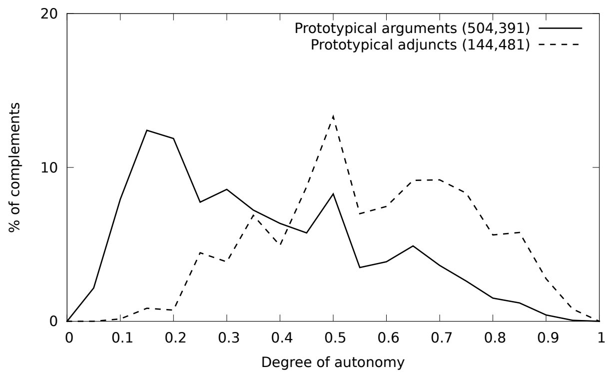

Figure 1: Complement distribution along the continuum according to the (h, c) productivity measure.

4.3 Discussion

For the results given in Table 3, the ordering criterion C1 performed the best. This can be seen as confirming the intuition expressed by [9] that arguments tend to be closer to the verb than adjuncts. This also confirms the intuition expressed by [3] that there is some natural order for verb complements.

The joint productivity criterion C2 performed very poorly. In a further analysis, we observe separately the two measures of productivity, presented in equation (3) and (4).

Figure 1 shows the distribution of the complements along the continuum that we obtained according to the (h, c) productivity measure[3].

It appears that most prototypical arguments are assigned a lower degree of autonomy than prototypical adjuncts. Thus the result complies with what we were expecting in the first place.

Figure 2 shows the distribution of the complements along the continuum that we obtained according to the (c, v) productivity measure.

It appears that most prototypical arguments are assigned a higher degree of autonomy than prototypical adjuncts thus the criterion proposed by [2] does not seem to apply to Japanese.

According to this result (c, v) pairs exhibiting a high degree of productivity tend to indicate an adjunct (and not an argument), *i.e.*, in the case of Japanese, arguments tend to belong to smaller semantic classes than adjuncts thus yielding a lower (c, v) productivity. As for implementation as an AAD criterion, it corresponds to equation (4) without the difference to 1.

[3]Due to the important number of values to plot. Values have been grouped by intervals of 0.05.

Table 3: Accuracy of the different distinction criteria with three different evaluation methods: at a lexical level with thresholding (M1), at a clause level with thresholding (M2), and continuum-based (M3).

		M1	M2	M3
Baseline	Selectional preference	43.60	13.60	35.20
	PMI	**62.66**	**39.47**	**62.68**
1st evaluation	Ordering (C1)	**79.29**	**69.68**	**81.90**
	Joint productivity (C2)	50.32	23.16	48.06
	μ(C1,C2)	73.03	56.58	78.70
2nd evaluation	Joint productivity (C3)	75.37	53.68	**88.23**
	μ(C1,C3)	**83.68**	**72.21**	87.79
	Oracle(C1,C3)	-	-	96.72

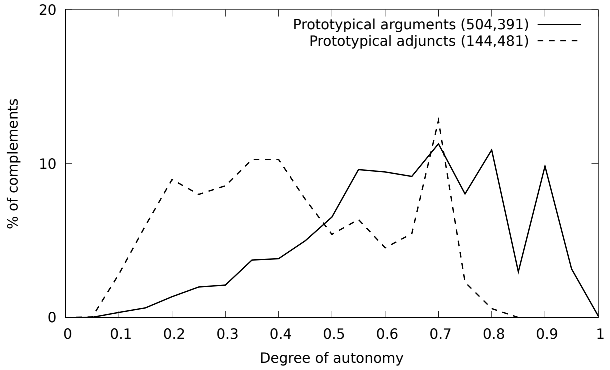

Figure 2: Complement distribution along the continuum according to the (c, v) productivity measure.

We update our joint productivity criterion with this new measure of productivity and run a new evaluation. Results are shown in Table 3 as "2nd evaluation".

It appears that the new joint productivity criterion C3 performs better than any other criterion.

However, we also observe that with both versions of joint productivity (*i.e.* C2 and C3) there is no improvement when combining them with the ordering criterion C1. Thus calculating the arithmetic mean of the degrees of autonomy does not seem to be the right way to combine criteria. We found that an optimized combination method (*i.e.* an oracle) could improve accuracy up to 96.72 percent with our continuum-based method of evaluation M3, as shown in the last line of Table 3.

5 Conclusion

The two AAD criteria we proposed in this paper have proved efficient and gave us a better understanding of argumenthood in Japanese.

In future work, we aim to apply our method to every type of complements to test whether our criteria can be generalized. The combination of different criteria remains to be studied, as we failed to combine our criteria to improve our results. Other distinction criteria should also be investigated. Finally, the model could be more accurate by using multiple predicate frame lexicons.

References

[1] Omri Abend and Ari Rappoport. Fully Unsupervised Core-Adjunct Argument Classification. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 226–236, 2010.

[2] Cécile Fabre and Cécile Frérot. Groupes prépositionnels arguments ou circonstants: vers un repérage automatique en corpus. In *TALN 2002*, pages 215–224, 2002.

[3] Naoki Fukui, Shigeru Miyagawa, and Carol Tenny. Verb Classes in English: a Case Study in the Interaction of Syntax, Morphology and Semantics. 1985.

[4] Yasunari Harada. On the Distinction between Complement and Adjunct in Japanese. In *The 6th Japanese-Korean Joint Conference on Formal Linguistics: Proceedings of the Sixth Japanese-Korean Joint Workshop*, pages 35–48, 1991.

[5] Yoko Hasegawa. A Diagnostic Test for the Complement/Adjunct Distinction in Japanese. In *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*, pages 66–77, 1988.

[6] Taku Kudo and Yuji Matsumoto. Japanese Dependency Analysis Using Cascaded Chunking. In *CoNLL 2002: Proceedings of the 6th Conference on Natural Language Learning 2002 (COLING 2002 Post-Conference Workshops)*, pages 63–69, August 2002.

[7] Kikuo Maekawa. Balanced Corpus of Contemporary Written Japanese. In *The 6th Workshop on Asian Language Resources*, pages 101–102, 2008.

[8] Christopher D. Manning. Probalistic syntax. In Rens Bod, Jennifer Hay, and Jannedy Stefanie, editors, *Probabilistic linguistics*, pages 289–341. The MIT Press, Cambridge, Massachusetts, 2003.

[9] Paola Merlo and Eva Esteve Ferrer. The Notion of Argument in Prepositional Phrase Attachment. *Computational Linguistics*, 32(3):341–377, 2006.

[10] Koichi Takeuchi, Kentaro Inui, Nao Takeuchi, and Atsushi Fujita. A Thesaurus of Predicate-Argument Structure for Japanese Verbs to Deal with Granularity of Verb Meanings. In *The 8th Workshop on Asian Language Resources*, pages 1–8, 2010.